Exploring the Effects of Threat and Error Management on Australian General Aviation

A Thesis submitted by

Seung Yong Lee, MAv

For the award of

Doctor of Philosophy

2018

Abstract

Since the emergence of Threat and Error Management (TEM), it has been well regarded as an effective method to improve aviation safety. International Civil Aviation Organisation's (ICAO) acknowledged the need for TEM training, and recommended that TEM be introduced to all pilot training syllabi. In response to ICAO's endorsement on TEM training, Civil Aviation Safety Authority (CASA) in Australia mandated TEM as an additional assessment item for various levels of flight tests and ground examinations, effective from July 2009. Although it has been more than eight years since the implementation of TEM in Australian general aviation, there are, to date, no definitive data available to suggest whether implementation of TEM training has been well received, and whether a positive effect of such training has been witnessed, experienced and translated as intended in Australian general aviation. The purpose of this thesis was to fill this knowledge gap by conducting a formal post-implementation review and assessment of TEM by conducting three separate studies.

The first study was exploratory in nature, and was the first step in examining how TEM is regarded among Australian general aviation pilots in terms of its use and effectiveness. A total of 59 general aviation pilots participated in a survey, and the results indicated a variable uptake of TEM principles, and differing opinions regarding its effectiveness. This warranted further study to be conducted in regard to TEM implementation and practice.

The second study was a qualitative (first) phase of mixed methods research approach using an exploratory sequential design. This two-phase design involved a collection of qualitative data followed by separately collecting quantitative data. The exploratory sequential design was preferred as there were a lack of definitive data and theory established after TEM was introduced in Australian general aviation. The study aimed to gather insights before and after TEM was implemented to enable appreciation as to how it was introduced and the effects of the use of TEM training. Five highly experienced flight examiners from Australian general aviation were interviewed. Based on thematic analysis, four themes arose from the interview data: impracticality, lack of support and guidance, TEM implementation and TEM in practice. The results

indicated that all participants shared very similar or the same views that TEM was not implemented well because of seemingly impractical TEM principles and a lack of guidance and support. All participants collectively agreed that the first three themes would have adversely affected the way TEM was taught and practised in Australian general aviation.

Based on the above four themes, a survey was devised to verify and test hypotheses generated from the findings in Study 2, and this was the quantitative (second) phase of mixed methods sequential design. A total of 97 survey responses were analysed, and the majority of hypotheses based on Study 2 were supported. Based on structural equation modelling, it was found that both impracticality and a lack of guidance and support adversely affected the way TEM implementation was viewed, and the lack of guidance and support was also found to adversely affect how TEM was practised.

The collective findings suggested that TEM was poorly implemented in Australian general aviation. This would likely have caused negative views on the use of TEM among general aviation pilots in Australia. However, the collective responses from Study 3 suggested that the survey participants highly valued and applied TEM principles when undertaking flying activities, and consequently considered TEM an important part of the flight preparation process. This is an encouraging finding because considering TEM principles and concepts prior to and during flying will further improve safety within general aviation in Australia.

Certification of Thesis

This Thesis is entirely the work of Seung Yong Lee except where otherwise acknowledged. The work is original and has not previously been submitted for any other award, except where acknowledged.

Principal Supervisor: Professor Paul Bates

Associate Supervisor: Hon. Professor Patrick Murray

Associate Supervisor: Dr. Tarryn Kille

Student and supervisors' signatures of endorsement are held at the University.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my principal supervisor, Professor Paul Bates, for his continued guidance and support throughout my PhD journey. Paul, you have been patiently guiding me throughout my study, as well as providing timely and thoughtful advice on my career progression to date, and I am forever indebted to you for your advice, support and encouragement.

I would also like to express my special appreciation to my associate supervisors, Hon. Professor Patrick Murray and Dr Tarryn Kille. Pat, you are like a human factors encyclopaedia, who always has answers to my questions and gave me insightful comments and encouragement throughout my study. Tarryn, thank you for patiently reviewing my chapters with constructive and valuable comments throughout this thesis.

Besides my supervisors, I would like to thank Dr Barbara Harmes, my English angel, and Dr Enamul Kabir, my statistics advisor, for their help with proofreading this thesis and providing statistical advice. Without your assistance I could not have completed this thesis. Also, this thesis was edited by Elite Editing, and editorial intervention was restricted to Standards D and E of the *Australian Standards for Editing Practice*.

Of course, this study could not be completed without the participants in the three separate studies. I thank you all for your participation and support of my study.

Last, but not the least, a special thank-you goes to my family, in particular my wife, Ria, and my two kids, Luke and Amy for their support. I am truly sorry to have been saying ‘no’ and ‘maybe later’ on numerous occasions, particularly for the last half of this year. Thank you for your patience. Let me make it up over the summer break.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANAO	Australian National Audit Office
APA	Airline Pilots Association
AQP	Advanced Qualification Program
ASA	Aviation Safety Advisor
ATO	Approved Testing Officer
ATC	Air Traffic Controller
ATPL	Air Transport Pilot Licence
ATSB	Australian Transport Safety Bureau
CAAP	Civil Aviation Advisory Publication
CAQDAS	Computer-assisted Qualitative Data Analysis Software
CASA	Civil Aviation Safety Authority
CFA	Confirmatory Factor Analysis
CFI	Chief Flight Instructor
CMAQ	Cockpit Management Attitudes Questionnaire
CP	Chief Pilot
CPL	Commercial Pilot Licence
CRM	Crew Resource Management
EASA	European Aviation Safety Agency
EFA	Exploratory Factor Analysis
FAA	Federal Aviation Administration
FE	Flight Examiner
FI	Flight Instructor
FOI	Flying Operations Inspectors
FSAQ	Flight Safety Attitudes Questionnaires
GAPAN	Guild of Air Pilots and Air Navigators
GFPT	General Flying Progress Test
IATA	International Air Transport Association
ICAO	International Civil Aviation Organisation
IFALPA	International Federation of Air Line Pilots’ Associations
IFI	Incremental Fit Index
LOFT	Line Oriented Flight Training

LOSA	Line Operations Safety Audit
MI	Modification Indices
NAA	National Aviation Authority
PDP	Professional Development Program
PM	Pilot Monitoring
PPL	Private Pilot Licence
RAA	Recreational Aviation Australia
RFDS	Royal Flying Doctor Service
RMSEA	Root Mean Square Error of Approximation
RPAS	Remotely Piloted Aircraft Systems
SEM	Structural Equation Modelling
SOP	Standard Operating Procedure
SPSS	Statistical Package for the Social Sciences
TCAS	Traffic Collision Avoidance System
TEM	Threat and Error Management
TLI	Tucker-Lewis Index
UAS	Undesired Aircraft State
US	United States
UTHFRP	University of Texas Human Factors Research Project
VFR	Visual Flight Rules
WHO	World Health Organisation

Chapter 1: Introduction

> *In the early afternoon of 27 March 1977, two Boeing 747s—Pan American World Airways flight 1736 and KLM flight 4805—diverted from their intended destination, Gran Canaria Airport, because of terrorist activity, and landed at Los Rodeo Airport on Tenerife Island. In the late afternoon of this day, KLM flight 4805 was cleared to enter Runway 12, and backtracked to the end of the runway to line up on Runway 30. After KLM's Boeing 747 entered the runway, Pan American flight 1736 was also cleared to enter Runway 12 and exit on the third taxiway to allow KLM flight 4805 to take off. With low visibility because of heavy fog and many other causal and contributing factors presented on the day, KLM flight 4805 initiated a take-off roll without clearance, while Pan American flight 1736 was still backtracking on the runway. The two Boeing 747s collided on the runway, causing the deaths of 583 people on board. This is still the world's deadliest commercial aviation disaster.*

Despite major technological advances and enhancements, this accident was one of several high-profile accidents that marked the human factors era and placed a stronger focus in aviation on human factors issues, such as the human–machine interface. Continuous efforts to improve safety initiatives and training to address these human factors issues have resulted in civil aviation to be considered an ultra-safe industry. According to Amalberti, Auroy, Berwick and Barach (2005), the rate of catastrophic accidents per flight is better than 1×10^{-6}, compared with other safety critical industries, such as surgery and road safety (both 10^{-4}). To view this from the Australian perspective, in 2010, Australia experienced 14 fatal aviation accidents, compared with 1,248 fatal accidents involving road transportation (ATSB, 2017; Pink, 2012). This impressive safety record for civil aviation is attributed to a constant drive to develop and improve safety initiatives and training. One of the recent safety initiatives was the introduction of Threat and Error Management (TEM). The Civil Aviation Safety Authority (CASA) (2008) defined TEM as "the process of detecting and responding to threats and errors to ensure that the ensuing outcome is inconsequential, i.e. the outcome is not an error, further error or undesired state" (p. 5).

TEM emerged during the development of the Line Operations Safety Audit (LOSA) and was developed to capture the full operational complexity of a flight (Merritt & Klinect, 2006). The concept has attracted strong interest from airlines, regulatory authorities and academia, and the benefits of such training have been widely witnessed (Merritt & Klinect, 2006). Consequently, the International Civil Aviation Organisation (ICAO) endorsed and recommended that TEM training be an integral part of pilot training and a licencing requirement for all pilot licences, and it has since become a requirement for initial and recurrent flight crew training (CASA, 2008; Maurino, 2005b).

In addition, anecdotal evidence suggests that pilots who continuously receive TEM training as part of their flight training from the ab initio[1] stage demonstrate earlier achievement of the required competencies, compared with pilots who receive insufficient and/or irregular TEM training (P. Murray, personal communication, July 29, 2009). This is consistent with previous studies (e.g., Bove, 2002) that demonstrated a higher achievement of post-training performance among trainees exposed to TEM.

In response to ICAO's acknowledgement of the need for TEM training and recommendation that TEM be introduced to all pilot training syllabi, CASA (2009) mandated TEM as an additional assessment item for various levels of flight tests and ground examinations, effective from 1 July 2009. It has been more than eight years since the introduction of TEM training in general aviation in Australia; thus, it is timely to conduct a formal post-implementation review and assessment of TEM in terms of the effectiveness of TEM training and its current practice in Australian general aviation.

1.1 Statement of Problem

Since the very first flight, it has been clear that there is an element of risk involved in flying, and this risk occasionally results in an incident or accident. Broadly, the risk has been considered to stem from deficiencies involving either technical or human

[1] Oxford dictionary (2018) defines 'ab initio' as starting from the beginning. This is a commonly used term to describe student pilots in an early stage of flight training.

factors. In the early days, risk was perceived to be mainly due to inadequate advancement of technology, although several studies (e.g., Feggetter, 1985; Hobbs, 2004; Murray, 1997) suggested that human factors have been the primary safety issue since the early days of aviation. Following the advancement and enhancement of technology, it has become clear that the main cause of the majority of accidents is the flight crews in the cockpit (Helmreich & Foushee, 2010).

It has been suggested that approximately 70 to 80% of aviation accidents are attributed, at least partly, to human error (Wiegmann & Shappell, 2003). As a result of the ubiquitous nature of human error, efforts have been made to prevent or eliminate pilot errors (Thomas, 2004). However, it has been acknowledged that it is unrealistic to believe that errors could be totally eliminated, given the physical and psychological limitations of humans. Adams (2006) stated that even competent humans completing a simple task continue to make errors. However, in most cases, these errors are recognised and corrected, so the results are inconsequential. Therefore, efforts have been made to identify errors and, more importantly, develop and refine training to appropriately manage those errors to avoid or mitigate negative consequences. Merritt and Klinect (2006) argued that the overarching objective of error management—later renamed TEM—provides the best possible support for pilots in managing everyday threats and errors.

Airlines around the world have devoted many resources to human factors training—particularly Crew Resource Management (CRM) and TEM, which are both endorsed and recommended by ICAO. In addition, more and more airlines of different sizes around the world are conducting LOSA to proactively determine their performance strengths and weaknesses (P. Murray, personal communication, September 25, 2018). Despite the positive effects of extensive human factors training in the airline sector, it often appears to be a neglected component in general aviation. Although there has been growing recognition of its importance in general aviation there are, to date, only a limited number of formal ongoing human factors training courses offered to general aviation, when compared with the airline environment. Sarter and Alexander (2000) suggested that, to improve overall aviation safety, the focus needs to be on the weakest links—that is, single pilots, low-technology and less rigorously trained pilots in general aviation. The lack of appropriate, formal human factors training and recurrency training offered in general aviation may explain why overall aviation safety has not significantly improved.

According to the National Transportation Safety Board (2014), general aviation accidents in 2011 accounted for approximately 95% of all aviation accidents and 94% of fatal aviation accidents in the United States (US), while only accounting for approximately half of the total industry flight time. Although the discrepancy between the commercial and general aviation sectors in Australia does not seem to be as severe as the US's statistics, the accident statistics in Australia involving general aviation indicate a similar trend. According to the Australian Transport Safety Bureau (ATSB) (2018), the numbers of fatal accidents and fatalities between 2007 and 2016 in the commercial sector were 13 and 18, respectively, while accounting for 1.6 times more hours flown than general aviation (ATSB, 2018). In contrast, the numbers of fatal accidents and fatalities during the same period within general aviation were 142 and 207, respectively (ATSB, 2018). These statistics highlight the importance of placing a stronger focus on the weakest link to improve overall aviation safety.

1.2 Purpose of the Thesis

Since the emergence of TEM, it has been well regarded as an effective method to improve aviation safety. Following ICAO's acknowledgment and recommendation of the need for TEM training for all pilot training (ICAO, 2006), CASA (2008) mandated the Day Visual Flight Rules (VFR) Syllabus on 1 March 2008 to reflect this initiative. In addition, CASA (2008) mandated from 1 July 2009 that TEM be formally assessed for the General Flying Progress Test (GFPT), Private Pilot Licence (PPL), Commercial Pilot Licence (CPL) and Air Transport Pilot Licence (ATPL) flight tests.

It has been more than nine years since TEM training was implemented in Australian general aviation. However, to date, there are no definitive data available to suggest whether implementation of TEM training has been well received, and whether a positive effect of such training has been witnessed, experienced and translated as intended in the Australian general aviation sector. This thesis aims to fill this gap by conducting a formal post-implementation review and assessment of TEM in terms of the effectiveness of TEM training and current practice in Australian general aviation. It is expected that the findings will add to the body of knowledge to better inform CASA and regulatory authorities in other countries as to how the requirement of TEM training is addressed. This will then allow National Aviation Authorities (NAA) to appropriately develop and enhance education and training materials in the area of TEM and human factors accordingly. In addition, the findings have the potential to provide information regarding pilot attitudes to safety initiatives more generally, which will be valuable in shaping future safety implementation in other growing sectors of general aviation, such as the Recreational Aviation Australia (RAA) and Remotely Piloted Aircraft Systems (RPAS) sectors.

1.3 Outline of the Thesis

This thesis comprises eight chapters, including the current chapter as the introduction. A brief description of each chapter is provided below.

Chapter 2: Literature Review—This chapter provides a literature review of three main topics: LOSA, CRM and TEM. The concept of TEM was originally conceived as a data collection framework for LOSA after it was realised that flight crews deal with threats and errors on normal everyday flights. The TEM concept was also a framework for the sixth generation of CRM that further improved conventional CRM skills as an effective means to manage everyday threats and errors.

Chapter 3: Methodology—This chapter describes four elements of the research process when developing this study. The first two elements—epistemology and theoretical perspective—are described with the three most prevalent research paradigms: post-positivism, constructivism and pragmatism. The third element—methodology—is then described that corresponds to the aforementioned research paradigms. The methodologies are quantitative, qualitative and mixed methods, respectively. The fourth element—methods—is also described. The chapter concludes by justifying why mixed methods was chosen as the preferred methodology, and exploratory sequential design was selected as the preferred method.

Chapter 4: Study 1 (Exploratory Study)—The research in this chapter served two purposes. The first was to determine whether a further post-implementation review of TEM training on a large scale was worthwhile. The second was for the researcher to select appropriate methodology and methods following the trial run. A total of 59 participants completed a short survey that aimed to collect information on the ways the requirement for TEM training was addressed, and gather views on the benefits and effectiveness of TEM. The main findings were published as a journal article in December 2016.

Chapter 5: Study 2 (Qualitative Study)—This chapter presents the qualitative (first) phase of mixed methods research approach using an exploratory sequential design. The study aimed to gather insights before and after TEM was implemented to gain an in-depth appreciation as to how it was introduced and the effects of the use of TEM training. A total of five highly experienced Flight Examiners (FE) were interviewed, and four main themes emerged from the interview data: impracticality, lack of guidance and support, TEM implementation and TEM in practice. Each theme is thoroughly discussed in this chapter. general aviation.

Chapter 6: Study 3 (Quantitative Study)—This chapter presents the quantitative (second) phase of mixed methods research approach using an exploratory sequential design. The study aimed to verify and test hypotheses generated from the findings in Study 2. To achieve this aim, a survey was devised based on the four themes, and a total of 97 responses were analysed. The combined findings from Studies 2 and 3 were expected to provide a better understanding of how TEM is implemented and its effects on the way TEM has been practised in Australian general aviation.

Chapter 7: Overall Discussion—This chapter presents a general discussion of the key findings from the three studies. The first major finding was that the perceived impracticality of TEM was a major contributor to the collective views among the study participants regarding poor TEM implementation. Lack of guidance and support from CASA was also found to have contributed to the collective views of participants on poor TEM implementation. The second finding, perhaps counterintuitively, was that TEM is highly valued and regularly used when flying activities are undertaken.

Chapter 8: Conclusion—This chapter provides a summary of key findings from the studies. The findings suggested that the effect of TEM implementation in Australian general aviation have not been experienced and translated as intended. However, despite the poor TEM implementation, it was found that TEM principles were highly valued among the study participants, which is an encouraging finding, as the ultimate aim of TEM is to further improve safety in Australian general aviation.

Chapter 2: Literature Review

TEM is described as a process of detecting and responding to threats and errors by maintaining adequate safety margins to ensure that outcomes are inconsequential (CASA, 2008). Simply stated, TEM is a required everyday exercise for pilots to fly an aircraft safely between two points (Maurino & Murray, 2009). This chapter provides a review of the literature on three topics: LOSA, CRM and TEM. The concept of TEM was originally conceived as a data collection framework for LOSA (Klinect, 2005) after it was realised that flight crews deal with threats and errors on normal everyday flights. The TEM concept was also a framework for the sixth generation of CRM that further improved conventional CRM skills as an effective means to manage everyday threats and errors (Maurino & Murray, 2010).

2.1 Line Operations Safety Audit (LOSA)

LOSA is a safety management tool, as well as a methodology for monitoring airline safety performance by collecting and analysing data on managed, mismanaged and/or unmanaged threats and errors occurring during everyday normal operations (Earl, Bates, Murray, Glendon, & Creed, 2012; Klinect, 2005). Klinect (2005) argued that, previously, although airlines continually collected various safety performance data, the majority were either reactive (e.g., incident reports) or incomplete in terms of safety performance. For instance, data gathered from quick access recorders provide information on flying parameters, such as aircraft speed and altitude, yet do not provide insights, such as pilot proficiency or flight crew harmony and behaviour (Klinect, 2005). This led to the development of a proactive and targeted cockpit observational method based on the TEM framework that 'gathers system safety and flight crew performance data during regularly scheduled flights in normal operations' (Klinect, 2005, p. 4).

There exist a number of indicators to illustrate airlines' performance in terms of safety. Some examples include incident and accident reports, collection of flight data and line checks. The first is considered reactive, as they identify system deficiencies after incidents and accidents, which is often too late. In addition, they may mislead airlines to falsely assume that their safety performance is high when there is an absence of incidents and accidents. Moreover, the use of incident and accident reports naturally poses certain limitations. These include, but not limited to, missing key information and different interpretations of causal and/or contributing factors for very similar accidents (Lee, Bates, Murray, & Martin, 2017).

In contrast, the other two indicators (i.e., collection of flight data and line checks) are considered proactive, yet do not portray a complete picture of normal operations. Flight data collected often lack context regarding the situation in which the data were collected. For instance, data may suggest a deviation from an approach profile during an instrument approach, yet do not necessarily provide context for this deviation, such as the flight crew's level of competency, fatigue level or group harmony (Klinect, 2005). In addition, these data do not explicitly show external influences, such as poor ATC instructions. Simply stated, the data show 'what', yet do not show 'how' and 'why'. Consequently, it is difficult to provide accurate and appropriate intervention.

Considering the above concerns, line checks are generally a good source of uncovering proficiency weaknesses (Klinect, 2005). However, they do not represent true normal operations, as these line checks are directly related to flight crews' employment status where failures may result in the termination of employment and, thus, subsequent behaviours, so-called 'angel behaviours', do not necessarily reflect what the crew actually do during normal operations (Klinect, 2005). These known limitations, alongside over-reliance on reactive safety data sources, necessitated the need for another innovative approach to collect reliable safety-related data (Klinect, 2005). Thus, LOSA was developed to complement other safety performance data to provide 'a unique perspective of flight crew performance' during normal flight operations (Klinect, 2005, p. 5).

The birth of LOSA can be traced back to the period after the Federal Aviation Administration (FAA) introduced the Advanced Qualification Program (AQP)—a voluntary program that allowed airlines to develop individual airline-specific training syllabi (Helmreich, Merritt, & Wilhelm, 1999). In return, airlines were required to provide both CRM and Line Oriented Flight Training (LOFT) to all flight crews, and fully incorporate CRM concepts into technical training (Helmreich et al., 1999). Delta

Air Lines first responded to the FAA's initiatives by developing a new CRM course for their flight crews in 1994; however, questions were raised with regard to its transferability as to whether the newly developed CRM course worked as intended (Klinect, 2005). Delta management was concerned that the CRM performance data from regular line checks and LOFT during flight simulator sessions would not provide line pilots' true CRM performance standards during normally scheduled flight operations (Klinect, 2005). This promoted a collaborative partnership between Delta and the University of Texas Human Factors Research Project (UTHFRP) team to develop a mechanism to better understand the actual CRM performance of Delta's flight crews (Klinect, Murray, Merritt, & Helmreich, 2003).

The first LOSA was conducted within Delta Air Lines in 1994, and involved over 450 jump seat observations on regularly scheduled flights (Klinect et al., 2003). Each observer provided a written narrative based on the phase of flight, and provided behavioural performance marker ratings, which together presented a complete picture of the flight crew's CRM performance in terms of their strengths and weaknesses (Klinect, 2005). The results from the first LOSA provided Delta Air Lines with an operational baseline of CRM strengths and weaknesses for the airline to prioritise areas to be improved for the revised CRM training (Klinect et al., 2003). The findings also assured Delta management of the reliability of CRM data (Klinect et al., 2003).

The second LOSA was conducted within Continental Airlines in 1996, and was the first LOSA that was based on the TEM framework for data collection (Klinect, 2005), The TEM-based LOSA was on the basis of notions that everyday flight crews encounter threats and commit errors that have the potential to develop into an incident or accident (Klinect et al., 2003). The LOSA recorded threats and errors, and, more importantly, indicated how these were managed by flight crews (Klinect et al., 2003). In response to the findings, Continental Airlines established operational CRM targets and implemented certain changes, such as introducing error management training for its flight crews (Klinect et al., 2003).

A follow-up LOSA in Continental Airlines was conducted in 2000, and positive outcomes were noted when compared with the results from the earlier LOSA, with a greater improvement in the area of checklist usage, a dramatic reduction in unstable approaches and an increase in overall crew performance (Gunther, 2002). Klinect (2005) highlighted that the LOSA not only provided a 'diagnostic snapshot of operational performance', but also offered benchmarks for the 'effectiveness of organisational safety changes' when it was conducted routinely (p. 33). This operational performance can also be compared with other airlines, as more and more LOSA data are gathered.

As the LOSA methodology continued to develop and mature, it was formally endorsed by ICAO (2002) as a proactive organisational strategy. Findings from LOSA could provide airlines with operational strengths and weaknesses to guide organisational strategies or tactics with regard to training and operations to achieve a higher level of safety (ICAO, 2002). Other governing organisations—such as the International Air Transport Association (IATA), International Federation of Air Line Pilots' Associations (IFALPA) and Airline Pilots Association (APA)—also endorsed LOSA as a key driver to monitor and diagnose airlines' safety performance during normal operations, and proactively develop appropriate safety interventions (Eames-Brown & Collis, 2007).

2.1.1 Line Operations Safety Audit data measures.

Each observational flight gathers four types of LOSA data based on the TEM framework: demographic, narrative, CRM behavioural markers and TEM measures (Klinect, 2005). Demographic measures refer to several flight characteristics, such as city pairs, flight time and aircraft type (Klinect, 2005). Any information that could jeopardise anonymity is not collected, such as flight numbers and date of flight (Klinect et al., 2003). Narrative measures refer to textual information provided by LOSA observers when they are unable to find a suitable code or they consider it appropriate to provide written narratives to supplement coded information (Klinect, 2005). CRM behavioural markers, as the name suggests, refer to a predetermined set of behaviours that indicate aspects of CRM performance (Flin & Martin, 2001). TEM measures refer to observed data on the flight crew's TEM performance, which encompasses three elements: threat management, error management and Undesired Aircraft State (UAS) management (Klinect, 2005). Given that this thesis's primary focus is TEM, the remainder of this section provides further elaboration on TEM measures.

The first element of TEM measures is threat management, which is divided into three variables: threat type, threat response and threat outcome (Klinect, 2005). Table 2.1 presents the three variables and their categorical levels. For LOSA observers to determine whether a threat is managed or not, the observers first need to determine which type of threat presented. The nature of threats can be expected, unexpected, latent or overt. In the University of Texas TEM framework, there are two types of threats: environmental and airline threats (Klinect, 2005). The former is outside the airline's direct control, while the latter—also referred as systemic threats—originates within the flight operations of the airline (Merritt & Klinect, 2006). There are four categories of environmental threats and seven categories of airline threats (Klinect, 2005), as presented in Table 2.1.

Table 2.1 *Primary Threat Management Variables.*

Variables	Categorical levels	
	Environmental threats	Airline threats
Threat type	• Adverse weather • Airport • ATC • Environmental operational pressure	• Aircraft malfunction • Airline operational pressure • Cabin • Dispatch/paperwork • Ground/ramp • Ground maintenance • Manuals/charts
Threat response	Was the threat discussed or planned before it was encountered? Yes/no	
Threat outcome	Inconsequential Linked to flight crew error	

Source: adapted from by Klinect (2005, p. 51).

The second variable of threat management is threat response, and this is the first part of threat management (Klinect, 2005). For LOSA observers, this is simply recorded as 'yes' or 'no' (Klinect, 2005). A 'yes' answer indicates that the flight crew either discussed or planned for a threat before it was encountered, while a 'no' answer indicates that either a threat was presented unexpectedly or the flight crew failed to plan ahead for the threat (Klinect, 2005).

The third variable of threat management, threat outcome, is the final quantitative measure of threat management (Klinect, 2005). The threat outcome is either inconsequential or linked to flight crew error (Klinect, 2005), and is depicted in Figure 2.1. Inconsequential threat outcome refers to a threat that was successfully managed or did not evolve into an error or UAS. An example of this threat is an Air Traffic Controller (ATC) failing to provide adequate lateral spacing between two aircraft (environmental threat), yet the flight crew of the following aircraft changing their course slightly to maintain the required spacing. For threat management, failure of the ATC to provide the adequate lateral spacing was not expected, but was nevertheless well managed; thus, the outcome was inconsequential. The second threat outcome is linked to flight crew error and refers to a mismanaged threat that has developed to a threat-induced error (Klinect, 2005). For example, a radio call from an ATC may interrupt the flight crew when they are performing a pre-landing checklist, and the crew consequently miss an item in the pre-landing checklist. This indicates an unexpected threat that was not managed, which escalated into the flight crew committing a threat-induced error of omitting a pre-landing checklist item.

Following threat management, the second element of TEM measures is error management, which is again divided into three variables: error types, error responses and error outcomes (Klinect, 2005). Table 2.2 displays these three variables, the error types and their categorical levels. In the TEM framework, there are three types of errors: aircraft handling errors, procedural errors and communication errors (Klinect, 2005). Aircraft handling errors refer to errors in the area of 'the flying, direction, speed and configuration of the aircraft' (Klinect, 2005, p. 52). There are five categories under aircraft handling errors, as presented in Table 2.2. Procedural errors involve deviation from rules, regulations and set procedures, such as Standard Operating Procedures (SOP) (Klinect, 2005). There are seven categories under procedural errors (Table 2.2). Communication errors concern poor communications or an absence of communications between flight crews or between the flight crew and external agents, such as the ATC, as presented in Table 2.2 (Klinect, 2005).

Table 2.2 *Primary Error Management Variables.*

Variables	Categorical levels		
	Aircraft handling errors	Procedural errors	Communication errors
Error type	• Automation • Flight control • Ground navigation • Manual handling • Systems/radio/ instruments	• Briefings • Callout • Checklist • Documentation • Pilot flying/pilot not flying duties • SOP cross-verification • Other procedural	• Crew to external • Pilot to pilot
Error response	Detected and actioned Failing to respond		
Error outcome	Inconsequential Additional error UAS		

Source: adapted from Klinect (2005, p. 57).

For LOSA observers to determine whether an error was managed or not, the LOSA observer first needs to determine which type of error was committed. The nature of errors can either be a simple slip or lapse (non-threat-induced or spontaneous error) or 'a by-product of the threat environment' (threat-induced error) (Klinect, 2005, p. 51). For instance, if a pilot simply misses a checklist item without apparent threat presented, this would be considered a non-threat-induced error. In contrast, if the same pilot misses a checklist item because a flight attendant interrupts while the flight crew is completing a particular checklist, this would be considered a threat-induced error.

The second variable of error management is error response, and this is the first part of error management (Klinect, 2005). In LOSA, there are two types of responses: (i) detected and actioned and (ii) failing to respond (Klinect, 2005). The former refers to an error that was committed, yet actively managed by the flight crew. For example, a flight crew is instructed by the ATC to climb to 25,000 feet, but the Pilot Monitoring (PM) incorrectly reads this back as 24,000 feet. If the PM then immediately realises his or her incorrect read-back and subsequently reads back the correct altitude of

25,000 feet, this would be considered an error response being detected and appropriately actioned. The other type of error response is failing to respond, and this involves the flight crew either failing to detect or ignoring an error, thereby leaving the error unmanaged (Klinect, 2005). For example, if the PM incorrectly reads back the newly assigned altitude and there is no subsequent correction, this would be considered an error response of failing to respond.

The third variable of error management, error outcome, is the final piece of information to construct overall error management (Klinect, 2005). The three types of outcomes are inconsequential, additional error and UAS, as depicted in Figure 2.1. Similar to an inconsequential threat, an inconsequential error refers to an error that does not develop into the flight crew committing additional errors or UAS. For instance, using the same example above, the PM incorrectly reads back the newly assigned altitude, and neither the Pilot Flying (PF) nor the PM realise their read-back error. However, the ATC corrects their read-back error and the PM subsequently reads back the assigned altitude correctly. This is considered an inconsequential error. The second outcome is committing an additional error and causing an error chain to develop. For instance, the flight crew incorrectly reads back the new assigned altitude and this error is not noticed by the flight crew or corrected by the ATC. The final outcome is an error or a series of errors developing into UAS, which is defined as an undesired 'crew-error-induced aircraft state' in which the required safety margins are compromised (Klinect, 2005, p. 58). Using the same example above, the incorrect read-back of an assigned altitude and subsequent setting of incorrect altitude on the mode control panel will result in the aircraft levelling off at the incorrect altitude, which will reduce safety margins, such as reduction in vertical and/or lateral separation from other aircraft in the vicinity.

Following error management, the third element of TEM measures is UAS management, which is again divided into three variables: UAS types, UAS responses and UAS outcomes (Klinect, 2005). Table 2.3 presents the three variables and their categorical levels. As per the definition of UAS in the earlier paragraph and Figure 2.1, UAS management is the final opportunity for the flight crew to return the aircraft to its intended and/or optimal state, so that the event does not escalate to an incident or accident.

Table 2.3 *Primary UAS Management Variables.*

Variables	Categorical levels
USA type	Aircraft handling Ground navigation Incorrect aircraft configurations
UAS response	Detected and actioned Failing to respond
UAS outcome	Inconsequential Additional error

Source: adapted from Klinect (2005, p. 60).

For LOSA observers to determine whether UAS was managed, the LOSA observers first need to determine the types of UAS. An example of aircraft handling UAS includes unstable approach (Klinect, 2005). An example of ground navigation UAS includes taxiing above the speed limit (Klinect, 2005). An example of incorrect aircraft configurations UAS includes incorrect altitude selection on the mode control panel (Klinect, 2005).

The second variable of UAS management is UAS response, and this is the first part of UAS management (Klinect, 2005). Similar to the TEM described above, there are two types of responses: (i) detected and actioned and (ii) failing to respond (Klinect, 2005). The former refers to the detection of UAS and acting appropriately to manage UAS. An example of the former response (detected and actioned) is as follows. After the flight crew levels off at an incorrect altitude, the crew realises their error and takes appropriate action, such as requesting to climb/descend to the correct altitude. The second type of response—failing to respond—involves the failure of the flight crew to return the aircraft to its intended and/or optimal state, thereby resulting in safety margins being compromised. An example of the latter response (failing to respond) is the same as the previous example, except the flight crew does not realise their error and no further action is taken until a Traffic Collision Avoidance System (TCAS) warning alerts the flight crew.

The final variable of UAS management, UAS outcome, is either inconsequential or linked to additional error (Klinect, 2005). The earlier example of levelling off at the incorrect altitude, yet taking appropriate action, such as requesting to climb/descend to the correct altitude, would be considered inconsequential if this vertical deviation did not escalate into an incident or accident. An example of the second UAS outcome (linked to additional error) is as follows. After an aircraft deviates from its intended vertical profile (either too high or too low) during an instrument approach, the flight crew fixates on regaining the correct vertical profile and does not notice that the aircraft has deviated from its intended lateral profile (either left or right of its intended track).

Collecting reliable data, such as TEM measures, is an essential aspect of LOSA. For LOSA to be successful and to collect quality data, it involves 10 core operating characteristics, each of which has equal importance (Klinect, 2005; Klinect et al., 2003). These characteristics are described in the next section.

2.1.2 Ten Line Operations Safety Audit operating characteristics.

Klinect (2005) described that, in response to years of field work experience and continued research in the observational research design, LOSA was developed in its current form with 10 operating characteristics that define the UTHFRP LOSA. These 10 operating characteristics are as follows (Klinect, 2005, pp. 34–35):

1. jump seat observations of regularly scheduled flights
2. voluntary flight crew participation
3. anonymous, confidential and non-punitive data collection
4. joint management/union sponsorship
5. secure data collection repository
6. trusted and trained observers
7. systematic observation instrument
8. data verification roundtables
9. data-driven targets for enhancement
10. feedback of results to line pilots.

To ensure that the data collected are of high quality, all LOSA observations are performed during regularly scheduled flights, which is the first operating characteristic of LOSA (Klinect, 2005). Formal line checks and evaluation flights are off limits because they will not exhibit the true behaviour of flight crews (Klinect et al., 2003).

To ensure the success of LOSA, the project needs to gain pilot trust. The more trust LOSA gains from its participating flight crews, the more likely the LOSA data will reflect the flight crews' normal behaviours (Klinect et al., 2003). The second, third, fourth and sixth operating characteristics are essential to increase the level of pilot trust. Voluntary participation reinforces that LOSA is not another form of evaluation flight (Klinect et al., 2003). Anonymous, confidential and non-punitive data collection ensures that no potentially identifying information—such as flight numbers and dates—are collected, and observers do not discuss their observations in public or private (Klinect, 2005). A formal agreement between airline management and the pilot's association further strengthens pilot trust in LOSA (Klinect et al., 2003).

Another essential characteristic to achieve successful LOSA and gain pilot trust is to appoint trusted and well-trained observers who are typically 'unobtrusive and non-threatening' (Klinect et al., 2003, p. 666). The majority of observers recruited are regular line pilots in the airline, with a high level of operational familiarity and technical expertise. However, the team of LOSA observers can also include other individuals, such as retired line pilots or external observers (Klinect 2005; Klinect et al., 2003). Potential observers are identified separately by the airline management and pilots' association, and individuals who appear on both lists are approached to participate as observers (Klinect et al., 2003).

The use of regular line pilots as observers can potentially result in collecting subjective data, such as 'what should have been done' instead of 'what was done' (Klinect, 2005). This will clearly affect the validity and reliability of the LOSA findings. Therefore, it is important to establish a systematic observation instrument that minimises the opportunity for subjective biases to emerge within the LOSA data (Klinect, 2005). Additionally, there needs to be a trusted data collection point, either jointly managed by representatives from both management and the pilots' association, or use of an independent third party, as well as a secure data repository to assure confidentiality (Klinect, 2005; Klinect et al., 2003).

After collecting all the LOSA data, a joint data verification roundtable is convened to check the consistency and accuracy of the data before analysing them (Klinect, 2005). The roundtable typically consists of three to five representatives from various sections of the airline (e.g., a fleet manager, pilot representative and so forth) who scan the LOSA data for any errors or inconsistencies (Klinect, 2005). After the roundtable session, the data are checked for consistency with the airline's SOPs (Klinect et al., 2003). During data analysis, specific patterns emerge, such as frequent errors during a particular phase of flight. These patterns serve as safety targets that the airline will work towards, as well as a benchmark for follow-up LOSAs, so that the subsequent findings can be measured against safety targets to assess whether relevant intervention plans and strategies implemented were successful (Klinect, 2005). The patterns, safety targets and relevant intervention plans and strategies must be communicated in a timely manner to relevant stakeholders, such as flight crew within the airline (Klinect, 2005).

LOSA has been well received as a proactive methodology to collect safety data that provide a snapshot of the strengths and weaknesses of an airline and its operations. For LOSA to be successful, data are collected from normal operations after gaining pilot trust by ensuring that the data collection is voluntary and non-punitive. The quality, validity and reliability of the data are further enhanced by certain characteristics, such as data verification roundtables. The findings provide areas of safety concerns on which to focus and develop organisational strategies or tactics to further improve the safety performance of the airline. In addition, LOSA data from an airline can be compared with de-identified findings from other airlines that already undertook LOSA to assess the airline's performance against other airlines. These characteristics and the proactive nature of LOSA were endorsed by ICAO and other international governing organisations, such as IATA, IFALPA and APA.

2.2 Crew Resource Management

The birth of CRM can be traced back to a 1979 workshop, titled 'Resource Management (RM) on the Flight Deck', which addressed the fact that the majority of accidents resulted from so-called pilot error, and primarily involved failures of interpersonal communication, decision-making and leadership (Helmreich et al., 1999). The workshop also challenged the long-held belief that 'pilot proficiency plus aircraft reliability equals to safe flight' (Maurino & Murray, 2009, pp. 10–13). Since then, CRM has evolved through different generations with varying foci to reach its current state. Maurino and Murray (2009) argued that it is undisputed that CRM is a significant contributor to the safety and efficiency of the aviation system. The following provides a brief description of each generation of CRM and the changes that occurred over time.

2.2.1 Generations of crew resource management.

The first generation of CRM mainly focused on pilot personality and its influence on effective crew coordination (Klinect, 2005). The topics covered in the first generation of CRM were predominantly psychological in nature (Helmreich et al., 1999), with a strong focus on individual behaviour, style, communication, managerial effectiveness and psychological testing (Maurino & Murray, 2009; Salas, Burke, Bowers, & Wilson, 2001). The underlying safety paradigm was that the level of safety was directly and exclusively related to the performance of the flight crew (Maurino & Murray, 2009). Hence, the first generation of CRM primarily focused on changing individual styles, and identifying and correcting deficiencies in the flight crew's behaviour—commonly called the 'wrong stuff' (Salas et al., 2001). The primary focus on individuals with the 'wrong stuff' caused resistance from some pilots who deemed this approach to be an attempt to manipulate their personalities. As a result, a revised version was presented as the second generation of CRM (Helmreich et al., 1999), which shifted from focusing on personality to training pilots by offering examples of desirable crew behaviours and relating them to the daily operational contexts (Klinect, 2005).

One distinct change in the second generation of CRM was a change of the first letter 'C' from meaning 'cockpit' to meaning 'crew', as it was recognised that other aviation personnel also required such skills, including cabin crew (Helmreich et al., 1999). The refinement process included, but was not limited to, greater use of specific aviation concepts and a stronger team orientation (Helmreich et al., 1999). However, similar to the first generation, the second generation of CRM maintained two separate dichotomies: technical training and non-technical (e.g., CRM) training (Maurino & Murray, 2009). The level of acceptance of the second generation of CRM greatly improved compared with the first generation of CRM, and this, along with the introduction of glass cockpits, led to the third generation of CRM (Klinect, 2005).

The commencement of the third generation of CRM coincided with the introduction of glass cockpits, and some of the highlights of this generation included an attempt to integrate CRM into technical training and to broaden its audience further to include other groups, such as dispatchers and maintenance personnel (Helmreich et al., 1999). The third generation also considered systemic influences (e.g., organisational culture and regulatory influence) and their effect on flight crews' performance, although it was recognised that these influences were external and beyond flight crews' control (Klinect, 2005). In addition, the third generation 'revisited the human–machine interface and introduced the concepts of mental models, stress and fatigue management, automation management, vigilance and human reliability' (Koeppen, 2012, p. 36). The attempt to fully integrate CRM into technical training led to the next generation of CRM.

The fourth generation of CRM in the 1990s encompassed a full integration of CRM into technical training, with the aim of solving the problems of human errors (Helmreich et al., 1999), as well as focusing on cultural issues, with particular attention to issues involving multinational crews (Maurino & Murray, 2009). In addition, in this generation of CRM, FAA (1991) introduced the Advanced Qualification Program (AQP), which allowed airlines to tailor their fully integrated CRM and technical training, as appropriate, to fit each airline's specific requirements (Salas et al., 2001). However, there were rising concerns that the key objective of CRM had been diluted because of the too-broad strategic concepts of safer flight and ways to achieve it. Hence, the fifth generation of CRM was presented (Maurino & Murray, 2009) after a re-examination of the original CRM principles discussed at the 1979 National Aeronautics and Space Administration (NASA) workshop (Klinect, 2005) and the introduction of LOSA (Koeppen, 2012).

The fifth generation of CRM saw a shift from training for error prevention to error management, after it was realised and accepted that error was a ubiquitous and natural part of every flight and could not be totally eliminated (Thomas, 2004). The fifth generation also marked 'back to basics' by narrowing the goals of CRM to specifically consider the goal that CRM was fundamentally intended to achieve on a tactical level—error management (Maurino & Murray, 2009). This was a result of two major findings from LOSA. The first was that human error is inevitable, and the second was that flight crews employ specific countermeasures to fly safely from Point A to B (Maurino & Murray, 2009). To achieve its aim, the focus of the fifth generation of CRM was improving teamwork skills that would promote: '(a) error avoidance, (b) early detection of errors, and (c) minimisation of consequences resulting from CRM related errors' (Salas et al., 2001, p. 642).

Continued research on LOSA and CRM suggested that the variability of the operational environment, which was beyond flight crews' control, had a significant effect on the way error management was exercised (Maurino & Murray, 2009). Koeppen (2012) added that, although potentially adverse environments were outside flight crews' control, the situations or events still required flight crews to manage them. Thus, error management was developed into TEM, which became the framework for the current generation of CRM. CRM training has become a core type of human factors training in airlines and other communities, such as medical and offshore oil production sectors, which have increasingly adopted and modified CRM principles and concepts in their activities (Flin, O'Connor & Crichton, 2008). This trend highlighted the need for a systematic review of CRM's effectiveness, which is described in the following section.

2.2.2 Crew resource management effectiveness.

Salas et al. (2001) conducted a systematic review of 58 published articles on CRM training, using Kirkpatrick's typology (Kirkpatrick, 1994), to determine CRM training's effectiveness in the aviation industry. O'Connor et al. (2008) acknowledged the utility of the framework to assess the effects of post-training intervention in an organisation, as it considers multiple levels. Kirkpatrick's evaluation approach consists of four levels of evaluation: (i) participants' reaction to a newly implemented program or training initiatives (reaction), (ii) changes in participants' attitude and improved skills and knowledge (learning), (iii) changes in behaviour (behaviour) and (iv) achievement of the predetermined final objectives of a program (results) (Kirkpatrick & Kirkpatrick, 2006).

The first level, reaction, involves gauging a level of satisfaction from, for instance, participants in a new training initiative to assess the effectiveness of the new initiative and the areas for improvement (Kirkpatrick & Kirkpatrick, 2006). The satisfaction level results from participants' different perceptions of a program, such as whether it was worthwhile attending, relevant to their job and interesting to hear (Salas et al., 2001). The review found that a total of 27 studies (46%) collected reaction data with the main focus of affective feelings towards CRM programs (12 of 27), the usefulness of these programs (nine of 27) and a combination of both (seven of 27) (Salas et al., 2001). The review suggested that CRM training was found to produce positive reactions, which was an encouraging finding because this would likely improve the credibility and necessity of CRM training and enhance trainees' motivation to learn (Salas et al., 2001).

The second level, learning, involves measuring the learning outcomes imparted, such as improvement in participants' knowledge and skills and participants' desired attitude change (Kirkpatrick & Kirkpatrick, 2006). Some of the outcomes can be assessed easily and often immediately after training (e.g., improvement in knowledge), whereas others are more gradual (e.g., changes in attitude) (Kirkpatrick & Kirkpatrick, 2006). The review found that a total of 30 studies (52%) collected learning data, and the majority of studies assessed changes in attitude after CRM training using the Cockpit Management Attitudes Questionnaire (CMAQ) or a modified version of it (Salas et al., 2001). The overall findings indicated that CRM training produced positive changes to participants' knowledge and targeted attitudes (Salas et al., 2001). The findings thus far have been positive and encouraging, yet further evaluation is needed to assess whether positive reaction and learning from CRM training result in changes to participants' desirable behaviour on the job.

The third level, behaviour, is described as the assessment of changes in behaviour (e.g., knowledge, skills and attitude) after learning has occurred (Kirkpatrick & Kirkpatrick, 2006). The review found that a total of 32 studies (55%) collected some form of behavioural data using the measurements of CRM-related behaviour during a training session (e.g., LOFT), online assessment or combination of both (Salas et al., 2001). The overall findings suggested that CRM training had a positive effect on crews' CRM performance, such as improving decision-making, enabling greater adaptability and improving team communication (Salas et al., 2001).

The fourth level, results, involves determining the achievement of the predetermined final objectives of a program, such as CRM training. The results are assessed in several ways, such as considering improved quality, reduced accidents and incidents, and higher profits (Kirkpatrick & Kirkpatrick, 2006). The review found that only six studies (10%) collected data at this level (Salas et al., 2001) which is understandable given the difficulties in terms of time and resources. The rare occurrences of incidents and accidents for gathering definitive, and often longitudinal, data to determine how successfully a program (e.g., CRM training) has achieved its predetermined final objectives (Salas et al., 2001). The six studies and their findings based on anecdotal evidence (e.g., accident/incident reports) and longitudinal studies (e.g., trends in quarterly air carrier discrepancy reports) weakly suggested that improved CRM behaviours would reduce the effects of human and mechanical errors on the level of an organisation's safety (Salas et al., 2001).

The review of 58 published articles on CRM training—using Kirkpatrick's typology (Kirkpatrick, 1994) to determine CRM training's effectiveness in the aviation industry—indicated that, at each level, the findings were positive, as the CRM training was found to produce positive reactions by the participants, which led to positive changes in their attitude and behaviour in the desired direction (Salas et al., 2001). Salas, Wilson, Burke, and Wightman (2006) conducted another review of 28 published articles both within the aviation industry and in other industries, such as medicine and offshore oil production and maintenance, as CRM can be conceptualised as a team training instructional strategy to improve coordination and performance among team members. The following describes the findings from this updated review.

As stated above, the updated review (Salas et al., 2006) identified 28 published articles within and outside the aviation industry, and the breakdown of the domains was as follows: commercial aviation (two studies), military aviation (seven studies), air traffic control (one study), aviation maintenance (three studies), medicine (11 studies), offshore oil production and maintenance (one study), shipping/maritime navigators (two studies) and nuclear (one study). The updated review also used Kirkpatrick's typology (Kirkpatrick, 1994) to determine the effectiveness of CRM within the aforementioned domains (Salas et al., 2006).

A total of 13 studies (46%) evaluated the effectiveness of CRM training at the lowest level of Kirkpatrick's evaluation approach—reactions—mostly using self-report surveys, such as CMAQ or a modified version (Salas et al., 2006). The findings indicated that, similar to the earlier review, the trainees found the training enjoyable and useful (Salas et al., 2006). Similar findings were noted in a study that performed a meta-analysis using 16 published CRM evaluation research outputs (O'Connor et al., 2008). These were considered encouraging findings, as this would likely improve CRM training's credibility and enhance trainees' motivation to learn (Salas et al., 2001). Indeed, O'Connor, O'Dea and Keogh (2013) found that participants in healthcare responded positively to CRM-type training which led to a large effect on the participants' knowledge.

In terms of learning—the second level of Kirkpatrick's evaluation approach (Kirkpatrick, 1994)—12 studies (43%) evaluated the effectiveness of CRM training at this level, and there was evidence of learning occurring in all domains with varying degrees (Salas et al., 2006). There were mixed findings that indicated positive transfer, where positive reactions resulted in positive learning and desirable attitudes, yet there were also findings that suggested a neutral result, where scores on knowledge and attitude were similar between trained and untrained groups (Salas et al., 2006). Ford, Henderson, and O'Hare (2014) conducted a similar study using Flight Safety Attitudes Questionnaires (FSAQ) to evaluate the effectiveness of newly developed CRM training for flight attendants at an Asia-Pacific airline. The findings indicated a significant improvement in understanding job roles and perceptions of the importance of team work and communication, which led to greater attitudinal change, particularly for senior and short-haul cabin crews (Ford et al., 2014). These findings further supported the importance of CRM training within and, most likely, outside the aviation industry.

Learning and subsequent positive changes in attitude likely lead to a change in behaviour in the intended direction, which is the third level of Kirkpatrick's evaluation approach (Kirkpatrick, 1994). A total of 16 studies (57%) examined a transfer of learning to the desirable behaviour after CRM training, and, unlike the earlier review (Salas et al., 2001), the updated review noted inconsistent transfer of behaviour, with some studies identifying partial and negative transfer, and others noting positive transfer (Salas et al., 2006). Salas et al. (2006) added that, even with an initial positive transfer of behaviour, there was a tendency for the behaviour to regress to the pre-training state. This tendency highlights the importance of conducting such training on a regular basis.

The fourth level of Kirkpatrick's evaluation approach (Kirkpatrick, 1994) is results, and, similar to the earlier review, the number of studies that evaluated the effectiveness of CRM training at this level was small (five studies), with conflicting findings regarding achievement of the predetermined final objects of the CRM training (Salas et al., 2006). Three studies' findings indicated a positive effect on safety, based on the reduced number of errors and incidents, yet the findings from two other studies were neutral (Salas et al., 2006). Understandably, with limited definitive data, there was weak indication that CRM training had a positive effect on improving safety. This requires continued research on the effectiveness of CRM training and its evaluation, as the adoption and use of CRM grows in various industries.

The two reviews (Salas et al., 2001; Salas et al., 2006) suggested that, despite the variability of the findings, CRM training generally appears to be effective and has a positive effect in a number of high-reliability industries, both within and outside the aviation industry. Therefore, the following section describes some of the high-reliability industries that have adopted the principles and concepts of CRM training, and their experiences.

2.2.3 Crew resource management in high-reliability industries.

O'Connor et al. (2008) proposed that the aviation industry has been playing a pivotal role in developing training programs that aim to ultimately make the industry safe. One such training program is CRM, which was originally confined to flight crews in the cockpit, yet quickly moved to other parts of the aviation industry through its evolution. With continued research, the concepts of CRM and training were extended beyond the aviation industry into other high-reliability industries, such as air medical, healthcare and nuclear power plant industries. Fisher, Phillips, and Mather (2000) acknowledged that appropriately instituted CRM was expected to promote safe and efficient operations within these industries. This section describes some of the research on CRM implementation and evaluation in these industries.

The primary role of the air medical industry is using aircraft to transport patients from one place to another, such as an accident scene to a hospital. Fisher et al. (2000) administered a survey asking whether each participant from the air medical teams received three identified areas of training (i.e., CRM training, team building exercises and effective communication training). Approximately half of the participants responded that they had received all three types of training. The findings indicated that the participants who completed all three types of training exhibited attitudes and behavioural patterns that would promote and encourage open communication and safer operations within the air medical environment (Fisher et al., 2000). These findings were consistent with findings (e.g., Salas et al., 2001) in the aviation industry after CRM training was introduced.

Continued research and a growing body of evidence on the benefits of CRM resulted in the concepts and principles of CRM extending to the healthcare industry, such as the operating room (Wakeman & Langham, 2018). The operating room is a place where a group of professionals with different backgrounds and skillsets work in a coordinated and harmonised manner to deliver optimum patient care (Wakeman & Langham, 2018). Therefore, it is important for each medical professional to understand their duties and perform them accordingly. One of the ways to improve this area is through preoperative briefing. According to Wakeman and Langham (2018), teams in the operating room were found to perform significantly more briefings, surgical pauses and debriefings after CRM training, which led to greatly improved communication among medical professionals, and a reduction in annual procedural mortality and all-cause morbidity.

Introduction of CRM principles and a training program was also observed in the nuclear power plant industry. A CRM training program with a focus on the area of non-technical skills—such as leadership, teamwork and communication—was developed because it was identified that team performance was one of the key factors to promote safety in the digitalised control room of a nuclear power plant (Kim & Byun, 2011). The main findings indicated that individual attitude and team performance improved after the CRM training program was implemented (Kim & Byun, 2011). In addition, the CRM training program was found to be more effective than other similar programs that aimed to enhance team skills in the field of crew coordination and communication (Kim & Byun, 2011).

Maurino and Murray (2009) argued that CRM was one of the significant contributors to the safety and efficiency of the aviation system. Continued research activity and the success of CRM have resulted in the principles and concepts of CRM being extended to other high-reliability industries, such as the healthcare and nuclear power plant industries. Maurino and Murray (2009) added that the evolution of CRM led to varying foci within each CRM generation, and led to its current sixth generation, which includes TEM as its framework.

2.3 Threat and Error Management (TEM)

TEM is described as a process of detecting and responding to threats and errors to ensure that the outcome of an event is inconsequential by maintaining adequate safety margins (CASA, 2008). Simply stated, TEM is a required day-to-day exercise for pilots to fly an aircraft safely between two points (Maurino & Murray, 2009). TEM emerged during the development of LOSA, with the aim of capturing the full operational complexity during flights (Merritt & Klinect, 2006). The following sections describe the theoretical framework of TEM in more detail.

2.3.1 Introduction of TEM in Australian aviation.

TEM is endorsed by ICAO as an integral part of pilot training at all licence levels, and has been generally accepted in the airline industry as an effective method of improving flight safety (CASA, 2008). In addition, anecdotal evidence has suggested that pilots who continuously receive TEM training in a structured manner as part of their flight training from the ab initio stage demonstrate earlier achievement of required competencies, compared with pilots who receive insufficient and/or irregular TEM training (P. Murray, personal communication, July 29, 2009). This is consistent with previous studies that demonstrated a higher achievement of post-training performance among trainees exposed to TEM (Bove, 2002). In response to ICAO's acknowledgement of the need for TEM training and recommendation that TEM be introduced to all pilot training syllabi, CASA (2009) mandated TEM as an additional assessment item for various levels of flight tests and ground examinations, effective from 1 July 2009.

2.3.2 Threat and error management model.

Maurino (2005a) described a TEM model as the theoretical foundation that 'assists in understanding, from an operational perspective, the inter-relationship between safety and human performance in dynamic and challenging operational contexts' (p. 1). CASA (2008) stated that TEM involves more than the traditional role of airmanship, as it encourages pilots to be more proactive in identifying and managing threats and errors, even before the flight commences. The model consists of three major components: threat, error and Undesired Aircraft State (UAS).

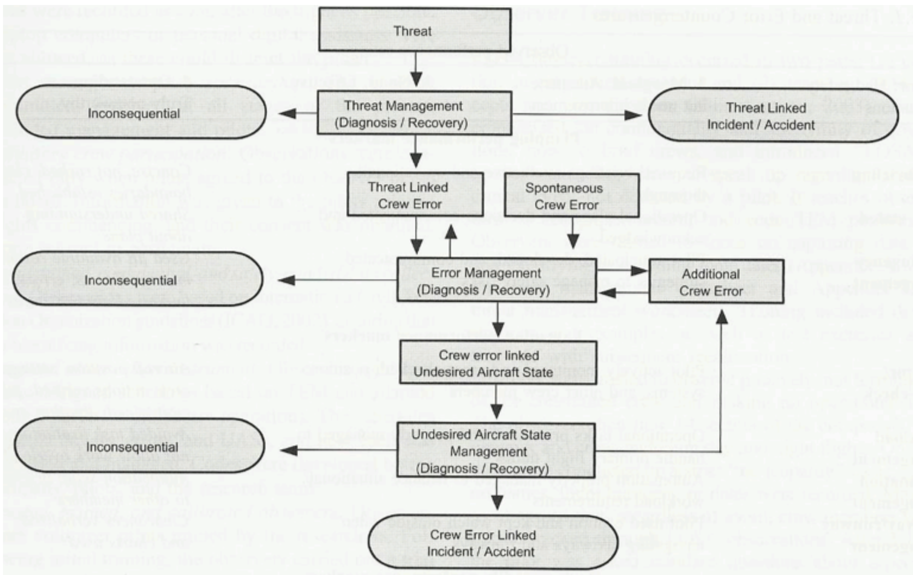

Figure 2.1. TEM model. Reproduced with permission from the LOSA Collaborative.

A threat is defined as an event that occurs outside the influence of the flight crew, and consequently increases the operational complexity of a flight (Merritt & Klinect, 2006). Hence, the crew's attention to and management of the threat is required so that the safety margins can be maintained (Merritt & Klinect, 2006). Two types of threats are environmental threats and airline threats. The former is outside the airline's direct control, while the latter—also referred as systemic threats—originates within the flight operations of the airline (Merritt & Klinect, 2006). There are four categories of environmental threats and seven categories of airline threats (Klinect, 2005). These categories and their examples are presented in Appendix 1.

The nature of threats can be classified as overt or latent (Maurino & Murray, 2009). Overt threats are easily observable and identifiable external events that are either expected or unexpected, while latent threats are not readily observable and are frequently not easily fixed at the crew level (Maurino & Murray, 2009). Examples of overt threats include adverse weather (environmental threats) and aircraft system malfunctions (airline threats). Some adverse weather conditions can be expected based on weather forecasts, such as thunderstorms and icing. In contrast, sudden engine failures or other system failures may be unexpected, yet flight crews are nevertheless required to manage these issues to maintain the required level of safety. Examples of latent threats include poor ATC training (environmental threats) and an inadequate and/or inappropriate training manual (airline threats). Understandably, complete avoidance of these types of threats is almost impossible. Thus, proper management of the threats is needed to maintain the required safety margins; otherwise, a mismanaged threat has the potential to develop into an error or result in a threat-induced accident or incident. According to Klinect (2016), approximately 25% of threats develop into an error with greater mismanagement of threats (46%) when threats are not anticipated. One of the ways to minimise the occurrence of mismanagement of threats is to ensure good pre-flight briefings, planning and active system monitoring.

The second major component of the TEM model is error. Errors are defined as flight crews' actions or inactions that lead to a deviation from crew or organisational intentions or expectations, and consequently reduce the required safety margins (Merritt & Klinect, 2006). This creates the potential for adverse operational events to develop while the flight crew are on duty (Merritt & Klinect, 2006). As displayed in Figure 2.1, the origin of errors can be either spontaneous (non-threat induced) or threat induced. In the TEM framework, errors are categorised based on the primary interaction of the flight crew at the time an error is committed (ICAO, 2005b). There are three types of errors: aircraft handling errors, procedural errors and communication errors. The types of errors and their examples are presented in Appendix 2. According to Klinect (2016), the highest mismanagement of errors involves aircraft handling errors. Any mismanaged errors can develop into a UAS or additional error, thereby creating an error chain. Thus, proper management of errors is essential to maintain the required safety margins.

According to Klinect (2005), error management focuses on the environmental influences and psychological mechanisms involved after an error is committed. There are three phases involved in the error management process after an error is committed: detection, localisation and correction (van der Schaaf, 1995). The first phase, detection, is the most important phase of the error management process because undetected errors cannot be corrected and managed. The second phase, localisation, essentially involves understanding the error, including its origin. The final stage, correction, refers to planning and devising timely and effective counter-actions to return to a stable status (van der Schaaf, 1995) or at the very least prevent the error from causing further complications (Klinect, 2005).

The third major component of the TEM model is UAS, which is defined as the safety-compromising state of an aircraft that has deviated from its intended and/or optimal state because of mismanagement of the flight crew's error, leading to a consequent reduction of the required safety margins (Merritt & Klinect, 2006). ICAO (2005b) added that this is a transitional state between a normal operational state and an outcome (e.g., incident or accident). This is the result of ineffective management of threat and/or error, and is the last opportunity for the flight crew to return the aircraft to its intended and/or optimal state, so that the event is inconsequential. Types of UAS and their examples are presented in Appendix 3.

Thus far, the three major components of the TEM model have been described. Now, this section moves on to describe three accidents involving aircraft in Australian general aviation, as well as a brief analysis using the TEM model. The first accident occurred as follows:

On 29 September 2011 at 1240 Central Standard Time, a Gippsland Aeronautics GA-8 Airvan, registered VH-AJZ, departed Marree, South Australia with one pilot and six passengers for a scenic charter flight over Lake Eyre and surrounding regions. About 45 minutes after takeoff and while flying at a height of about 500 ft above ground level the pilot felt a shudder through the airframe, then heard a loud pop and the propeller stopped. The pilot carried out a successful forced landing on the Birdsville Track, approximately 135 km north-north-east of Marree. (ATSB, 2013)

In this example, based on the TEM model, an airline threat (aircraft malfunction) was present, and the threat was correctly managed by successfully performing a forced landing without power. The second accident occurred as follows:

On 29 September 2012, a Cessna P206B aircraft, registered VH-EGG, departed Gympie on a private flight to Monduran, Queensland. On board the aircraft were the pilot and four passengers. On arrival at the Monduran aeroplane landing area (ALA), the pilot noted the windsock was indicating gusty wind conditions, from about 310–320°, and elected to land on runway 02. When on the base and final legs of the circuit, the pilot reported that wind gusts in excess of 20 kts were experienced, along with moderate to severe turbulence. During the landing, at about 10 ft above the runway, the flare was commenced. Immediately after, a significant wind gust was experienced, resulting in a hard landing on the main landing gear. The aircraft bounced and the pilot applied a small amount of power in an attempt to regain control. A second wind gust of greater intensity then occurred. The aircraft stalled and touched down hard, collapsing the nose landing gear. The pilot maintained directional control and the aircraft came to a stop. The aircraft sustained damage to the propeller, nose landing gear and lower engine cowls. (ATSB, 2012)

In this example, based on the TEM model, an environmental threat (gust of wind in turbulent conditions) was clearly present in the accident report, and the threat was not adequately managed. A threat-induced error (aircraft handling error) followed, but was not correctly managed; hence, a UAS occurred, where the aircraft stalled, and the outcome was collapsed nose landing gear. The third accident occurred as follows:

> *On 7 November 2012, at about 1000 Eastern Standard Time a Cessna 172N registered VH-JGR (JGR) departed Archerfield Airport, Queensland on a training flight. The purpose of the flight was to conduct solo pre-test revision, prior to the pilot performing the Private Pilot Licence (PPL) flight test. The aircraft was booked for 2 hours. During the flight, the aircraft impacted terrain and was substantially damaged. The pilot reported that he had very little memory of the flight, but did recall that he intended to fly to the southern training area, to practice holding heading and altitude for his upcoming flight test. The pilot stated that his usual practise was to track via Jimboomba, Beaudesert and Boonah before returning to Archerfield via Jimboomba. The pilot had not submitted a flight plan or left a flight note with a responsible person or lodged a Search and Rescue Time (SARTIME) with the Airservices Australia. Following the accident, the pilot recalled regaining consciousness and crawling to JGR to broadcast a distress call on the aircraft radio. The aircraft was fitted with a personal locator beacon (PLB) but the pilot was unable to locate it after the accident to activate it. At about 1330, an aircraft in the area reported hearing two mayday calls on the Brisbane Centre Frequency, the calls were very faint and not heard by Brisbane Centre. The area controller requested the pilot of another aircraft in the area to track south from Kagaru to investigate. At about 1410, a Cessna 172 was sighted in a paddock in uneven tussock strewn country about 2.5 Km south of Kagaru aeroplane landing area (ALA) on the runway heading. The pilot was the only person on board and suffered severe injuries as a result of the accident and at about 1500, he was airlifted to hospital. (ATSB, 2013)*

In this example, based on the TEM model, there was no clear evidence of a threat present; however, it is likely that a spontaneous aircraft handling error was not adequately managed. Hence, an error-induced accident occurred. This is one of many accidents in which threats are not prerequisites for errors to occur. Maurino and

Murray (2009) confirmed that approximately 50% of errors are independent of the presence of threats.

2.3.2.1 *Threat and error management applications.*

There are various ways in which TEM can be applied and adopted in areas such as system health checks, safety analysis and teaching. In terms of performing system health checks, Reason (2000) explained that there are two ways to determine the level of system health: the person's approach and the system's approach. The person's approach focuses on the unsafe acts of people, typically at the sharp end (e.g., pilots), and provides countermeasures that aim to reduce unwanted variabilities in human behaviour (Reason, 2000). Conversely, the system's approach focuses on the blunt end (e.g., airlines) that supports and shapes the people within it (Dekker, 2006). The TEM model accommodates both approaches in an integrated manner.

Maurino (2005a) argued that, because the TEM model is descriptive and diagnostic in nature, it can simultaneously focus on human and system performance. Kharchenko, Chynchenko, and Raychev (2007) added that the TEM model is descriptive because it can capture and assess the performance of both humans and systems in the normal operations context. Further, the TEM model is considered diagnostic, as it allows the quantification of complexities related to describing human and system performance in the operational context (Kharchenko et al., 2007). TEM, when combined with LOSA, can then produce an understanding of systemic patterns within a sequence of events, thereby helping to clarify human performance, needs, strengths and vulnerabilities (Earl et al., 2012). The findings can subsequently be proactively incorporated into the existing initial and/or recurrent training syllabi, and, if necessary, additional training can be provided. Maurino (2005) further highlighted the adaptability of the TEM model, in that it can be used for different levels and sectors within an organisation and across different organisations within the aviation industry.

Another application of TEM is to use it as a safety analysis tool focusing on a single event, such as analysing accidents/incidents, or a large set of events, similar to conducting operational audits. Wiggins (2005) described that a number of factors can surface as an accident investigation proceeds. The use of TEM as a safety analysis tool can highlight specific areas of failure within the system, where further investigation can be conducted to achieve tangible improvement in overall system safety (Maurino, 2005; Wiggins, 2005).

Another application of TEM is in the instructional area, and the use of the TEM model in the training environment has had positive effects (Thomas, 2004). Thomas (2004) explained several benefits associated with the interpretation and analysis of simulation-based training syllabi that could well be translated into the actual aircraft cockpit. Incorporating types of common operational threats increases the realism of training and training efficiency, and allows trainees to adopt systematic approaches to error management training (Thomas, 2004). Klinect (2005) added that, although TEM was originally developed as a data collection measure for LOSA, it has become widely recognised and accepted as a foundation for human factors and CRM training in the airline industry.

Similar to CRM and its expansion to other high-reliability industries, TEM has also seen application in those industries, such as offshore drilling and healthcare, despite the shorter time it has been around. However, it is worth noting that, because of its shorter period of existence, at the time of writing, there were very limited, if any, studies available evaluating TEM. Therefore, the following section primarily focuses on the ways in which the principles and concepts of TEM have been applied in some high-reliability industries.

2.3.2.2 Threat and error management in high-reliability industries.

It is generally accepted that the aviation industry has played a pivotal role in helping other industries (such as the healthcare industry) develop and shape numerous techniques to continue to improve (patient) safety (O'Connor et al., 2008). The application of TEM principles and concepts in the healthcare industry was no exception. Ruskin et al. (2013) developed a TEM-based predictive risk taxonomy applicable to anaesthesiologists. The taxonomy contained a list of threats for each stage of surgical anaesthesia to better prepare and manage situations that would otherwise develop into undesired patient states (Ruskin et al., 2013). The study also had an emphasis on identifying specific threats and associated errors and their management to be considered when developing or revising education programs and other quality improvement initiatives for current and future healthcare professionals (Ruskin et al., 2013).

Catchpole et al. (2006) conducted LOSA-style research that involved a direct observation methodology based on a TEM model to examine the systemic sources of threats and errors that led to minor and major failures in paediatric cardiac surgery. Based on 24 successful operations, the study identified a total of 366 minor failures that were associated with 406 threats and 218 errors. The most frequent type of threat was a task-related threat, such as equipment failure, and the most frequent type of error was a non-technical error, such as breakdown in coordination and communication among the surgical team (Catchpole et al., 2006). The study also found that seven operations could have been unsuccessful because of major failures associated with a greater number of minor failures, and cumulative effects, such as operations being high risk and/or long in duration (Catchpole et al., 2006). The findings from the study supported the application of the TEM model to identify threats and errors that develop into minor and major failures. Importantly, a closer examination of threat and error linked minor and major failures and lessons learned from it can inform recommendations on potential safety-related interventions and education programs for relevant healthcare authorities.

Kvalheim and Haugen (2014) examined the application of the TEM framework to the offshore drilling industry by using case studies. Similar to aviation, the offshore drilling industry has realised the importance of diverting away from the 'traditional sharp end compliance-based approach towards gaining a deeper understanding of the context' that drilling crews experience on a daily basis (Kvalheim & Haugen, 2014, p. 390). To this end, the TEM framework was adopted to help unfold complex accidents/incidents, so that areas of concern could be identified, followed by implementing appropriate strategies to properly address these concerns, identify the challenges of real-life drilling situations, and importantly determine how drilling crews manage such challenges (Kvalheim & Haugen, 2014). These findings are being used as a base to improve current training content and practices in the offshore drilling industry.

2.4 Chapter Summary

This chapter has provided a review of the literature on LOSA, CRM and TEM. LOSA is a proactive and targeted observational methodology to gather data on system safety and flight crew performance during normal operations. The findings from LOSA are predominantly non-technical skills and CRM related, which are a major focus to improve the safety and efficiency of the system in the aviation industry. Both CRM and TEM training are endorsed and recommended by ICAO for flight crews to effectively use all the resources available to avoid, mitigate and manage everyday threats and errors that present during normal operations. The benefits of such training and concepts within aviation have attracted other safety critical industries, such as healthcare, to adopt them and appropriately modify them as required, so they can further improve current training programs and practices.

Chapter 3: Methodology

Crotty (1998) suggested that two fundamental questions need to be answered when planning a research study: which methodologies and methods should be adopted, and how is this choice justified? The provision of adequate and appropriate justification regarding the methodologies and methods planned or proposed to be adopted is based on the researcher's theoretical perspective, which is embedded, in and informed by, the proposed methodology and epistemology (Crotty, 1998). Figure 3.1 below illustrates these four different process elements.

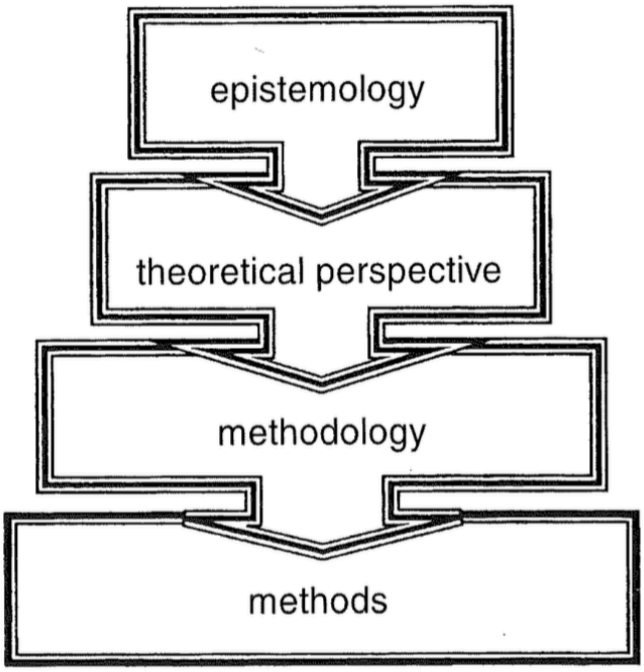

Figure 3.1. Four elements of research process when developing a study. Source: adapted from Crotty (1998, p. 4).

Crotty (1998) argued that a solid understanding of the elements of the research process and correctly following the research process will ensure the soundness of the proposed research, a robust research process and convincing outcomes. Thus, it is appropriate to provide a brief description of these elements for the current study.

3.1 Epistemology and Theoretical Perspective

Maynard (1994) described epistemology as a set of philosophical groundings that informs researchers' decision-making on the types of knowledge they are seeking, and ensures their research and findings are adequate and legitimate. Simply stated, epistemology describes 'a way of understanding and explaining how we know what we know' (Crotty, 1991, p. 3). The term 'epistemology' has also been referred to as a 'worldview' (Creswell, 2014), 'paradigm' (Lincoln, Lynham, & Guba, 2011) and 'evolving position to research methodologies' (Neuman, 2011). A number of epistemologies inform different types of methodologies. The following three subsections briefly describe post-positivism, constructivism and pragmatism, within which three widely used methodologies are embedded: quantitative, qualitative and mixed methods, respectively (Creswell, 2014).

3.1.1 Post-positivism.

Crotty (1998) described post-positivism—the thinking after positivism that is strongly influenced by unchallengeable natural laws—as a less attenuated form of positivism, claiming a certain level of objectivity, instead of absolute objectivity. This shift in viewpoint arose after scientists, including positivist scientists, realised that the world addressed by positivist science often differed from the real world (Crotty, 1998). This was particularly the case when studying the behaviour and actions of humans, and trying to solve issues empirically (Creswell, 2014). Neuman (2011) argued that, in terms of social science, humans exhibit qualitative differences from one another and from the types of objects studied, which led to uncertainty about whether differences in human characteristics and the environment in which they live require small or entire adjustments to the research approach. Despite such limitations, post-positivism still holds a philosophy that is deterministic and reductionist (Creswell, 2014). The former involves cause-finding research that aims to determine and/or identify causes for effects, while the latter refers to reducing ideas to a small, discrete set for the purpose of testing hypotheses and research questions (Creswell, 2014). This type of epistemology is more strongly embedded in research that is quantitative in nature.

3.1.2 Constructivism.

While positivism and post-positivism base their viewpoints on the objectivism that is embedded in quantitative research, constructivism rejects the idea of discovering objective reality and meaning that is independent from, for instance, variability among individuals (Crotty, 1998). Rather, constructivism is based on subjectivism, and argues that truth or meaning is constructed and comes in existence while individuals interact with reality (Crotty, 1998). Therefore, constructivist researchers focus on understanding others' subjective viewpoints of the world to generate a theory or pattern of meaning (Creswell, 2014). This inductive approach is typically aligned with qualitative research. It is noteworthy that, although constructivism has had substantial influence as a key epistemology on the development of qualitative research, this is not exclusive, as there are other epistemologies (e.g., interpretivism) that have influenced, to a certain degree, diverse ranges of qualitative research (Lee, 2012).

3.1.3 Pragmatism.

While the paradigm–method fit issue between two conflicting epistemologies (post-positivism and constructivism) was debated, a third epistemology—pragmatism—first surfaced in the 1950s (Creswell & Plano Clark, 2011) and rose to prominence in the late 1980s (Creswell, 2014). The central focus of pragmatism is practicality and the importance of gaining both objective and subjective knowledge, thereby allowing researchers to adopt a pluralistic approach to derive knowledge about the research problem, rather than subscribing to a certain epistemology-embedded methodology (Creswell, 2014). This flexibility allows researchers, for example, to commence a study using the deterministic and reductionist nature of post-positivism to obtain specific variables and empirical measures to test a theory. This is followed by taking advantage of the in-depth and contextual nature of qualitative findings, within a single study, to supplement and/or complement the overall research findings (Creswell, 2014). The reverse order or alternating of more than two epistemologies can also be adopted within a single study. Ultimately, use of such a dialectical approach in a field such as aviation is dependent on the research questions and answers sought (Ferroff, Mavin, Bates, & Murray, 2012). This approach involves epistemology associated with a mixed methods methodology.

3.2 Methodology

Crotty (1998) defined methodology as 'the strategy, plan of action, process or design lying behind the choice and use of particular methods and linking the choice and use of methods to the desired outcomes' (p. 3). As aforementioned, there are three methodologies commonly used in research: quantitative, qualitative and mixed methods. Simply stated, quantitative research involves numeric data, whereas qualitative research involves written data. Research adopting mixed methods as a methodology implements an integrative approach that 'attempts to consider multiple viewpoints, perspective, position and [paradigmatic] standpoints of qualitative and quantitative research' (Johnson, Onwuegbuzie, & Turner, 2007, p. 113).

3.2.1 Preferred methodology.

Ferroff et al. (2012) identified that, although the quantitative research methodology is preferred in aviation research, inevitable loss of detail through the numeric representation of the qualitative process is often experienced. Bryman (2006) added that the purpose of illustrating findings is 'often referred to as putting meat on the bones' (p. 106) of quantitative findings, so that research questions can be more accurately and completely answered. Therefore, the qualitative process within a single study—whether at the beginning of or after the quantitative process—is desirable so that generalised and representative findings can be supplemented and complemented with in-depth, contextual findings.

Neuman (2011) supported this pluralistic approach by using an example of surveyors, who observe and measure a distance between objects from multiple angles to obtain a good fix on the object's true location. Neuman (2011) emphasised the importance of the process of triangulation—a direct comparison between two different, yet often supplementary and complementary, datasets (quantitative and qualitative data) to promote synergism, where the weakness of using a single dataset is reduced, while the strengths of each approach are further enhanced. This results in a better understanding of the studied phenomenon and subsequently more credible corroboration and validation of the findings (Neuman, 2011).

Bryman (2006, pp. 104–107) conducted a content analysis of 232 social science articles that integrated two research approaches (quantitative and qualitative), and provided the 16 rationales propounded for this integration, as listed below:

1. triangulation or greater validity—refers to the traditional view that quantitative and qualitative research might be combined to triangulate findings so that they may be mutually corroborated
2. offset—refers to the suggestion that the research methods associated with both quantitative and qualitative research have their own strengths and weaknesses, so that combining them allows the researcher to offset the methods' weaknesses to draw on the strengths of both
3. completeness—refers to the notion that the researcher can bring together a more comprehensive account of the area of enquiry in which he or she is interested if both quantitative and qualitative research are employed
4. process—quantitative research provides an account of structures in social life, yet qualitative research provides sense of process
5. different research questions—the argument that quantitative and qualitative research can each answer different research questions
6. explanation—one approach is used to help explain the findings generated by the other approach
7. unexpected results—the suggestion that quantitative and qualitative research can be fruitfully combined when one generates surprising results that can be understood by employing the other
8. instrument development—refers to contexts in which qualitative research is employed to develop questionnaires and scale items, for example, so that better working or more comprehensive closed answers can be generated
9. sampling—refers to situations in which one approach is used to facilitate the sampling of respondents or cases
10. credibility—refers to suggestions that employing both approaches enhances the integrity of findings
11. context—refers to cases in which the combination is rationalised in terms of qualitative research providing contextual understanding, coupled with either generalisable, externally valid findings or broad relationships among variables uncovered through a survey

12. illustration—refers to the use of qualitative data to illustrate quantitative findings, often referred to as putting ‘meat on the bones’ of ‘dry’ quantitative findings
13. utility or improving the usefulness of findings—refers to a suggestion (that is more likely to be prominent among articles with an applied focus) that combining the two approaches will be more useful to practitioners and others
14. confirm and discover—entails using qualitative data to generate hypotheses and quantitative research to test the hypotheses within a single project
15. diversity of views—this includes two slightly different rationales: (i) combining researchers’ and participants’ perspectives through quantitative and qualitative research, respectively, and (ii) uncovering relationships between variables through quantitative research, while also revealing meanings among research participants through qualitative research
16. enhancement or building on quantitative/qualitative findings—this entails a reference to making more of augmenting either quantitative or qualitative findings by gathering data using a qualitative or quantitative research approach.

The majority of Bryman’s (2006) rationales are closely aligned with the central aim of this thesis—exploring the effects of TEM on Australian general aviation. It has been more than nine years since TEM was implemented in Australian general aviation. However, there has been very limited formal review, assessment or evaluation to examine the situation after the introduction of TEM in Australian general aviation. To properly and adequately examine this issue, a dataset is required (such as recounts of experienced pilots in general aviation), and this dataset needs to be supplemented with another dataset such as survey results (Bryman’s credibility and context). This can be achieved by designing a study so that the weaknesses arising from a single methodological approach are minimised, while the strengths gained from a combination of two methodological approaches are enhanced (Bryman’s triangulation, offset, completeness and enhancement). Therefore, a mixed methods approach was selected to be the appropriate methodology for this thesis.

3.2.2 Brief history of mixed methods.

Similar to the development of the aviation industry to date, the mixed methods approach underwent several stages to reach its current form. According to Johnson et al. (2007), although the use of mixed research (as it was referred to in the early stage of the development of mixed methods) was adopted in research during the early years of the twentieth century, Campbell and Fiske's (1959) article was often credited as the first to formalise the practice of using multiple research methods. The authors employed multiple quantitative datasets in a single study when validating an underlying psychological phenomenon or trait (Johnson et al., 2007). Denzin (1978, as cited in Johnson et al., 2007) added that, despite the use of multiple datasets in a within-methods triangulation study (use of either multiple quantitative or qualitative approaches in a single study), the limited value of the study was soon realised because of the inherent weakness stemming from the use of a single approach. This subsequently encouraged researchers to adopt between-methods triangulation (use of both quantitative and qualitative approaches in a single study). Creswell and Plano Clark (2011) observed more and more researchers advocating for and advancing this methodological pluralism, while it underwent several periods of transformation—formative, paradigm debate, procedural development, advocacy, expansion and reflective—for it to receive appropriate credence as the third methodology within day-to-day research practice.

3.2.3 Definition of mixed methods.

While mixed methods as a methodology underwent its developmental stages, various terms and definitions were provided. Johnson et al. (2007) held a discussion via electronic means with 21 distinguished and highly regarded researchers in the field of mixed methods, and reviewed 19 definitions of mixed methods to consolidate and provide a composite definition of mixed methods. Johnson et al. (2007) offered the following definition:

> The type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference

> techniques) [in a single study or in multiple phases of a program of study] for the broad purposes of breadth and depth of understanding and corroboration. (p. 123)

Johnson et al. (2007) also identified that, although some definitions were homogenous while others were heterogeneous, the definitions shared common themes. Creswell and Plano Clark (2011, p. 5) provided the following consolidated list of common characteristics of mixed methods found in the diverse range of viewpoints. Mixed methods research:

- persuasively and rigorously collects and analyses both qualitative and quantitative data (based on research questions)
- mixes (or integrates and links) the two forms of data concurrently by combining them (or merging them) sequentially by having one build on the other, or embedding one within the other
- gives priority to one or both forms of data (in terms of what the research emphasises)
- uses these procedures in a single study or in multiple phases of a program of study
- frames these procedures within philosophical worldviews and theoretical lenses
- combines the procedures into specific research designs that direct the plan for conducting the study.

These characteristics were the result of many years of evolution in mixed methods as the third methodology. It is noteworthy that, because of the practicality and flexibility found in mixed methods designs, there are different types of mixed methods designs. Section 3.2.4 and its subsections discuss the four major mixed methods designs that are commonly used in research.

3.2.4 Four major mixed methods designs.

After a researcher has determined that the use of mixed methods is most appropriate for given research questions, the researcher must select the most appropriate mixed methods design from the different types of designs, as ‘each major design has its own history, purpose, considerations, philosophical assumptions, procedures, strengths, challenges, and variants’ (Creswell & Plano Clark, 2011, p. 53). The four most common mixed methods designs are: convergent parallel design, explanatory sequential design, exploratory sequential design and embedded design (Creswell & Plano Clark, 2011). Primarily, the differences among these designs are based on the point at which the collected quantitative and qualitative data are mixed, which is referred to as the point of interface or timing of data collection. The following subsections provide a brief description of each design and a summary of research that has adopted each design as an example.

3.2.4.1 Convergent parallel design.

Convergent parallel design (Figure 3.2) is the most well-known, and perhaps the most common, approach to mixed methods (Creswell & Plano Clark, 2011). The design refers to the concurrent, yet separate, collection and analysis of quantitative and qualitative data within the same phase of the research process, prior to merging the two sets of results to provide the overall interpretation (Creswell & Plano Clark, 2011). Both quantitative and qualitative data are given equal priority, and this differentiates the convergent parallel design from another similar type of design—embedded design, which is described below (Section 3.2.4.4). The point of interface is after the quantitative and qualitative data are independently collected and analysed. A brief summary of a study that adopted this type of design is presented below as an example.

Wittink, Barg, and Gallo (2006) conducted a study using convergent parallel design to better understand the aspects of the doctor–patient relationship that may influence the way patients discuss their depression status. The first phase of the study was quantitative in nature, and compared the doctors’ assessments of the patients’ level of depression with the results from a 36-item short-form health survey specifically devised for this study (Wittink et al., 2006). In the second phase, transcribed data from semi-structured interviews were separately analysed (qualitative) to develop themes. The point of interface occurred after these two data sources were separately analysed. The authors' recount suggested that the additional interview data provided possible contributing factors to better illustrate the dynamic process of the doctor–patient interaction on depression (Wittink et al., 2006).

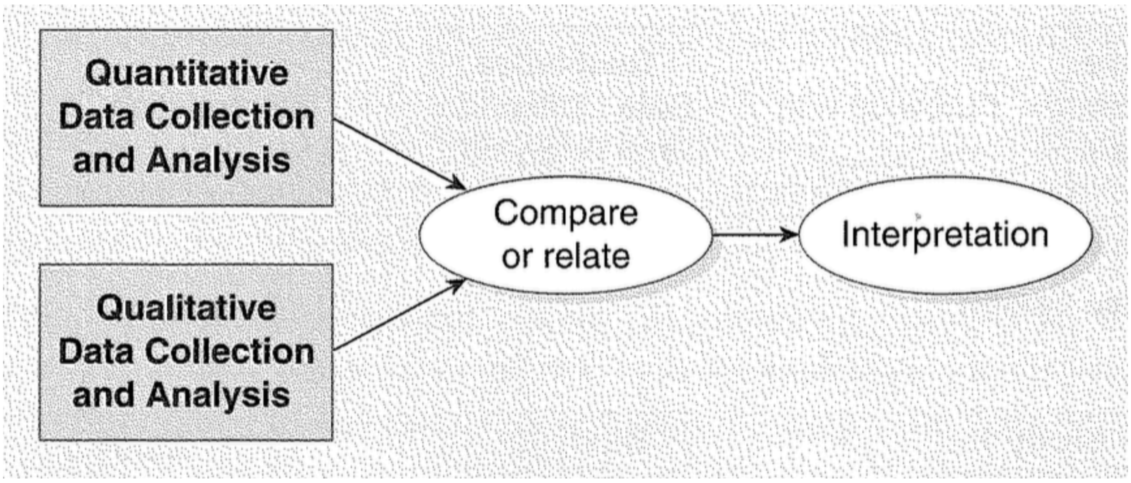

Figure 3.2. Convergent parallel design. Source: adapted from Creswell and Plano Clark (2011, p. 69).

3.2.4.2 *Explanatory sequential design.*

Explanatory sequential design (Figure 3.3) refers to a two-phase, mixed methods design that includes collecting and analysing quantitative data first, followed by collecting and analysing qualitative data (Creswell & Plano Clark, 2011). The main reason for employing this type of follow-up design is to use the qualitative phase of study to explain the findings from the initial quantitative phase—hence the name of the design. The quantitative findings that are of interest to the research inform the development and refinement of the design of the follow-up qualitative phase (Creswell & Plano Clark, 2011). After collecting and analysing the secondary qualitative data, the two datasets are merged to provide more complete answers to the research questions. A brief summary of a study that adopted this type of design is presented below as an example.

Friday (2011) adopted explanatory sequential design in an attempt to better answer the central research question: how could the current FAA certification process be modified for a private pilot to more safely operate in the National Airspace System with the introduction of advanced technology in general aviation aircraft? This study sequentially collected and analysed both quantitative and qualitative data. The author first collected and analysed quantitative data to gather participants' demographic information (e.g., professional qualifications and flight experience) and overall perceptions about the influence of advanced technology in certain key areas of interest (Friday, 2011). The first part of study was followed by qualitative interviews to further investigate the areas of interest and explore potential solutions to the central research question.

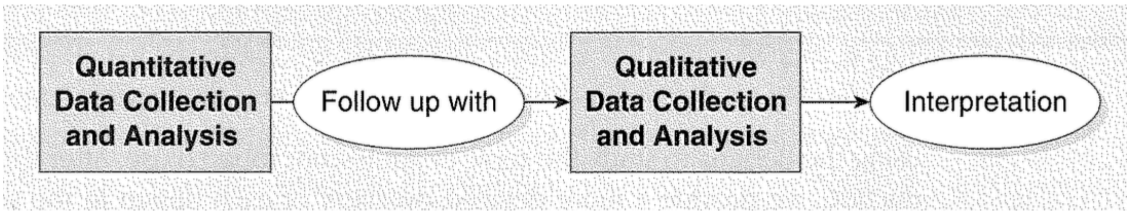

Figure 3.3. Explanatory sequential design. Source: adapted from Creswell and Plano Clark (2011, p. 69).

3.2.4.3 *Exploratory sequential design.*

This design (Figure 3.4) is also referred to as a two-way, instrument development (e.g., survey) mixed methods design; however, the sequence is opposite to explanatory sequential design. The design begins by devising the first phase of the study, which includes collection and analysis of qualitative data to explore a topic (hence its name), followed by developing an appropriate instrument and identifying variables or stating propositions for testing to generalise the findings from the first phase of the qualitative study (Creswell & Plano Clark, 2011). A brief summary of a study that adopted this type of design is presented below as an example.

Martin (2013) chose mixed methods as a preferred methodology for his doctoral thesis on the effects of pathological phenomena—startle, freeze and denial—on a situation outcome. The study used an exploratory sequential design, where qualitative data were obtained through semi-structured interviews with pilots who had experienced emergency situations or unexpected critical events. The data were analysed to create a clear picture of the concepts and likely involvement of the phenomena in aviation accidents and incidents (Martin, 2013). This was followed by conducting an experiment to collect quantitative and qualitative data from 'startle' experiments in a Boeing 737 simulator (Martin, 2013). The findings were combined and interpreted to generate a complete picture of the potential threats that the pathological phenomena presented.

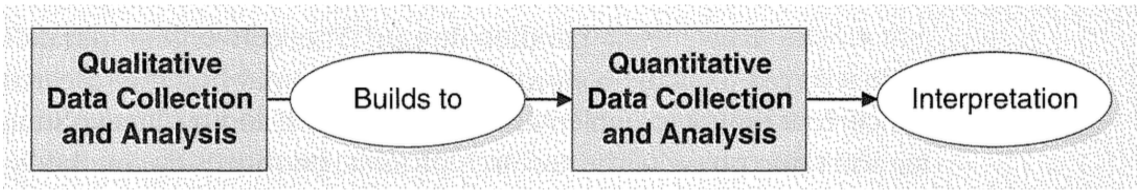

Figure 3.4. Exploratory sequential design. Source: adapted from Creswell and Plano Clark (2011, p. 69).

3.2.4.4 *Embedded design.*

This design (Figure 3.5) is similar to the convergent parallel design previously described. However, the difference is that embedded design does not collect secondary data, whether quantitative or qualitative, that are given the same priorities and keeping them independent (Creswell & Plano Clark, 2011). Instead, supplemental quantitative or qualitative data are collected and added to major data collection—qualitative or quantitative, respectively—to enhance the overall design and findings. These supplemental data can be collected prior to an experimental trial to provide preliminary exploration (sequential/before), during the experimental trial to provide a more complete understanding of the experimental trail (concurrent/during) and after the experimental trial to provide follow-up explanations (sequential/after) (Creswell & Plano Clark, 2011). A brief summary of a study that adopted this type of design is presented below as an example.

Foster, Curtis, Mitchell, Van, and Young (2016) conducted a prospective, two-year longitudinal study involving the parents of children (aged zero to 12) who were severely injured, and using the embedded design as part of a program of research on paediatric trauma. Foster et al. (2016) considered data collected from semi-structured interviews as their primary source, and these were collected at four time points: the acute hospitalisation phase and then six, 12 and 24 months after the severe injury of the child. These data were expected to provide an in-depth understanding of parents' emotional wellbeing and to ascertain the factors that supported or impeded the parents' ability to manage their child's injury (Foster et al., 2016). The primary qualitative strand was supplemented with four self-report standardised measures: a paediatric inventory for parents to measure their own experiences of emotional functioning (e.g., stress); a parent quality-of-life measure with regard to parents' physical and mental health; depression, anxiety and stress measures to assess parents' risk of psychological distress; and the Connor Davidson resilience scale to assess parents' resilience (Foster et al., 2016). The study findings were expected to provide evidence-based recommendations to better support the parents and families of injured children and to strengthen their capacity during this traumatic time (Foster et al., 2016).

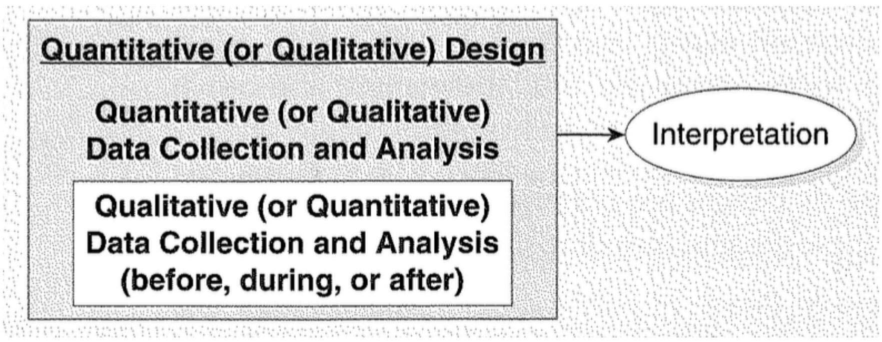

Figure 3.5. Embedded design. Source: adapted from Creswell and Plano Clark (2011, p. 69).

3.2.5 Preferred mixed methods design.

The emergence of TEM is well regarded as an effective method of improving aviation safety, with TEM becoming a mandatory assessment item for various levels of flight tests and ground examinations in Australia. Although it has been more than nine years since TEM training was implemented in Australian general aviation, there are no definitive data available to establish whether TEM training has been well received or whether a positive effect of this training has been witnessed, experienced or translated in the general aviation sector. The present study intends to fill the gap by exploring the effect of introducing TEM among general aviation pilots in Australia. The overall results from the exploratory study (Study 1 in Chapter 4) indicated a variable uptake of TEM principles and differing opinions regarding its effectiveness. Given the lack of definitive data or theory established, and the differing opinions in this area, exploratory sequential design was adopted in this thesis to initially better understand how TEM has been implemented in Australian general aviation, followed by generalising the findings. Figure 3.6 below illustrates the research design adopted in this thesis.

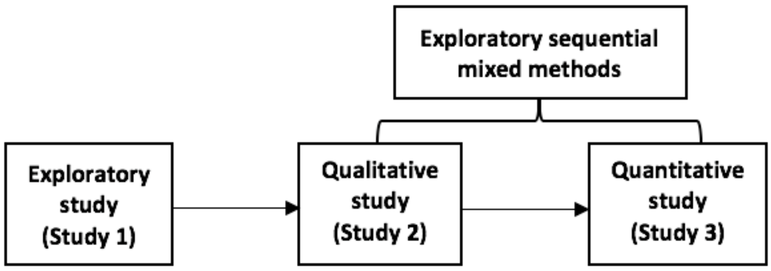

Figure 3.6. Research design adopted for the thesis.

3.3 Chapter Summary

This chapter has provided a brief overview of epistemology and the theoretical perspective of three methodologies commonly used in research (quantitative, qualitative and mixed methods). Justification for the preferred methodology (mixed methods) and descriptions of the four commonly adopted mixed methods designs were then provided. The current study adopted ‘exploratory sequential design’ because this was considered to appropriately agree with the central aim of this thesis—conducting a post-implementation assessment to examine the after-state of TEM implementation, in terms of the effectiveness of introducing TEM in Australian general aviation.

Chapter 4: Study 1 (Exploratory Study)

4.1 Purpose

This study was exploratory in nature and sought to identify the way in which TEM training was conducted and regarded among general aviation pilots. The main focus was its benefits and effectiveness, after it was mandated in July 2009. The findings better informed the researcher about whether further study on this topic on a larger scale was considered worthwhile. In addition, the study allowed the researcher to select an appropriate methodology and methods after this trial run.

4.2 Methods

4.2.1 Design overview.

The primary purpose of the study was to collect preliminary information on how the requirement for TEM training is currently being addressed, and to gather general consensus on the benefits and effectiveness of TEM training. To collect the required data, the researcher devised a survey with his supervisors using questions based on an ATSB report (Cheng, Inglis & Godley, 2009). The survey contained several common questions (e.g., 'I feel that appropriate use of TEM training improves non-technical skills, such as situational awareness and decision-making') for all pilot groups (trainee pilots, private pilots, trainer pilots and other pilots, such as charter pilots), as well as group-specific questions (e.g., 'I feel that incorporation of TEM has improved the way that airmanship is taught'). An open-ended question was included at the end of the survey to offer the participants the opportunity to share their opinions on TEM training.

4.2.2 Participants.

A total of 63 participants completed the survey; however, four participants failed to indicate their consent. Consequently, responses from 59 participants were analysed for this study. These participants comprised 26 trainee pilots working towards a PPL or higher licence, five private/recreational pilots, 21 trainers (e.g., flight/ground instructors) and seven other pilots (e.g., a bible pilot and a charter pilot).

4.2.3 Materials.

Although no separate headings were used, the survey comprised three main parts. The first part aimed to categorise the general aviation pilots into four groups (trainee pilots, private pilots, trainer pilots and other pilots, such as charter pilots). The second part intended to gain insights into how TEM training was delivered. The third part aimed to determine how the benefits and effectiveness of TEM training were regarded among general aviation pilots. This part contained seven common questions (Table 4.1) and group-specific questions. Table 4.2 presents the group-specific questions for the trainee pilots and trainers.

Table 4.1 *Seven Common Questions.*

Q1. I feel that the number of accidents/incidents reduced in my organisation after TEM training was introduced.
Q2. I feel that TEM training has improved overall aviation safety in general aviation.
Q3. I use TEM principles in my day-to-day flying.
Q4. I feel that CASA produces adequate training and guidance materials for TEM training for general aviation.
Q5. I feel that appropriate use of TEM training improves my technical skills (i.e., aircraft handling skills).
Q6. I feel that appropriate use of TEM training improves my non-technical skills, such as situational awareness and decision-making.
Q7. I feel that the benefits of TEM training are overrated.

Table 4.2 *Group-specific Questions for the Trainee and Trainer Groups.*

Group	Questions
Trainees	Q1. I feel that TEM training is well integrated into the flight training syllabus in my organisation. Q2. I feel that the TEM training methods/techniques used among flight instructors in my organisation are well standardised. Q3. I feel that my flight instructor strongly emphasises TEM during pre-/post-flight briefings, as well as during instructional flights. Q4. I feel that TEM is appropriately assessed during dual flights. Q5. I feel that TEM is appropriately assessed during flight tests.
Trainers	Q1. I feel that TEM is well integrated into the flight training syllabus in my organisation. Q2. I feel that the TEM training methods/techniques used among flight instructors in my organisation are well standardised. Q3. I feel that my students value TEM training as part of flight training.

Q4. I feel that TEM is appropriately assed during flight tests.
Q5. I feel that incorporation of TEM has improved the way that airmanship is taught.

The survey questions in the third part allowed six response options: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree and unsure. For the purpose of statistical analysis, the following numerical values were assigned to each of the potential responses: 'strongly disagree' = 1, 'disagree' = 2, 'neither agree nor disagree' and 'unsure' = 3, 'agree' = 4 and 'strongly agree' = 5.

4.2.4 Procedure.

Invitations to participate in this study were sent to 14 flight training organisations, of which seven organisations agreed to participate (50%). As a result of the limited resources available, the majority of flight training organisations contacted were within an approximately 100 km radius of the area in which the researcher resided. The survey was designed to be used online; however, the survey questions were printed and sorted for appropriate groups because of the possibility of unreliable internet connections at some of the intended survey sites. A purposive sampling technique was used when recruiting potential participants.

When approaching each participant, the overall purpose of the research was explained (the overall purpose was also available on the first two pages of the survey), and it was emphasised that participation in the survey was completely voluntary, and that participants' anonymity and confidentiality would be strictly maintained. No remuneration was offered as an incentive to participate in the survey.

This study was considered to involve the minimal level of risk and ethical issues in accordance with the National Statement of Ethical Conduct in Research Involving Humans. Therefore, an online Expedited Ethical Review Checklist was completed when applying for ethical clearance. Acknowledgement of ethics approval is presented in Appendix 4.

4.3 Results and Discussion

A manuscript based on this exploratory study was submitted and subsequently accepted by the *International Journal of Training Research*. This peer-reviewed journal article was published in December 2016, and is presented in Section 4.3.3. Readers are referred to this journal article for the main findings and discussion. The results based on the group-specific questions are presented below.

4.3.1 Results (group-specific questions).

Statistical Package for the Social Sciences (SPSS) (Version 23 for Apple Macintosh computers) was used with the level of significance (alpha) set to be $p < .05$ for all statistical analyses. The results were presented in a combination of descriptive and inferential statistics.

A mean score for each group-specific question was compared with the mean score of the scale of 3, which would indicate that the participants neither agreed nor disagreed and/or were unsure about the effect of TEM training and its assessment. Trainees' responses to each of the five questions are presented in Table 4.3, which indicates the mean difference score, one-sample *t*-test, degrees of freedom and significance for the five questions. The results presented in Table 4.3 suggest that the participants were more in agreement with all five questions.

Table 4.3 *Trainee Group-specific Questions (Mean Difference from Neutral, One-sample t-test, Degrees of Freedom and Significance for Individual Questions).*

Trainee group-specific questions	Mean difference from neutral = 3 (*SD*)	t	df	p
Q1. I feel that TEM training is well integrated into the flight training syllabus in my organisation.	1.31 (.74)	9.06	25	< .001
Q2. I feel that the TEM training methods/techniques used among flight instructors in my organisation are well standardised.	1.08 (.94)	5.88	25	< .001
Q3. I feel that my flight instructor strongly emphasises TEM during pre-/post-flight	1.08 (1.09)	5.03	25	< .001

briefings, as well as during instructional flights.				
Q4. I feel that TEM is appropriately assessed during dual flights.	1.12 (.91)	6.26	25	< .001
Q5. I feel that TEM is appropriately assessed during flight tests.	1.08 (1.2)	4.59	25	< .001

The same analysis was conducted for the group-specific questions for the trainer group. Table 4.4 presents the mean difference score, one-sample *t*-test, degrees of freedom and significance for the five questions. Again, the results presented in Table 4.4 suggest that the participants were more in agreement with the majority of the questions.

Table 4.4 *Trainer Group-specific Questions (Mean Difference from Neutral, One-sample t-test, Degrees of Freedom and Significance for Individual Questions).*

Trainer group-specific questions	Mean difference from neutral = 3 (*SD*)	t	df	p
Q1. I feel that TEM is well integrated into the flight training syllabus in my organisation.	.95 (.74)	5.9	25	< .001
Q2. I feel that the TEM training methods/techniques used among flight instructors in my organisation are well standardised.	.81 (.87)	4.25	25	< .001
Q3. I feel that my students value TEM training as part of flight training.	.81 (.81)	4.56	25	< .001
Q4. I feel that TEM is appropriately assessed during flight tests.	.67 (1.2)	2.55	25	= .019
Q5. I feel that incorporation of TEM has improved the way that airmanship is taught.	.86 (.91)	4.32	25	< .001

Three questions from Tables 4.3 and 4.4 were identical. These questions were:

- I feel that TEM is well integrated into the flight training syllabus in my organisation.
- I feel that the TEM training methods/techniques used among flight instructors in my organisation are well standardised.
- I feel that TEM is appropriately assessed during flight tests.

A Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the trainee and trainer groups. The results indicated that they were not significantly different for all three questions: $U = 197.5$, $z = -1.809$,

$p = .071$, $r = -.26$; $U = 216.5$, $z = -1.341$, $p = .18$, $r = -.20$; and $U = 201.5$, $z = -1.646$, $p = .10$, $r = -.24$. The results suggested that both the trainee and trainer groups were more in agreement that TEM was well integrated into the flight training syllabus in their organisations. In addition, both groups were more in agreement that the TEM training methods and techniques used among Flight Instructors (FI) in their organisations were well standardised. These are considered positive findings. However, there were mixed responses among the participants in the trainer group with regard to the appropriate assessment of TEM during flight tests.

4.3.2 Interim discussion.

The study findings suggested that participants from both the trainee and trainer groups were more in agreement with the group-specific questions. That is, the majority of participants agreed that TEM was well integrated into the flight training syllabus, that training methods/techniques were well standardised, that instructors emphasised TEM during pre-/post-flight briefings and during instructional flights, and that trainees valued TEM training as part of their flight training. In addition, the majority of trainees and trainers considered that TEM was appropriately assessed during dual flights and flight tests, although there were mixed responses among participants in the trainer group regarding the latter. These are considered encouraging findings because positive views of safety initiatives, such as TEM implementation, can strongly influence safety behaviours and the environment.

4.3.3 Results and discussion (main findings).

INTERNATIONAL JOURNAL OF TRAINING RESEARCH, 2016
http://dx.doi.org/10.1080/14480220.2016.1259006

Routledge
Taylor & Francis Group

An exploratory study on the post-implementation of threat and error management training in Australian general aviation

Seung Yong Lee, Paul R. Bates, Patrick S. Murray and Wayne L. Martin

Aviation and Logistics, University of Southern Queensland, Springfield, Australia

ABSTRACT

Threat and Error Management (TEM) training, endorsed and recommended by the International Civil Aviation Organisation (ICAO), was mandated in Australia with the aim of improving aviation safety. However, to date, there has been very limited, if any, formal post-implementation review, assessment or evaluation to examine the 'after-state' in terms of the effectiveness of the introduction of TEM. This exploratory study, is a first step in examining how TEM is regarded among Australian general aviation pilots in terms of its use and effectiveness. Fifty-nine general aviation pilots participated in a survey. The survey results indicated a large decline in positive support for TEM compared with the findings of an earlier study. Other results from the current survey indicated a variable uptake of TEM principles and differing opinions as to its effectiveness, suggesting further study should be conducted in respect of TEM as an important safety initiative within Australian general aviation.

ARTICLE HISTORY
Received 7 May 2016
Accepted 25 October 2016

KEYWORDS
Threat and error management; general aviation; aviation safety; human factors training

Introduction

Approximately 70–80% of aviation accidents are attributed, at least in part, to human error (Wiegmann & Shappell, 2003). Besco (1992) suggested that because few human (e.g., pilot) induced errors were the consequence of technical skill deficiencies, the entire industry needed to focus on knowledge and attitude to improve pilot performance. Consequently, airlines around the world have devoted many resources to human factors training. Threat and Error Management (TEM), endorsed by the International Civil Aviation Organisation (ICAO), has been one of these recent human factors training initiatives.

TEM emerged during the development of the Line Operations Safety Audit (LOSA) and was developed to capture the full operational complexity of a flight (Merritt & Klinect, 2006). Since then, 'TEM has been generally accepted in the airline industry as an effective method of improving flight safety and is now required by the International Civil Aviation Organisation (ICAO) as an integral part of pilot training at all licence levels through to air transport pilot' (Civil Aviation Safety Authority [CASA], 2008, p. 37). In addition, anecdotal evidence has suggested that pilots who continuously receive TEM training as part of their flight training from the *ab initio* stage demonstrate earlier achievement of required competencies when

CONTACT Seung Yong Lee paul.lee@usq.edu.au

compared with pilots who receive insufficient and/or irregular TEM training. This is consistent with other studies that have demonstrated a higher achievement of post-training performance among trainees exposed to TEM (Bove, 2002).

It has been more than five years since the introduction of TEM training in Australian general aviation; thus, it is considered timely to examine how the requirement for TEM training is addressed. Areas of importance include how TEM training is regarded among Australian general aviation pilots in terms of its use and effectiveness.

Threat and error management

TEM is described as a process of detecting and responding to threats and errors to ensure that the outcome is inconsequential by maintaining adequate safety margins (CASA, 2008). Simply put, this is a required day-to-day exercise for pilots to fly an aircraft safely between two points (Maurino & Murray, 2009).

Development of TEM was a by-product of the LOSA and was developed to capture the full operational complexity of a flight (Merritt & Klinect, 2006). Since the first full, TEM-based LOSA in 1996, the concept has attracted strong interest from airlines, regulatory authorities and academia, and the benefits of such training have been widely witnessed (Merritt & Klinect, 2006). Consequently, ICAO has recommended that TEM training be a licensing requirement for all pilot licences and it has become a requirement for initial and recurrent flight crew training (Maurino, 2005b).

'The TEM model [see Figure 1] is a conceptual framework that assists in understanding, from an operational perspective, the inter-relationship between safety and human performance in dynamic and challenging operational contexts' (Maurino, 2005a, p. 1). The Civil Aviation Safety Authority (2008) emphasised that this was more than the traditional role of airmanship, as it encouraged pilots to be more proactive in identifying and managing threats

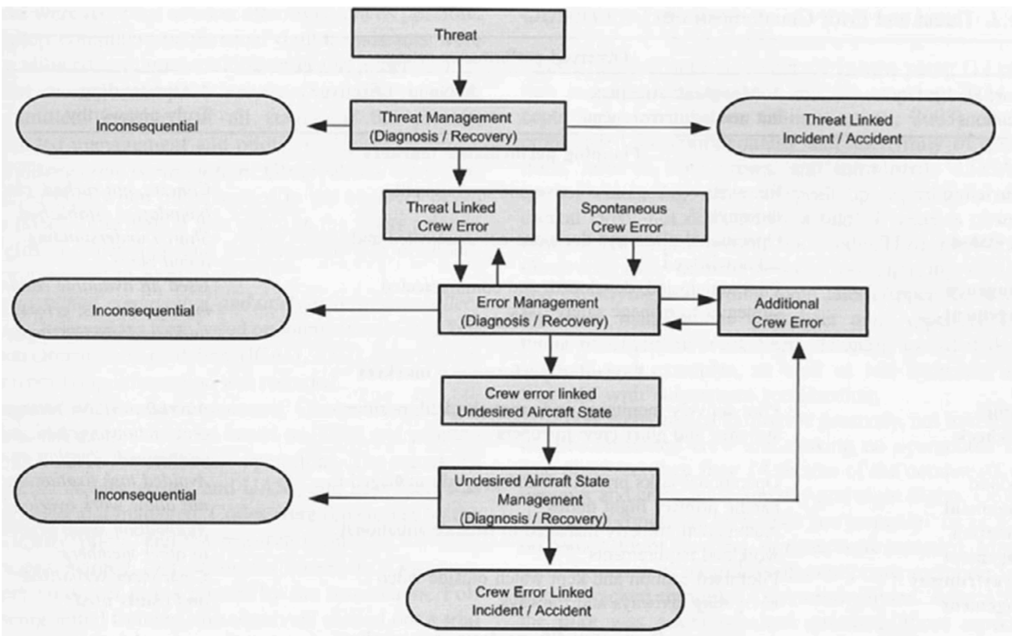

Figure 1. TEM model. Reproduced with permission from the LOSA Collaborative.

and errors even before the flight commences. The model consists of three components: Threat, Error and Undesired Aircraft State (UAS).

A threat is defined as an event that occurs outside the influence of the flight crew and consequently increases the operational complexity of a flight (Merritt & Klinect, 2006). Hence, the crew's attention to and management of threat is required so that the safety margins are maintained (Merritt & Klinect, 2006). Two types of threats are environmental threats (e.g., adverse weather condition and poor airport signage) and airline threats (e.g., aircraft system malfunctions and inadequate training manual).

Errors are defined as a flight crew's actions or inactions that lead to a deviation from the crew or organisational intentions or expectations and consequently, they reduce required safety margins (Merritt & Klinect, 2006). This could develop a higher probability of adverse operational events while the flight crew are on duty (Merritt & Klinect, 2006). Three types of errors are aircraft handling errors (e.g., incorrect flap settings and speed deviations), procedural errors (e.g., omitted checklist items and failure to cross-verify automation inputs) and communication errors (e.g., misinterpretation of instruction and within-crew miscommunication).

UAS is defined as the safety-compromising state of an aircraft that has deviated from its intended and/or optimal state due to mismanagement of the flight crew's error, consequently, required safety margins are reduced (Merritt & Klinect, 2006). This is the last opportunity for the flight crew to return the aircraft to its intended and/or optimal state so that the event is inconsequential. Types of UAS include aircraft handling and ground navigation.

Use of TEM

Reason (2000) argued that there are two ways of viewing human error problems: the person approach and the system approach. The person approach focuses on the unsafe acts of the people, typically at the sharp end (e.g., pilots), and provides countermeasures that aim to reduce unwanted variabilities in human behaviour (Reason, 2000). The system approach, conversely, focuses on the blunt end (e.g., airlines) that supports and shapes people within it (Dekker, 2006). The appropriate countermeasures are eliminated, leading to a focus on the conditions within which people work (Reason, 2000). The TEM model aims to incorporate both approaches.

Maurino (2005a) argued that the TEM model is both descriptive and diagnostic in nature and focuses simultaneously on human and system performance. TEM, when combined with LOSA, produces an understanding of systemic patterns within a sequence of events, helping to clarify human performance, needs, strengths and vulnerabilities (Earl, Bates, Murray, Glendon, & Creed, 2012). The findings identified can be proactively incorporated into the existing initial and/or recurrent training syllabi and, if necessary, additional training can be provided. Maurino (2005a) further argued for the adaptability of the TEM model, in that it can be used at different levels and sectors within an organisation, and across different organisations within the aviation industry. In addition, the TEM model can be used as a safety analysis tool focusing on a single event – for instance, analysing accidents/incidents – or on systemic patterns within a large set of events, as is the case with operational audits.

Use of the TEM model within the training environment has had positive effects. Thomas (2004) explained several benefits associated with the interpretation and analysis of simulation-based training syllabi that could well be translated into the actual aircraft cockpit.

Incorporating types of common operational threats increases the realism of training and training efficiency, and allows trainees to adopt systematic approaches to error management training (Thomas, 2004).

Statement of the problem

Despite the positive effects of extensive human factors training in the airline sector, it often appears to be a neglected component within the general aviation sector. Indeed, to date, there is, relatively, a smaller and limited number of formal ongoing human factors training courses in general aviation, when compared with the airline environment, although there has been a growing recognition of its importance in the general aviation sector. Sarter and Alexander (2000) suggested that the overall safety of the system was determined by the performance of the weakest link (i.e., single pilot, low-technology, less-rigorously trained pilots in general aviation). A lack of formal human factors training and re-currency training offered to the weakest link (i.e., lower end of pilots in the general aviation) perhaps explains why overall aviation safety has not significantly improved.

According to a report from the National Transportation Safety Board in the USA, general aviation accidents accounted for 95% of all aviation accidents, 94% of fatal accidents and 92% of all US civil aviation fatalities in 2011, while only accounting for just over half of the estimated total flight time (National Transportation Safety Board, 2014). The report also identified that, although the accident rate involving general aviation fluctuated over the period between 2002 and 2011, the overall trend remained stable for the same period (National Transportation Safety Board, 2014). Although the accident rate in Australia for the similar period did not seem to be as severe as that in the USA, the accident statistics in Australia involving general aviation showed a similar trend, where the general aviation accident rate was approximately three times higher than the accident rate involving airlines (Australian Transport Safety Bureau, 2011).

However, it is noteworthy that although the accident rate for general aviation is clearly higher, accident rates within general aviation vary by segment (i.e., different purposes of flights) because general aviation is composed of pilots with a wide variety of experience, a wide range of aircraft types and wide diversity of flight activities (Hunter, 1999). Among the wide range of aircraft types and their involvement in accidents, the National Transportation Safety Board (2014) report identified that fixed-wing aircraft accounted for approximately 88% of all general aviation accidents in 2011, while helicopters accounted for just over 8% for the same year. Hence, those general aviation segments flying fixed-wing aircraft for their operations should be an immediate focus of further research.

Traditional approaches to aviation accidents have focused on a series of single events identifying human errors that are ubiquitous in nature (Thomas, 2004), leading to a false notion of total error elimination strategies. However, it was later realised that it was unrealistic to believe that errors could be totally eliminated due to the physical and psychological limitations of humans (Thomas, 2004). Consequently, efforts have been made to identify errors and, more importantly, developments and refinement of training have been made to manage appropriately those errors in order to avoid or mitigate negative consequences. It was this notion that provided the overarching objective of error management, later renamed TEM, that was believed to provide the best possible support for pilots in managing everyday threats and errors (Merritt & Klinect, 2006).

Following ICAO's acknowledgement of the need for TEM principles to be embedded within all pilot training, CASA in Australia amended the Day Visual Flight Rules (VFR) syllabus on 1 March 2008 to reflect this initiative (CASA, 2008). In addition, CASA (2008) mandated that TEM be formally assessed on flight tests for the Private Pilot Licence (PPL), Commercial Pilot Licence (CPL) and Air Transport Pilot Licence (ATPL) from 1 July 2009. In preparation for the regulatory changes, the Guild of Air Pilots and Air Navigators (GAPAN) developed a TEM 'train-the-trainer' course for general aviation and low-capacity air transport operations, and conducted the course in 10 different locations throughout Australia between August and October 2007 (Australian Transport Safety Bureau, 2009). Following the course, two surveys, a post-training survey and follow-up survey, were administered and the findings showed that many respondents believed TEM would improve safety and the majority stated that they used TEM in their day-to-day activities (Australian Transport Safety Bureau, 2009). However, there has subsequently been very limited, if any, formal post-implementation review, assessment or evaluation to look at the 'after-state' in terms of the effectiveness of the introduction of TEM. This study, exploratory in nature, is a first step in examining how TEM training is regarded among Australian general aviation pilots in terms of its use and effectiveness.

Training evaluation

The importance of post-implementation evaluation of programmes or training initiatives has been recognised since the late 1700s and accordingly, the field of programme evaluation has continued to evolve to the twenty-first century (Hogan, 2007). In parallel, a number of evaluation approaches to newly implemented programmes or training initiatives have emerged, notably in and from the 1940s that coincided with the time period when Ralph Tyler, considered the father of educational evaluation, made a considerable contribution to evaluation (Hogan, 2007).

One of the evaluation approaches consists of four levels of evaluation: (1) participants' reaction to a newly implemented programme or training initiative (reaction); (2) changes in participants' attitude and improved skills and knowledge (learning); (3) changes in behaviour (behaviour); and (4) achievement of predetermined final objectives of a programme (Kirkpatrick & Kirkpatrick, 2006). The two surveys (i.e., the post-training survey, completed immediately after the TEM train-the-trainer course, and the follow-up survey, completed approximately eight months after the course) administered by the Australian Transport Safety Bureau (2009) are considered to have addressed the first two levels (reaction and learning) of Kirkpatrick's evaluation approach.

Method

The primary purpose of this study was to examine how TEM training was regarded among Australian general aviation pilots in terms of its use and effectiveness. Given the application of multiple methods being an emerging trend in research and the nature of the present study being exploratory, a short survey was devised and administered to make an initial assessment and to determine whether further evaluation on post-implementation of TEM training was deemed necessary. The survey aimed to collect information on how the requirement for TEM training was addressed and to gather general views on the benefits and effectiveness of TEM. Some of the findings from this survey serve as providing information on the third level of Kirkpatrick's evaluation approach (behaviour).

Participants

A total of 59 participants completed the survey. A purposive sampling technique was used when recruiting potential participants. These general aviation (GA) fixed-wing pilots comprised 26 trainee pilots working towards a PPL or higher licence, 21 trainers (e.g., flight/ground instructors) and 12 other GA pilots (e.g., private/recreational and charter pilots). This study was conducted in accordance with ethics approval (BPS/04/13/HREC).

Materials

Although no separate headings were used, the survey questionnaire comprised three parts. The first part was to categorise participants into the three groups as described previously. The second part aimed to gain insights into how TEM training was delivered. The third part aimed to gather information about how general aviation pilots regarded the benefits and effectiveness of TEM training. This part contained seven common questions as presented in Table 1, as well as group-specific questions. The last question of the survey was an open-ended type, giving respondents an opportunity to share their opinions on TEM training. Some of the comments are presented, where appropriate, in the Results section to provide further illustration of the findings.

The survey questions in the third part had six response options: *strongly agree*; *agree*; *neither agree nor disagree*; *disagree*; *strongly disagree* and *unsure*. For the purpose of statistical analysis, numerical values of 'Strongly disagree' = 1, 'Disagree' = 2, 'Neither agree nor disagree' and 'Unsure' = 3, 'Agree' = 4 and 'Strongly agree' = 5 were assigned to each of the potential responses. The Statistical Package for the Social Sciences (SPSS, version 23) was used with the level of significance, alpha, set to be $p < .05$ for all statistical analyses.

Results

The results are presented in a combination of descriptive and inferential statistics. Prior to performing inferential statistics, the Kolmogorov-Smirnov test was completed to check whether the distribution of responses deviates from a comparable normal distribution (Field, 2013). When the responses were significantly non-normal, appropriate non-parametric tests (e.g., the Kruskal-Wallis test) were used to assess differences among independent sets of group scores.

Table 1. Seven common questions.

Q1. I feel that the number of accidents/incidents has reduced in my organisation after TEM training was introduced
Q2. I feel TEM training has improved overall aviation safety in general aviation
Q3. I use TEM principles on my day-to-day flying
Q4. I feel CASA produces adequate training and guidance materials for TEM training for general aviation
Q5. I feel appropriate use of TEM training improves my technical skills (i.e. aircraft handling skills)
Q6. I feel appropriate use of TEM training improves my non-technical skills such as situational awareness and decision-making
Q7. I feel the benefit of TEM training is over-rated

Delivery of TEM training

From a total of 59 participants, less than half of participants ($n = 23$) were engaged in flying activities before TEM training was mandated. All the participants from flight training organisations (six trainees and eight trainers) who were engaged in flying activities before the mandating of TEM training confirmed that briefing and/or training syllabi were amended to incorporate TEM training. It is worth noting that, according to the follow-up survey by the Australian Transport Safety Bureau (2009), approximately 60% of respondents indicated that their organisation had trained the majority of their staff in TEM with an additional 36% intending to train staff in the future.

A series of questions was asked to gauge how TEM training was delivered. The majority of trainees and trainers indicated that TEM training was incorporated in pre-/post-flight briefings, ground school and during instructional flights. For those other GA pilots, the main source of becoming familiar with TEM initiatives/training was from the flying organisations (e.g., aeroclubs) to which they belonged.

Common questions

There were seven common questions in the survey. All 59 participants (26 trainees, 21 trainers and 12 other GA pilots) answered the questions and they are analysed below.

The first common question (Q1) was to gauge participants' opinion as to whether TEM training had led to a reduction in number of accidents/incidents in their organisations. In comparison with the post-training survey ($N = 209$) administered by the Australian Transport Safety Bureau (2009), where 87% of respondents either agreed or strongly agreed TEM would improve safety in their organisations, responses for the present study showed that just over half of participants ($n = 31$) either agreed or strongly agreed that TEM training had reduced the number of accidents/incidents in their organisations.

The results from inferential statistics, the Kruskal-Wallis test due to significantly non-normal responses (Table 2) to determine any differences among the three groups (trainee pilots, trainers and other GA pilots), suggested that overall responses for Question 1 varied according to groups ($H(2) = 7.26$, $p = .027$). A post hoc, pairwise comparison with adjusted p-values showed that there were no significant differences between trainee pilots and other GA pilots ($p = .072$, $r = .29$) or between trainee pilots and trainer pilots ($p = 1.00$, $r = .07$). However, there was a significant difference in overall responses between other GA pilots and trainer groups ($p = .009$, $r = .34$). Descriptive statistics supported this difference where 15 trainers, in contrast to a mere two other GA pilots, either agreed or strongly agreed that they felt the number of

Table 2. Test of normality, descriptive statistics and non-parametric tests for seven common questions.

	Kolmogorov-Smirnov			Descriptive			Kruskal-Wallis		
Common questions	Statistic	*df**	Sig.	*N*	Mean	SD	*H*	*df*	*p*
Q1	.262	59	.000	59	3.58	0.72	7.26	2	.027
Q2	.322	59	.000	59	3.86	0.71	2.45	2	.294
Q3	.260	59	.000	59	4.07	0.74	7.65	2	.022
Q4	.215	59	.000	59	3.44	1.25	14.07	2	.001
Q5	.211	59	.000	59	3.58	0.97	0.97	2	.616
Q6	.257	59	.000	59	4.20	0.89	12.03	2	.002
Q7	.290	59	.000	59	2.42	1.13	14.62	2	.001

*df = the sample size ($N = 59$) (Field, 2013).

accidents/incidents had reduced in their organisation after TEM training was introduced. This suggested that the other GA pilot group did not strongly feel that introduction of TEM training reduced the number of accidents/incidents in their organisations.

The second common question (Q2) was to gauge participants' opinion as to whether TEM training has improved overall aviation safety in general aviation. When compared with the follow-up survey ($N = 72$) administered by the Australian Transport Safety Bureau (2009), where approximately 90% of respondents either agreed or strongly agreed that TEM would improve aviation safety, overall responses for the present study showed that approximately two-thirds of participants agreed or strongly agreed that TEM training had improved overall aviation safety in general aviation.

Responses for the second common question were analysed again using the Kruskal-Wallis test and the result showed that there were no significant differences among the three groups ($H(2) = 2.45$, $p = .294$). Descriptive statistics supported these similar responses among the three groups where 19 trainees, 18 trainers and seven other GA pilots either agreed or strongly agreed that they felt TEM training had improved overall aviation safety in general aviation. This is an encouraging finding, as safety initiatives in the area of non-technical skills seem to be positively received by pilots in general aviation.

Results for the third common question (Q3) showed that responses varied according to groups ($H(2) = 7.65$, $p = .022$). A post hoc, pairwise comparison with adjusted p-values showed that there were no significant differences between trainee and other GA pilot groups ($p = .425$, $r = .19$) or between trainee and trainer groups ($p = .316$, $r = .21$). However, there was a significant difference in overall responses between the other GA pilot and trainer groups ($p = .019$, $r = .36$), suggesting the other GA pilot group was less inclined to use TEM principles in their daily flying activities. A possible explanation for this different view could be due to these pilots not being aware of TEM training, as suggested in the written comment below:

> Never had any TEM training (formal) and not sure what is involves. (participant 23, private pilot)

This finding is concerning because CASA's legislation requires all licensed pilots to undergo a Biennial Flight Review (BFR), conducted every two years by appropriately qualified flight instructors or testing officers, that 'contributes to continuing pilot proficiency and consequently the safety of flight' by assessing a range of skills, knowledge and behaviour (CASA, 2010, p. 5). The Civil Aviation Advisory Publication (CAAP 5.81–1) specifies the mandatory requirement of demonstration and application of TEM principles whilst BFR is conducted (CASA, 2010).

Results for the fourth common question (Q4) indicated that overall responses were again varied according to groups ($H(2) = 14.07$, $p = .001$). A post hoc, pairwise comparison with adjusted p-values showed that there was no significant difference between trainer pilots and other GA pilots ($p = 1.00$, $r = .04$). However, there were significant differences in overall response between trainee pilots and other GA pilots ($p = .008$, $r = .39$) and between the trainee and trainer groups ($p = .004$, $r = .42$). Descriptive statistics supported these varied responses among the groups where only eight trainers and three other GA pilots, compared with 21 trainee pilots, agreed or strongly agreed that they felt CASA produced adequate training and guidance materials for TEM training for general aviation. Two written comments below offered possible reasons as to why there were these differences:

> I have participated in many CASA seminars on TEM but still find it hard to find 'relevant and real world' information on TEM. I don't believe the GA and instructional industry focus enough on TEM due to the lack of information. (participant 13, trainer)

> Not being aware of TEM until today suggests it's been poorly promoted. In spite of that, I feel my training of 43 years ago [e]nsured a culture of airmanship which probably informally provided the essence of TEM without the formal structure and possible regulatory overhead. (participant 22, private pilot)

Responses for the fifth common question (Q5) were analysed and results showed that there were no significant differences among the three groups ($H(2) = .97$, $p = .616$). The mean score of 3.58 ($SD = 0.97$) suggests that participants did not strongly feel that appropriate use of TEM training improved their technical skills (i.e., aircraft handling skills). This response was somewhat expected as TEM training is more widely regarded as training in the area of non-technical skills. However, anecdotal evidence suggesting that pilots who continuously receive TEM training as part of their flight training from the *ab initio* stage demonstrate earlier achievement of required competencies when compared with pilots who receive insufficient and/or irregular TEM training was not supported.

In contrast to the previous question, results for the sixth common question (Q6) indicated that overall responses varied according to groups ($H(2) = 12.03$, $p = .002$). A post hoc, pairwise comparison with adjusted p-values for the question showed that there was no significant difference between trainer and trainee groups ($p = 1.00$, $r = .11$). However, there were significant differences in overall responses between the other GA pilots and trainer groups ($p = .006$, $r = .4$) and between the other GA pilots and trainee groups ($p = .004$, $r = .42$), suggesting that views of the other GA pilot group were not as positive when TEM training was considered to improve the area of non-technical skills. This result was likely due to these other GA pilots not being aware of TEM training and hence they did not have an opportunity to regularly apply TEM principles despite the legislative requirements.

Results for the last common question (Q7) indicated that overall responses varied according to groups ($H(2) = 14.62$, $p = .001$). A post hoc, pairwise comparison with adjusted p-values for the question showed that there was no significant difference between the trainer and trainee groups ($p = 1.00$, $r = .11$). However, there were significant differences in overall responses between the other GA pilots and trainer groups ($p = .001$, $r = .48$) and between the other GA pilots and trainee groups ($p = .005$, $r = .41$). This finding again confirmed the view that the other GA pilot group did not appreciate as much the perceived benefit of TEM training. Written comments below reflect these mixed views on the perceived benefit of TEM training:

> TEM has greatly improved my situational awareness and has got me better prepared before every flight. (participant 1, trainee)

> TEM is only common sense and CASA tries to trick students with formal words instead of logic. (participant 45, charter pilot)

A trainer-specific question

In addition to these common questions, a trainer-specific (i.e., flight instructor) question was asked: 'I feel incorporation of TEM has improved the way that airmanship is taught'. The question aimed to gauge whether trainers considered TEM training as a mere replacement for traditional airmanship or a way of re-labelling airmanship. The majority ($n = 16$) either agreed or strongly agreed that the incorporation of TEM had improved the way that airmanship was taught and this was supported by a binomial test, $p = .027$ (2-tailed). This was an encouraging finding, as a positive perception by trainers of this training has a stronger likelihood of better transfer of training to trainees. This finding supported CASA's position that TEM training was more than the traditional role of airmanship, as it encouraged pilots to be more proactive in identifying and managing threats and errors even before their flight commenced (CASA, 2008).

Discussion

It has been more than five years since TEM training was introduced in Australia. However, there are no definitive data available to suggest whether TEM training is well-received and whether a positive effect of such training is witnessed, experienced or translated within the general aviation sector. The principal objective of this study was to fill the gap by exploring the effect of the introduction of TEM among Australian general aviation pilots.

The results from the two post-training and follow-up surveys before the mandating of TEM suggested that almost 90% of respondents either agreed or strongly agreed that TEM would improve safety in their organisations and a similar proportion of respondents either agreed or strongly agreed that TEM would improve aviation safety (Australian Transport Safety Bureau, 2009). However, the current study noted a large decline where just over half of participants either agreed or strongly agreed that they felt the number of accidents/incidents had reduced in their organisations after TEM training was introduced. Approximately two-thirds of participants agreed or strongly agreed that TEM training had improved overall aviation safety in general aviation. This alarming finding is consistent with other studies (e.g. Helmreich, Merritt, & Wilhelm, 1999) that there had been a gradual decay in positive attitude to Crew Resource Management (CRM) over time. Helmreich et al. (1999) noted that 'proceduralising CRM' may have caused such a decline and this might have been the case for the perception of participants in the current study after TEM training was mandated.

An encouraging finding noted that the majority of participants felt that TEM training had improved their non-technical skills. This result was somewhat expected as TEM principles were widely known as improving non-technical skills (e.g., decision-making). A concern was also raised where other GA pilots (including private pilots) did not share the similar view on this, considering the majority of aviation accidents are, at least in part, due to deficiencies in non-technical skills and private pilots flying fixed-wing aeroplanes account for the majority of general aviation accidents. Greater investment on offering formal, ongoing, human factors training courses in general aviation is much needed as the positive effects of extensive human factors training have been witnessed in the airline sector.

Another encouraging finding was that the majority of trainers agreed that incorporation of TEM training had improved the way that airmanship was taught. This suggests that the majority did not consider TEM training as a mere replacement for, or a fancy way of re-labelling, airmanship.

Conclusion

A major limitation of this study was the small sample size, particularly the private pilot group, resulting in statistical weakening of the generalisability of the findings. It was clearly identified from the literature review that private pilots flying a fixed-wing aeroplane in general aviation were the weakest link of the safety chain, deserving immediate attention. Thus, a larger sample size for the private pilot group would have yielded a fuller understanding of how TEM training was regarded by them, which in turn would have provided some direction regarding translation of the positive impact of TEM to the weakest link.

Overall results from this exploratory study indicated a variable uptake of TEM principles and differing opinions as to its effectiveness. Hence, further studies should be conducted on a larger scale to ascertain how TEM training is regarded and whether a positive effect of TEM training can be determined within Australian general aviation. Further studies in this area would also have the potential to provide information on pilot attitudes to safety initiatives more generally that would be of value in shaping future safety implementation in general aviation.

Since the emergence of TEM, it has been generally well-regarded as an effective method to improve aviation safety and TEM became a mandatory assessment item in various levels of flight test and ground examinations in Australia. However, there are no definitive data available to confirm this view and this exploratory study identified a variable uptake of TEM principles and differing opinions as to its effectiveness. It is therefore timely to examine more extensively TEM training and how it is regarded in a larger study with a strong focus on private pilots flying fixed-wing aeroplanes.

ORCID

Seung Yong Lee http://orcid.org/0000-0002-2662-639X

References

Australian Transport Safety Bureau (2009). *Threat and error management: Attitudes towards training and applicability of TEM to general aviation and low capacity air transport operations (AR-2006-156)*. Canberra, Australia: Author.

Australian Transport Safety Bureau (2011). *Aviation occurrence statistics 2001 to 2010*. Canberra, Australia: Author.

Besco, R. O. (1992). Analyzing knowledge deficiencies in pilot performance. *International Journal of Aviation Psychology, 2*, 53–74.

Bove, T. (2002). *Development and validation of a human error management taxonomy in air traffic control (Riso-R-1378[EN])*. Roskilde, Denmark: Roso National Laboratory.

Civil Aviation Safety Authority. (2008). *Teaching and assessing single-pilot human factors and threat and error management (CAAP 5.59-1)*. Canberra, Australia: Author.

Civil Aviation Safety Authority. (2010). *Flight crew licencing: Flight reviews (CAAP 5.81-1)*. Canberra, Australia: Author.

Dekker, S. (2006). *The field guide to understanding human error*. Surrey, England: Ashgate.

Earl, L., Bates, P. R., Murray, P. S., Glendon, I., & Creed, P. A. (2012). Developing a single-pilot line operations safety audit. *Aviation Psychology and Applied Human Factors, 2*, 49–61.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: Sage.

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *International Journal of Aviation Psychology, 9*, 19–32.

Hogan, R. L. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education and Development, 2*, 1–14.

Hunter, D. R. (1999). The general aviation pilot: Variety is the spice of flight. In D. O'Hare (Ed.), *Human performance in general aviation* (pp. 25–46). Aldershot, UK: Ashgate.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco, CA: Berrett-Koehler.

Maurino, D. E. (2005a). Threat and error management (TEM). Paper presented at the Canadian Aviation Safety Seminar (CASS), Vancouver, BC, 18-20 April.

Maurino, D. E. (2005b). TEM & LOSA: The state of affairs. Paper presented at the 3rd IATA/ICAO LOSA/TEM Conference, Kuala Lumpur, Malaysia, 13-14 September.
Maurino, D. E., & Murray, P. S. (2009). Crew resource management. In J. A. Wise, D. Hopkins, & D. J. Garland (Eds.), *Handbook of aviation human factors* (2nd ed., pp. 10-1–10-20). Boca Raton, FL: CRC.
Merritt, A., & Klinect, J. (2006). *Defensive flying for pilots: An introduction to threat and error management. The university of texas human factors research project: The Losa collaborative*. Austin, TX: The University of Texas.
National Transportation Safety Board. (2014). *Review of U.S. civil aviation accidents: Calendar year 2011*. Washington, DC: Author.
Reason, J. T. (2000). Human error: Models and management. *British Medical Journal, 320*, 768–770.
Sarter, N. B., & Alexander, H. M. (2000). Error types and related error detection mechanisms in aviation domain: An analysis of aviation safety reporting system incident reports. *The International Journal of Aviation Psychology, 10*, 189–206.
SPSS Statistics (22.0) [Computer software]. https://www-01.ibm.com/software/uk/analytics/spss/downloads.html
Thomas, M. J. W. (2004). *Error management training: Defining best practice (No. 2004/0500 – ATSB aviation safety research grant scheme project)*. Adelaide, South Australia: University of South Australia.
Wiegmann, D., & Shappell, S. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Aldershot, UK: Ashgate.

4.4 Chapter Summary

TEM is well regarded as an effective method for improving aviation safety, and has become a mandatory assessment item for various levels of flight tests and ground examinations in Australian aviation. Although it has been more than nine years since TEM training was introduced and implemented, there has been very limited formal review, assessment or evaluation to examine the after-state of TEM implementation, in terms of the effectiveness of introducing TEM. Therefore, this exploratory study was conducted to examine the ways in which the requirements for TEM training are addressed and regarded among Australian general aviation pilots in terms of its use and effectiveness. The findings indicated variable uptake of TEM principles and differing opinions regarding its effectiveness. Thus, these findings warranted further study in this area on a larger scale.

Chapter 5: Study 2 (Qualitative Study)

5.1 Purpose

This study involved a qualitative (first) phase of mixed methods research approach using an exploratory sequential design (Figure 3.6). This two-phase design involved a collection of qualitative data (study 2) followed by separately collecting quantitative data (study 3). The study aimed to gather insights before and after TEM was implemented to gain an in-depth appreciation as to how it was introduced and the effects of the use of TEM training. To fulfil this purpose, five highly experienced general aviation FEs were interviewed. The findings from this study were used when formulating survey questions for Study 3 (quantitative study) to verify and generalise the findings from the current study.

5.2 Methods

Crotty (1998) defined methods as 'the techniques or procedures used to gather and analyse data related to some research question or hypothesis' (p. 3). After reflecting on the epistemology and the theoretical perspective that informed the appropriate methodology in Chapter 3, the next logical step was the selection and adaptation of a specific design that best fit the research questions in the study. Methods should describe the process in detail, so that other researchers, who are not involved in the research, can replicate the same methods without difficulties (Crotty, 1998). Wiersma and Jurs (2009) added that qualitative research is more concerned with the comparability of the research findings. Comparability refers to providing an adequate description of the research procedures and theoretical constructs so that the findings can be better understood (Wiersma & Jurs, 2009). In this study, semi-structured interviews were used to gather information, and thematic analysis was conducted when analysing the interview data.

5.2.1 Interview.

An interview is defined as 'a face-to-face verbal exchange, in which one person, the interviewer, attempts to elicit information or expressions of opinion or belief from another person or persons' (Maccoby & Maccoby, 1954, p. 449). Interviewing has a long history and circumstances frequently arise in daily life in which certain forms of interviewing are used to obtain information and/or elicit an individual's feelings (e.g., visiting a medical clinic). The first use of interview as a method for a research activity was conducted by Charles Booth, whose research involved the seminal study of poverty in London in the 1880s (Morris, 2015). Despite its current ubiquitous use in daily life, the interview as a method—including in-depth interviews—did not gain reputation and traction as a useful method for qualitative research until the late 1930s or early 1940s (Lee, 2008). Nowadays, the interview is often at the core of qualitative research (Morris, 2015). Brinkmann (2017) highlighted the important point that the prevalence of this normalised sociocultural practice of interview in the postmodern age eliminated the need to justify why this is the best approach for research projects in many disciplines.

Typically, there are three forms of interviews: relatively unstructured form, relatively structured form and semi-structured form (Brinkmann, 2017). The word 'relatively' is included in the first two forms to highlight that absolutely structured and unstructured forms can only exist in theory, and not in practice (Brinkmann, 2017). Parker (2005) explained that, in a structured interview, complete control over interviewees' responses to the predetermined questions is not possible. Likewise, there is no absolute unstructured interview form (the other extreme) because, at the very least, interviewers have an idea of the topics to be discussed, and this frames an interview structure 'in accordance with certain specific conversational norms rather than others' (Brinkmann, 2017, p. 579).

Morris (2015) described a relatively structured interview as being closely aligned to the survey method, with the central premise that all structured interviews follow a predetermined and standardised format to ensure minimisation of interviewer variation and error. In this form of interview, the interviewer is expected to read, using the same tones, the pre-structured questions exactly the way they are written to every participant, without addition or omission. This passive way of recording people's opinions and attitudes on a specific topic may diminish the dialogical potential for knowledge production within conversations (Brinkmann, 2017).

In contrast, a relatively unstructured interview lies at the other end of the spectrum, as a predetermined list of questions is unnecessary and often cannot be prepared, and interviewees tend to have considerable control over the entire process (Morris, 2015). In this form of interview, the interviewer's main role is to be an active listener and facilitator, while not interrupting the interviewee after the opening request for a narrative (Brinkmann, 2017).

As previously mentioned, absolute structured and unstructured forms of interviews only exist in theory (Brinkmann, 2017). Therefore, Parker (2005) concluded that, in practice:

> an interview in qualitative research is always 'semi-structured' because it invariably carries the traces of patterns of power that hold things in place and it reveals an interviewee's, a co-researcher's, creative abilities to refuse and resist what a researcher wants to happen. (p. 53)

The current study adopted a semi-structured in-depth interview as the preferred method.

5.2.2 Semi-structured in-depth interview.

Brinkmann and Kvale (2015) defined the semi-structured interview as 'an interview with the purpose of obtaining descriptions of the life [real] world of the interviewee in order to interpret the meaning of the described phenomena' (p. 6). As previously established, and as the term suggests, the semi-structured interview—the preferred method of data collection in this study—is located somewhere between the relatively structured and unstructured interview forms. Morris (2015) explained that the exact location of the semi-structured interview on the spectrum is dependent on many factors, such as interviewees' personality and experience, and the level of probing required. Similar to the relatively structured interview form, a list of predetermined questions is available to the interviewer, yet the interview format allows flexibility, and permits interviewers and interviewees to digress, which may promote a detailed discussion to facilitate in-depth understanding and clarification (Morris, 2015). This can enhance the quality and relevance of the data collected through the knowledge-producing conversational process. In addition, this conversational interview, with purpose and built-in flexibility, will lead to a further enhanced understanding of the topic the researchers are examining.

In-depth interviewing is the most common qualitative research method, as it is a powerful way to gather detailed data (Morris, 2015). In-depth interviewing generally involves two people (interviewer and interviewee) having a conversation on a research topic about which the interviewee has expertise. Therefore, the interviewee is at the centre of the interview, and the role of interviewer is to ensure that as much information as possible is extracted by asking probing questions that provide elaborated and further clarified detail on the research topic (Morris, 2015). As aforementioned, in-depth interviewing is an effective method to gather data on a research topic and, although there are some weaknesses, there are a number of strengths that contribute to this. A list of strengths and weaknesses is presented below (Figure 5.1).

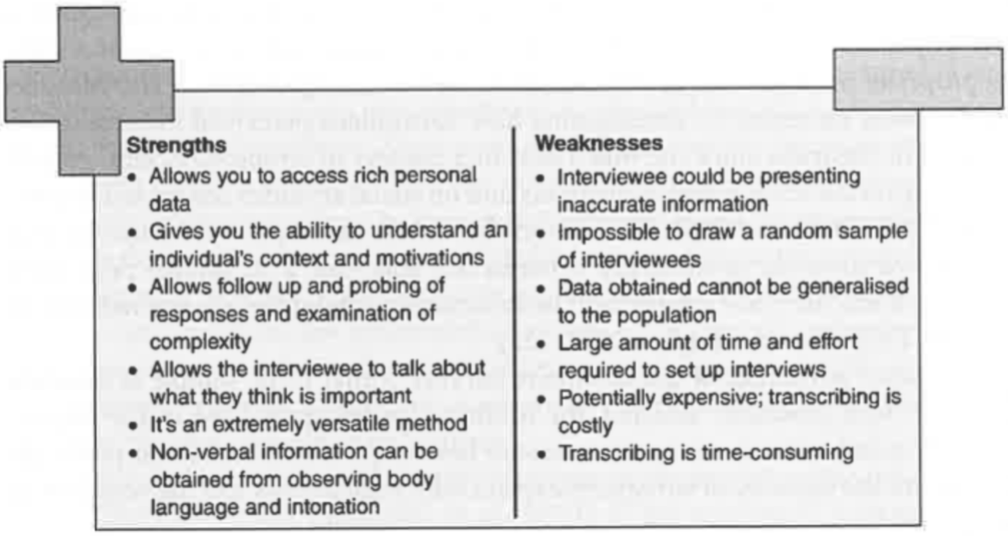

Figure 5.1. Strengths and weaknesses of in-depth interviews. Source: adapted from Morris (2015, p. 7).

Some of the weaknesses of in-depth interviews are inevitable (e.g., a large amount of time and effort is required, and transcribing is time consuming); however, other weaknesses (such as inability to generalise findings) can be overcome by using an appropriate methodology, such as an adoption of mixed methods.

The central purpose of this study was to gather insights on the TEM implementation process in Australian general aviation, and its effect on the practice of TEM. With this in mind, conducting an in-depth semi-structured interview with highly experienced FEs in general aviation was considered the most appropriate approach to gain the required knowledge and understanding.

5.2.3 Trustworthiness in research.

The purpose of research is for the research findings to advance the body of knowledge and/or provide solutions to specific issues. To achieve this, the findings need to be accepted by the intended audiences. Trustworthiness is one of the ways to persuade the intended audiences that the findings are 'worth paying attention to' (Lincoln & Guba, 1985, p. 290). According to Lincoln and Guba (1985), the assessment criteria used to achieve trustworthiness are truth value, applicability, consistency and neutrality. These criteria appear differently between quantitative and qualitative research because of differences in epistemology and theoretical perspectives in the two research approaches. This section provides a brief description of each assessment criterion applicable to qualitative research (Table 5.1), followed by more detailed description in Section 5.2.3.1.

Table 5.1 *Assessment Criteria for Trustworthiness in Qualitative Research.*

Criterion	Qualitative approach
Truth value	Credibility
Applicability	Transferability
Consistency	Dependability
Neutrality	Conformability

Source: adapted from Krefting (1991, p. 217).

Truth value refers to researchers' confidence in the truth of the findings, based on research design, participants and context (Krefting, 1991). In qualitative research, credibility takes the position of truth value. It is subject oriented in that it endeavours to discover the experience of participants as they have lived and perceived it (Krefting, 1991). Study 2 in this chapter used in-depth semi-structured interviews to elicit and gather information from five highly experienced FEs in Australian general aviation, and this was deemed to satisfy the assessment criterion.

Applicability refers to how successfully the findings can be generalised to alternative contexts, settings or groups (Krefting, 1991). In qualitative research, applicability is termed 'transferability' and this is supported when a researcher provides rich descriptive data about how particular research was conducted (Lincoln & Guba, 1985). Krefting (1991) added that generalisation of research findings is not relevant in qualitative research because this type of research is typically conducted in naturalistic settings with fewer controlling variables. Throughout this chapter, detailed descriptions of the methods used in this study are provided to satisfy the assessment criterion. Moreover, the follow-up study (Study 3) in Chapter 5 generalises the findings from the present study (Study 2), which also satisfies the assessment criterion of applicability.

Consistency is the third assessment criterion of trustworthiness applicable to qualitative research, and considers the consistency of the findings if a research study is repeated involving the same samples of participants or in a similar context (Krefting, 1991). In qualitative research, dependability is an equivalent term to consistency. According to Krefting (1991), the findings from a repeated study with the same participants may differ from the initial findings. Lincoln and Guba (1985) described this by using an analogy that 'one can never cross the same stream twice' (p. 299). Despite this, detailed descriptions of the methods used in a qualitative study are important, and these are provided throughout this chapter.

Neutrality is the fourth and final criterion of trustworthiness that determines the legitimacy of the findings of researchers without 'biases, motivation, interests or perspectives' (Lincoln & Guba, 1985, p. 290). Lincoln and Guba (1985) argued that, when determining the neutrality of qualitative research, focus should be placed on the data, rather than the researchers, as it is inevitable for qualitative researchers to be distanced from their subjects. Conformability is the alternative criterion for neutrality, which is achieved when credibility and transferability are established (Krefting, 1991).

Therefore, in-depth semi-structured interviews with highly experienced FEs in Australian general aviation, as well as provision of detailed descriptions of the methods and procedures used in this study, were considered to satisfy neutrality in this qualitative study.

5.2.3.1 Trustworthiness in qualitative research.

The previous section included a brief description of the four assessment criteria employed to ensure trustworthiness (credibility, transferability, dependability and conformability) in qualitative research. Lincoln and Guba (1985) offered a number of techniques to achieve these assessment criteria, and a summary of the techniques is provided below (Table 5.2). This section expands on each assessment criterion, including their respective techniques and how the current study satisfies the assessment criteria.

Table 5.2 *Summary of Techniques for Establishing Trustworthiness.*

Criterion	Technique
Credibility	Activities in the field that increase the probability of high credibility (prolonged engagement, persistent observation and triangulation—sources, methods and investigators) Peer debriefing Negative case analysis Referential adequacy Member checks (in progress and terminal)
Transferability	Thick description
Dependability	Dependability audit, including the audit trail
Conformability	Conformability audit, including the audit trail

Source: adapted from Lincoln and Guba (1985, p. 328).

Credibility is the first assessment criterion in qualitative research to satisfy trustworthiness (truth value). Krefting (1991) described credibility as a subject-oriented assessment criterion that endeavours to discover the experience of participants. Lincoln and Guba (1985, p. 301) offered five techniques to enhance the credibility of qualitative research:

1. an activity that increases the likelihood that credible findings and interpretations will be produced

2. an activity that provides an external check of the enquiry process
3. an activity aimed to refine working hypotheses as more and more information becomes available
4. an activity that renders it possible to check preliminary findings and interpretations against archived 'raw data'
5. an activity that enables direct testing of the findings and interpretations with the human sources from which they came—the constructors of the multiple realities being studied.

The first technique to achieve a higher probability of producing credible findings and interpretations involves three activities: 'prolonged engagement, persistent observation and triangulation' (Lincoln & Guba, 1985, p. 301). Prolonged engagement involves adequate submersion in the research settings to become intimately familiar with a research domain (Krefting, 1991), so that researchers can identify distortions—intended or accidental—from research participants and build trust, so that more sensitive information can be shared between researchers and participants (Lincoln & Guba, 1985). The researcher of the present study already knew three participants well before the interviews were conducted, and the other two participants were familiar with the researcher's background and extensive experience and involvement as an FI in general aviation. Therefore, prolonged engagement was achieved relatively easily for the researcher to build trust and develop an ability to identify any distortions that would adversely affect the trustworthiness of the data and findings.

Persistent observation is the second activity that Lincoln and Guba (1985) offered to achieve a higher probability of producing credible findings. The central purpose of persistent observation is extensively searching to identify relevant characteristics and elements that are the core aspects of the research in question, and delving more deeply into them (Lincoln & Guba, 1985). A simple description to highlight the major difference between prolonged engagement and persistent observation is simply that the former provides scope, while the latter adds depth of enquiry (Lincoln & Guba, 1985). The researcher of the present study has almost 20 years of experience in general aviation in Australia and New Zealand, during which the researcher observed a number of changes, including the introduction of TEM as a licence requirement in both countries. This prolonged engagement and observation opportunity in general aviation is expected to further enhance the credibility of the findings and accurate interpretation.

Triangulation is the third activity to improve the probability of producing credible findings and their interpretation (Lincoln & Guba, 1985). Neuman (2011) described triangulation as searching for answers to research questions with improved accuracy. There are a few ways to achieve triangulation. Two common forms are triangulation of measures and triangulation of method (Neuman, 2011). Triangulation of measures refers to collecting and measuring multiple datasets (e.g., several quantitative or qualitative data or both quantitative and qualitative data) in a single study. Triangulation of method refers to combining quantitative and qualitative approaches and their data. The present study adopted mixed methods as a preferred methodology, and sequential exploratory design as a preferred method. Therefore, it was expected that the combined findings and their interpretations would be more credible through triangulation.

The second technique to improve the credibility of qualitative research involves an activity (i.e., peer debriefing) conducted to review and discuss insights (e.g., research process and findings) of the particular research (Krefting, 1991; Lincoln & Guba, 1985). Krefting (1991) suggested employing an impartial colleague for this purpose. This chapter, as well as the other chapters of this thesis, were thoroughly proofread by a staff member within the university, who was independent from this study. This enhanced the credibility of this study and its interpretations and findings.

The third technique to increase the credibility of qualitative research is employing an activity (i.e., negative case analysis) that involves constant update and refinement of the working hypotheses as more information progressively becomes available (Lincoln & Guba, 1985). Brodsky (2008) described negative case analysis as a central data analytic approach in qualitative research involving appropriate management of a negative case (e.g., new data), so that the credibility of the research findings is supported, and the researcher biases in data analysis and theory generation are minimised. Broadsky (2008) further added that, after identifying negative cases, there are three steps to appropriately address them: understanding where and how these new data diverged from a working hypothesis, making appropriate revisions to address these new data in the revised hypothesis, and testing the applicability of the revised hypothesis in the overall research results. This process naturally occurred in the current study, particularly during a data analysis phase (thematic analysis with several iterations), and was appropriately addressed in accordance with Broadsky's advice.

The fourth technique to increase the credibility of qualitative research is engaging in an activity that involves examining preliminary findings and interpretations against archived raw data (Lincoln & Guba, 1985). Use of thematic analysis was a suitable activity to increase the credibility of this study. Thematic analysis involves identifying, analysing and reporting themes and subthemes within the data as researchers follow a six-phase, recursive, two-way linear process to ensure the comprehensiveness and thoroughness of the analysis process (Braun & Clarke, 2006). Through this process, the accuracy of the findings and interpretations are checked against raw data as they are fine-tuned. This study followed the six-phase process suggested by Braun and Clarke (2006).

Another way to satisfy this technique to increase the credibility of qualitative research is establishing the auditability of the research. Auditability suggests that, with an appropriate volume of information, replicated research will have a greater chance of attaining similar findings (Krefting, 1991), although Lincoln and Guba (1985) argued that 'one can never cross the same stream twice' (p. 199). Despite this, Shenton (2003) argued that researchers should endeavour to achieve repeatability—for example, by providing detailed descriptions of the phenomenon under study and presenting relevant documents (e.g., transcriptions, field notes or staged data analysis files) available for inspection if required. The current researcher provided a detailed description of the process used in this qualitative study in the present chapter, and saved all relevant documents in his computer, with a back-up copy stored in the university-approved Cloud storage for future reference.

According to Shenton (2004), the fifth technique to increase the credibility of qualitative research involves an activity (i.e., member checking) that provides an opportunity for research participants to confirm the accuracy of data (e.g., transcription) and the researcher's understanding of the phenomenon under study during conversation (e.g., interview). Lincoln and Guba (1985) considered this the most important technique for establishing credibility. In this study, after each interview, the entire recording was transcribed by the researcher and the completed transcripts were then returned to each respective participant for their perusal and correction. Two participants returned their transcripts with a few changes. The rest were satisfied with the transcriptions without alteration. In addition, this study adopted the semi-structured in-depth interview as the preferred method, as this type of interview provides the flexibility for interviewers and interviewees to digress, thereby promoting a detailed discussion to enable in-depth understanding and better clarification (Morris, 2015). Further, in-depth interviewing with participants allowed the researcher to ask probing questions that provided elaborated and more clarified detail about how TEM is implemented and practised in Australian general aviation. Therefore, credibility was appropriately and adequately established in this qualitative phase of the study.

Transferability is the second assessment criterion in qualitative research to satisfy trustworthiness, which, according to Lincoln and Guba (1985), is almost impossible to achieve in qualitative research. Shenton (2004) explained that, because the findings in qualitative research tend to derive from small specific contexts, demonstration of such findings and conclusions in other situations and/or populations is not achievable. However, despite the apparent impossibility of conventional generalisability in qualitative research, Shenton (2004) argued that it is important to provide sufficient contextual information, so that readers of the study can make informed decisions about the appropriateness and applicability of transferring such results and conclusions in other settings. The present study provides detailed information, such as information about the research process and interpretation of the findings. In addition, the aim of the present study was to identify key themes from which survey questionnaires could be devised for the subsequent quantitative study to enable generalisation of the findings. To this end, transferability was adequately addressed in the overall design of research in this thesis.

Dependability is the third assessment criterion in qualitative research to satisfy trustworthiness, which is very closely tied to credibility (Lincoln & Guba, 1985). This co-requisite nature of creditability and dependability suggests that adequately satisfying the former will also sufficiently satisfy the latter (Lincoln & Guba, 1985).

In this regard, the present study satisfied this assessment criterion. In addition, provision of detailed information—such as information on the research process and how the findings are interpreted—plays an important role in satisfying dependability. Further, Lincoln and Guba (1985) offered that overlapping and a dependability audit can be performed to further improve dependability. In the present study, as previously mentioned, detailed descriptions—such as information on the research process and interpretation of findings—are presented throughout this chapter. In addition, this study adopted mixed methods as the preferred methodology, which anticipates that the overall findings and their interpretation are more likely to be credible through overlapping data (triangulation).

Confirmability is the fourth assessment criterion in qualitative research to satisfy trustworthiness. Lincoln and Guba (1985) offered two courses of action to address confirmability: auditability and triangulation of methods and data sources. As previously mentioned, auditability suggests that, with an appropriate volume of information, replicated research will have a greater chance of attaining similar findings (Krefting, 1991). To this end, qualitative researchers provide detailed descriptions of the phenomenon under study (Shenton, 2004) and present relevant documents (e.g., transcriptions, field notes or staged data analysis files) available for inspection if required. In this regard, the researcher provided detailed descriptions of the process of this qualitative study in this chapter, and saved all relevant documents in his computer, with a back-up copy stored in the university-approved Cloud storage for a possible audit. In addition, this study is part of a series of studies that adopted mixed methods, which sufficiently satisfied triangulation to further improve the confirmability of the overall findings and conclusions. In addition, conformability was deemed to have been established in this study, as creditability and transferability were satisfied.

This section has described in detail the four assessment criteria to ensure trustworthiness in qualitative research, and provided certain techniques to further improve the assessment criteria. This section has also provided rationales regarding how each criterion was addressed and satisfied in this study. The following sections provide detailed descriptions on the participants and process of this study.

5.2.4 Participants.

The central purpose of this phase of the study was to gain insights before and after TEM was implemented, so that the researcher had a better understanding of how TEM was implemented, and its effect on TEM practice in Australian general aviation. The ideal candidates were individuals who were part of the Civil Aviation Safety Authority—the regulatory body for Australian aviation—with extensive experience in general aviation that included, but was not limited to, flight examinations. Given the need for these highly experienced participants, there were only limited numbers of potential participants available.

To recruit suitably experienced participants for this study, the researcher retrieved a list of Approved Testing Officers (ATO) who had delegation to conduct certain flight tests from the CASA website. The list provided each ATO's contact details (telephone number and email address), as well as flight test permissions that indicated which flight tests the ATOs were authorised to conduct (e.g., CPL and instrument rating). Initially, the researcher sent an email to an ATO who was the researcher's former colleague, asking him to take part in the study. After this interview, the researcher asked the participant to identify other suitable potential participants from the list. A total of five[2] further suitable candidates were identified within a radius of 150 km from the location where the researcher resided. Each potential participant was contacted via email and follow-up telephone calls. One of the participants advised via a return email that he had been retired for some time, and showed little interest in participating in this research. Therefore, no further attempt was made to recruit him. The remaining candidates all agreed to participate in this study.

All participants had considerable experience in both flight training and testing. Their combined experience in general aviation represented over 180 years, and two of the participants also had extensive military flight instructional experience. In addition, the majority of participants had extensive experience in flight testing over a wide

[2] After an email was sent to potential participants, the researcher learnt that one of the participants who agreed to participate no longer resided in Brisbane. However, given this individual's vast experience in the regulatory environment and flight training in both the military and general aviation sectors, it was agreed to conduct an interview via Skype.

geographical area in Australia. A summary of the participants' relevant experience and qualifications in general aviation is provided below:

- all participants had considerable experience as FIs and FEs, and two participants had extensive military flight instructional experience
- all participants were authorised to conduct a PPL flight test
- four participants were authorised to conduct a CPL flight test
- four participants were authorised to conduct a Command Instrument Rating flight test
- three participants were authorised to conduct a FI rating flight test
- four participants had previously worked for CASA
- four participants were currently exercising flight testing functions
- three participants were either a Chief Flying Instructor (CFI) or a Chief Pilot (CP) when TEM was mandated.

5.2.5 Procedures.

Each potential participant was initially contacted by email. In some cases, a follow-up email and/or telephone call was sent/made as a reminder. When approaching each potential participant, two documents—both approved by the university's ethics committee—were attached in the email. The first document was a participant information sheet (Appendix 5) that provided the potential participants with adequate information about the intended study, and offered an opportunity for them to ask questions and/or clarify any concerns. The second document was a consent form (Appendix 6) that represented the participants' formal agreement to participate in the research. No participants withdrew from the study.

On the day of the interview, hardcopies of the participant information sheet and consent form were provided by the researcher, who discussed them with the participants. Given that no signed consent form was received prior to the scheduled interview, this form was signed on the day of the interview, before the interview commenced. Three participants elected to be interviewed at their workplace when they had no work commitments. The fourth participant preferred to be interviewed at a cafeteria at an airport. The final participant was interviewed via Skype—one of the 'synchronous (real-time) methods' that a growing number of qualitative researchers are using as 'an alternative or supplemental choice' for interviewing (Janghorban, Roudsari, & Taghipour, 2014, para. 2).

On average, the interviews took approximately one hour. All interviews were audio recorded. All interview recordings were transcribed by the researcher and the completed transcriptions were sent back to each respective participant via email for verification purposes. This also allowed the participants to remove any transcribed conversation they did not wish to be used for analysis. Two participants returned their transcripts with a few changes. The rest of the participants were satisfied with the transcriptions without alteration. Part of the transcriptions is presented in Appendix 7. All data collected have been stored in the QRIScloud storage provided by the Queensland Cyber Infrastructure Foundation in accordance with the university's research data management policy.

5.2.6 Transcribing interview recordings.

Broadly, there are two ways in which interview data are transcribed: through a commercial transcribing service or self-transcribing. Each approach has inherent advantages and disadvantages. The use of a commercial transcribing service is quicker, yet more expensive. Self-transcribing allows researchers to have a better understanding of the data collected, yet a large time investment is required to transcribe interview data, particularly for individuals who are not native English speakers and/or are very new to transcribing interview recordings. Despite this, the researcher decided to transcribe all five pieces of interview recordings himself, after seeking advice from the relevant literature and other qualitative researchers, who suggested that self-transcribing allows the researcher to be more familiar with the interview data.

On average, each interview transcription was about 20 pages. Given that the researcher is not a native English speaker, and to offer participants the opportunity to review their interview in a written form, a copy of the transcript was returned to each respective participant for their perusal, correction, addition and endorsement prior to inclusion in this study. No participant objected to the inclusion of their transcribed data, and two participants returned the transcripts with very minor typographical errors corrected and comments added. These were incorporated before the transcripts were analysed.

5.3 Qualitative Data Analysis Method

Qualitative data are often referred to as data in the form of words, and are typically collected through interviews, observations, documents and artefacts (Miles, Huberman, & Saldana, 2014). These rich and holistic qualitative data, collected in naturally occurring settings, have the potential to provide in-depth appreciation of real issues beyond a mere description of ‘what’ and ‘how many’ (Miles et al., 2014). Prior to analysing qualitative data, they need to be appropriately prepared, and such preparation typically involves transcribing audio recordings, expanding raw field notes and saving the data in a computer using Computer-Assisted Qualitative Data Analysis Software (CAQDAS). Section 5.3.2 provides in-depth descriptions of CAQDAS.

Data analysis is ‘a systematic search for meaning’ through which the information that was learnt can be effectively and structurally communicated to intended audiences (Hatch, 2002, p. 148). It involves ‘organising and interrogating data in ways that allow researchers to see patterns, identify themes, discover relationships, develop explanations, make interpretations, mount critiques, or generate theories’ (Hatch, 2002, p. 148). Miles et al. (2014) categorised this iterative concurrent flow of activity into three phases: data condensation, data display, and drawing and verifying conclusions. These interactive and cyclical processes occur before, during and after data collection (Figure 5.2).

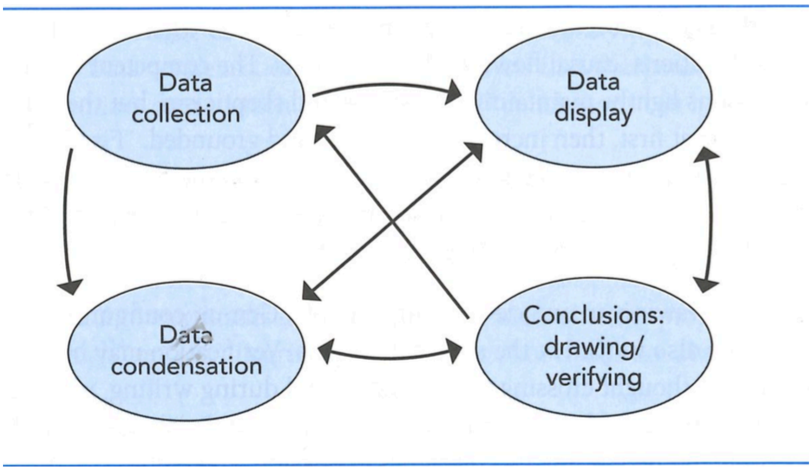

Figure 5.2. Components of data analysis: interactive model. Source: adapted from Miles et al. (2014, p. 14).

The first phase of data analysis is data condensation. Miles et al. (2014, p. 12) defined data condensation as 'the process of selecting, focusing, simplifying, abstracting, and/or transforming the data that appear in the full corpus (body) of written-up field notes, interview transcripts, documents, and other empirical materials'. This process is embedded within the full lifecycle of any qualitatively oriented project from the beginning of research planning to anticipatory data condensation, until the completion of the final report (Miles et al., 2014). For instance, the process informs qualitative researchers on selecting appropriate research design and questions, data collection methods, and other associated works, often without researchers being fully aware of this (Miles et al., 2014). The process continues and affects researchers' and/or data analysts' decisions when collected data are analysed through several iterations of coding, theming and writing a final draft that draws final conclusions (Miles et al., 2014).

The second phase, which is interwoven with the first and third phases, is termed 'data display'. This activity refers to 'an organised, compressed assembly of information that allows conclusion drawing and action' (Miles et al., 2014, pp. 12–13). Research findings presented in a long list of numbers and/or textual information are very difficult for researchers/data analysts to examine during data analysis, and even more difficult for readers to understand and draw conclusions based on them. Therefore, clear display of organised information in the form of images—such as graphs, charts and/or networks—is essential for researchers/data analysts to draw justified conclusions or inform them to conduct further analyses (Miles et al., 2014). Miles et al. (2014) further added that this phase is a ‘major avenue to robust qualitative analysis’ (p. 13).

The third phase, which again continuously interacts with the first two phases, is drawing and verifying conclusions. Preliminary conclusions are drawn as the data are collected and data analyses proceed. The justifiable final conclusions are reached after the transformation of data, and are verified through various methods, such as crosschecking field notes and/or memos for their confirmation (Miles et al., 2014). This improves the overall reliability and validity of qualitative research, as described above.

In summary, the coding of data, as part of qualitative data analysis, leads researchers/data analysts to categorise data and subsequently identify themes and their relationships. These themes may be clearly displayed in the form of graphs and/or charts for researchers/data analysts to draw justified conclusions. Alternatively, they may be used to conduct further analyses (Miles et al., 2014). Final conclusions are then drawn and verified, which improves the overall robustness of qualitative data analysis.

There are many different data analysis methods that can be adopted when analysing qualitative data. One way to categorise the various types of analysis methods is to consider whether a method is tied to a specific theoretical or epistemological position (Braun & Clarke, 2006). This leads to ‘relatively limited variability in how the method’ is applied or whether the method is independent of theoretical or epistemological positions that are compatible with various schools of thought, while offering greater flexibility (Braun & Clarke, 2006, p. 79). Thematic analysis falls in the latter category (Braun & Clarke, 2006), and this was the preferred method for the qualitative data analysis of the current study because of its flexibility and close alignment to the theoretical perspective of pragmatism.

Braun and Clarke (2006) argued that, although there are some overlaps and differences among various types of qualitative data analysis methods, thematic analysis is nevertheless a ‘foundational method for qualitative analysis’ (p. 78). Therefore, this is the first qualitative data analysis method that commencing qualitative researchers should learn, so that the core skills that researchers gather while undertaking thematic analysis will provide a solid foundation when later conducting other forms of qualitative data analysis (Braun & Clarke, 2006). One such generic core skill is thematising the meanings of collected qualitative data (Holloway & Todres, 2003). Braun and Clarke (2006) further remarked that thematic analysis offers theoretical flexibility when analysing qualitative data, and this was closely aligned to the epistemology (pragmatism) of the methodology in this thesis.

5.3.1 Thematic analysis.

Thematic analysis involves identifying, analysing and reporting themes within data (Braun & Clarke, 2006). A theme refers to a pattern that typically appears across the dataset that relates to the research questions. Braun and Clarke (2006) cautioned that prevalence of a theme in terms of space and frequency is not necessarily dependent on a quantifiable measure. Rather, the researcher must identify meaningful themes, often with subthemes, that accurately and adequately describe a phenomenon in relation to a specific research question. To achieve this, Braun and Clarke (2006, p. 87) suggested the following six phases—presented in Table 5.3—with the emphasis that these phases do not necessarily suggest a linear process from one phase to another. Instead, the process is recursive, encouraging researchers/analysts to move back and forth to enable comprehensiveness and thoroughness in the analysis process (Braun & Clarke, 2006).

Table 5.3 *Six Phases of Thematic Analysis.*

Phase	Description of the process
1. Familiarising oneself with one's data	Transcribing data (if necessary), reading and re-reading the data, and noting initial ideas
2. Generating initial codes	Coding interesting features of the data in a systematic manner across the entire dataset, and collating data relevant to each code
3. Searching for themes	Collating codes into potential themes, and gathering all data relevant to each potential theme
4. Reviewing themes	Checking whether the themes work in relation to the coded extracts (Level 1) and the entire dataset (Level 2), and generating a thematic 'map' of the analysis

5.	Defining and naming themes	Ongoing analysis to refine the specifics of each theme and the overall story the analysis tells, and generating clear definitions and names for each theme
6.	Producing the report	Final opportunity for analysis—selection of vivid, compelling extract examples; final analysis of selected extracts; relating back the analysis to the research question and literature; and producing a scholarly report of the analysis

Source: Adapted from Braun and Clarke (2006, p. 87).

Prior to describing each phase, it is important to establish a working definition of a number of terms, as presented below (Braun & Clarke, 2006, p. 79):

- The 'data corpus' refers to all data collected for a particular research project.
- A 'dataset' refers to all the data from the corpus that are being used for a particular analysis.
- 'Data items' refer to each individual piece of data collected, which together make up the dataset or corpus.
- 'Data extract' refers to an individual coded chunk of data that has been identified within, and extracted from, a data item.

The first phase of familiarising oneself with one's data involves reading and, often, re-reading the written data until the researcher is familiar with 'the depth and breadth of the content' (Braun & Clarke, 2006, p. 87). It was noted that researchers might be tempted to skip or fast track this phase, particularly if the researchers collected and transcribed the interview data themselves, as they have some level of familiarity with the collected data (Braun & Clarke, 2006). However, this first phase is the foundation of thematic analysis. Therefore, repeated reading of the data and active searching for meanings and patterns by making rough notes and writing early impressions is essential (Maguire & Delahunt, 2017). The principal researcher collected and transcribed the interview data, yet still considered this advice by perusing the entire data corpus twice and then generating notes and ideas on the third reading, before moving to the second phase. During the first phase, the following ideas, as an example, were noted.

- TEM has been poorly introduced and implemented.
- CASA, a regulatory body for civil aviation in Australia, has an inadequate understanding of general aviation.
- There are differences among the FEs when assessing TEM elements during flight tests.

The second phase of thematic analysis involves organising data in a systematic way and producing initial codes based on each segment of data that relate to research questions (Maguire & Delahunt, 2017). A code is defined as 'a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data' (Johnny, 2009, p. 3). Table 5.4 below presents examples of data extracts and codes from the transcribed interview data. Braun and Clarke (2006, p. 88) described coding to be either 'data driven' or 'theory driven'—that is, explorative or explanative, respectively. The findings from the pilot study (Study 1) indicated a variable uptake of TEM principles and differing opinions regarding TEM's effectiveness. As such, this study adopted an explorative sequential design data-driven approach, which was considered when producing initial codes.

Table 5.4 *Examples of Data Extracts and Codes from the Second Phase of Thematic Analysis.*

Data extract	Code
'Industry needed, certainly from the outset, an education roadshow to go around the regional centres, set up workshops and explain what it was, what they wanted, what outcomes they wanted out of TEM implementation. So, no, it was not well done. It left the industry pretty much on their own devices' (Participant 1).	TEM was poorly introduced and implemented
'A regulator that is quite removed from industry, so if you look at the majority of people making the rules, or publishing the guidance material, they are many, many, many, many years removed from the cold face and things are quite often done reactively, rather than proactively, and without a huge amount of engagement' (Participant 4).	CASA's understanding of general aviation
'…to this day, I would be fairly confident in saying that you could ask 15 inspectors at CASA and you'll have 15 different opinions on threat and error management and their understanding' (Participant 4).	Differences among examiners when assessing TEM

The third phase of searching for themes commences after a long list of initial codes is produced. It involves grouping the different codes into potential themes and collating relevant data extracts accordingly (Braun & Clarke, 2006). There may be several main themes, and usually more subthemes. It was suggested that the use of visual representation can help researchers organise the different codes into appropriate theme piles (Braun & Clarke, 2006). As an example, Figure 5.3 below displays a preliminary thematic map that contains some of this study's themes and subthemes. This preliminary phase of the theming exercise finishes after all data extracts are coded and sorted into appropriate candidate themes (Maguire & Delahunt, 2017).

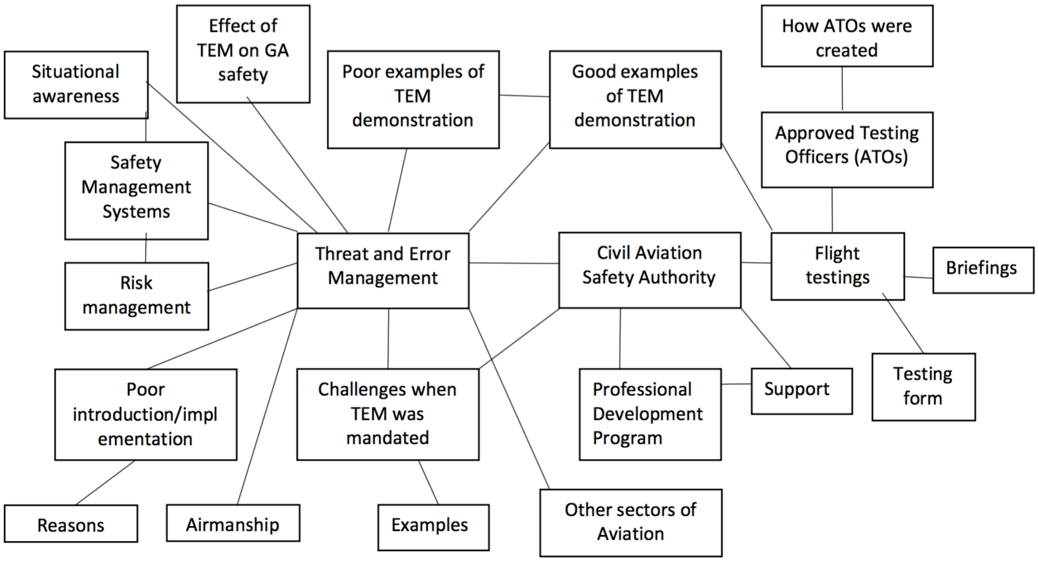

Figure 5.3. Preliminary thematic map of candidate themes of this research.

The fourth phase of reviewing themes involves a two-tiered approach of review and refinement of the candidate themes identified during the third phase, and these further undergo discarding, merging and/or dividing themes and subthemes (Braun & Clarke, 2006). Two criteria that can be used for this review and refinement phase are internal homogeneity and external heterogeneity (Patton, 1990). Internal homogeneity 'concerns the extent to which the data that belong in a certain [theme] hold together or "dovetail" in a meaningful way' (Patton, 1987, p. 154). This criterion is applied to the first-tiered approach of review and refinement in a localised sense, whereby the data extracts in each theme are carefully reviewed and refined for internal coherence and consistency. The second criterion of external heterogeneity 'concerns the extent to which differences among [themes] are bold and clear' (Patton, 1987, p. 154). This is applied to the second-tiered approach of review and refinement in a more global sense that considers datasets' and themes' distinction from one another. At the end of this phase, further refined thematic maps were developed based on their major themes. As an example, a developed thematic map based on this study's first main theme is presented in Figure 5.4.

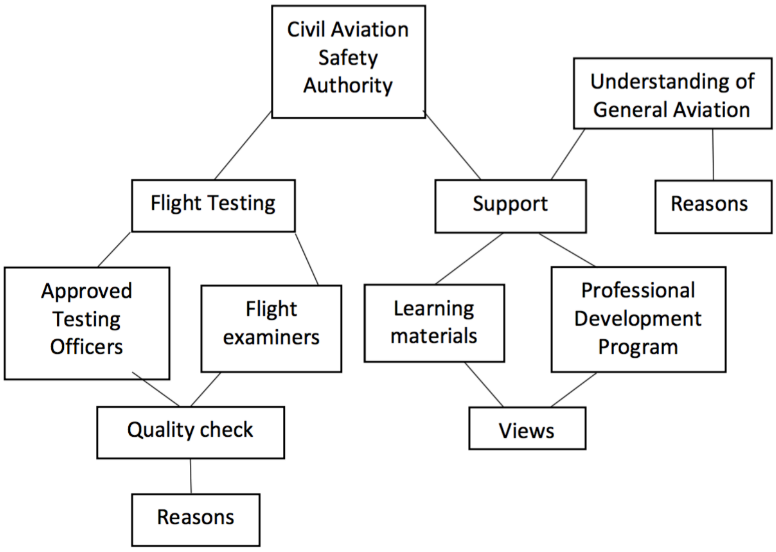

Figure 5.4. Developed thematic map on this study's first main theme—CASA.

The fifth phase of defining and naming themes involves 'identifying the essence' of the further defined themes' meaning, both individually and collectively, as well as 'determining' the specific aspects of the data each theme delivers (Braun & Clarke, 2006, p. 92). In doing so, it can be established what each theme is about as well as how themes relate to one another (Maguire & Delahunt, 2017). At the end of this phase, a working title for each theme is provided via a clear and concise permanent name (Braun & Clarke, 2006). A final thematic map is presented in Figure 5.5, which contains four main themes.

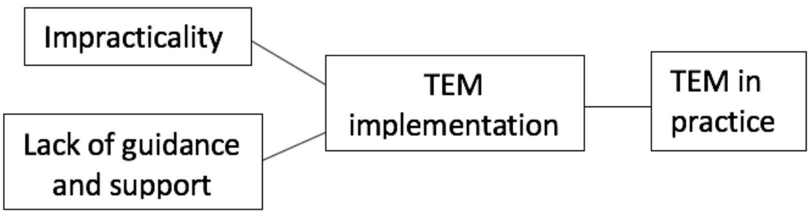

Figure 5.5. Final thematic map for this study with four main themes.

The sixth phase of producing the report involves the final analysis and subsequent writing up of definitive themes, subthemes and their relevant data extracts (Braun & Clarke, 2006). The final write up provides a 'concise, coherent, logical, non-repetitive and interesting account of the study … within and across themes', with appropriately embedded data extracts and analytic narratives that present plausible arguments to address the research questions (Braun & Clarke, 2006, p. 93). In order to further improve the quality of thematic analysis Braun and Clarke (2006, p. 96) offered the following criteria (Table 5.5). These criteria were used as checklist items when analysing and presenting the interview data for this study.

Table 5.5 *Fifteen Criteria for Good Thematic Analysis and its Outcomes.*

Process	No.	Criteria
Transcription	1	The data have been transcribed to an appropriate level of detail, and the transcripts have been checked against the tapes for 'accuracy'.
Coding	2	Each data item has received equal attention in the coding processes.
	3	Themes have not been generated from a few vivid examples (an anecdotal approach); rather, the coding process has been thorough, inclusive and comprehensive.
	4	All relevant extracts for all themes have been collated.
	5	Themes have been checked against each other and the original dataset.
	6	Themes are internally coherent, consistent and distinctive.
Analysis	7	Data have been analysed, interpreted and made sense of, rather than just paraphrased or described.
	8	The analysis and data match each other—the extracts illustrate the analytic claims.
	9	The analysis tells a convincing and well-organised story about the data and topic.
	10	A good balance between analytic narrative and illustrative extracts is provided.
Overall	11	Enough time has been allocated to complete all phases of the analysis adequately, without rushing a phase or giving a phase insufficient time.
Written report	12	The assumptions about, and specific approach to, thematic analysis are clearly explicated.
	13	There is a good fit between what the researchers claim they did, and what the researchers show they have done—i.e., the described method and reported analysis are consistent.
	14	The language and concepts used in the report are consistent with the epistemological position of the analysis.
	15	The researcher is positioned as *active* in the research process; themes do not just 'emerge'.

From the start to the end of thematic analysis, CAQDAS, NVivo, was used to assist the researcher with data analysis and presentation for this part of the study.

5.3.2 Computer-Assisted Qualitative Data Analysis Software (CAQDAS).

Computer-assisted basic content analysis of text became popular from the 1960s (Seale, 2008). Since then, the development of CAQDAS has continued, while improvement and refinement on various strands of qualitative data analysis methods were conducted. By 1990, there were over 26 qualitative analysis software packages available, all aiming 'to facilitate data management and promote the rigour of qualitative research' (Bazeley & Jackson, 2013, p. 6). Although earlier versions of these programs and their functionalities were somewhat limited (e.g., mainly text-retrieval tasks) because of the lower computer capabilities at that time, these limitations were gradually lifted with the advancement of computer technologies (Seale, 2008)

As a result of subsequent updates and versions, newer and more sophisticated features and functions were made available (Bazeley & Jackson, 2013). These improved features offer a broad range of support for the common tasks of qualitative data analysis, including organisation, exploration, interpretation and integration of gathered data (Gilbert, Jackson, & di Gregorio, 2014). It is noteworthy that the use of CAQDAS helps researchers/data analysts with the management, retrieval and analysis of qualitative data. However, it does not replace high-level cognitive exercises, such as defining and developing conceptual themes, writing memos and journals, gaining insights into phenomena and promoting theoretical understanding (St John & Johnson, 2000).

The use of CAQDAS is well established in the qualitative research domain because of the advantages listed below. However, there are some concerns regarding the use of CAQDAS. The following paragraphs describe some of the advantages, followed by the concerns raised. The advantages of using CAQDAS include:

- an ability to deal with large amounts of qualitative data
- a reduction in the amount of time needed for manual handling tasks
- an increase in flexibility and thoroughness in handling data
- the provision of more rigorous analysis of data
- the provision of a more visible audit trail in data analyses (St John & Johnson, 2000, p. 394).

The first two advantages are obvious, as the use of computers conveniently and effectively increases handling and processing speed when sorting and searching for pre-coded and categorised qualitative data. The use of software allows researchers/data analysts to spend more time completing high-level, creative and intellectual tasks, such as thinking about the 'meaning of data', rather than working through a large pile of photocopied papers with a range of different colours of highlighters and sticky notes (Seale, 2008, p. 235). In short, an aerial view of data is readily available to researchers/data analysts.

CAQDAS greatly improves flexibility and thoroughness when handling qualitative data. As the data analysis progresses, researchers/data analysts have greater ability to add notes, in either textual or audio-visual forms, when defining and developing concepts and ideas (St John & Johnson, 2000). This is particularly important when several researchers/data analysts are working on a project from different geographical locations. Codes are developed as the data analysis progresses, and these codes and coding systems can easily be modified, merged, moved, replaced and deleted (St John & Johnson, 2000). In addition, graphical representation of models and ideas will provide greater insights into and understanding of the rich, and often complex, qualitative data (St John & Johnson, 2000). This will subsequently improve the thoroughness of the analysis when handling particularly large complex data.

The use of CAQDAS helps researchers/data analysts demonstrate that their research findings and interpretations are based on rigorous analysis (Seale, 2008). This is achieved because CAQDAS examines entire data that relate to certain analytic questions, and will subsequently overcome certain assumptions and biases that qualitative researchers naturally bring into their research (St John & Johnson, 2000).

Consequently, the overall rigour and trustworthiness of research conclusions is improved by avoiding false or distorted accounts that are based on researchers'/data analysts' assumptions and worldviews (Neuman, 2011).

The final advantage of using CAQDAS in qualitative research is an improved visible audit trail in data analyses (St John & Johnson, 2000). Qualitative researchers are often accused of not being explicit in their analytical strategies when analysing qualitative data (St John & Johnson, 2000). This perceived limitation can be overcome with the use of CAQDAS because the entire dataset (e.g., interview transcription, interview notes, audio recording of interviews and memos) can be made available. More importantly, a clear audit trail is established when researchers vigorously document their enquiry process when analysing qualitative data with the use of CAQDAS 'through journaling and memoing, keeping a research log of all activities … and recoding data analysis procedures clearly' (Creswell & Miller, 2000, p. 128).

Despite the number of advantages described above, there are also concerns raised when using CAQDAS in qualitative research. The majority of concerns are associated with acceptance of the use of CAQDAS, without comprehensive appreciation of its potential shortcomings (St John & Johnson, 2000). These concerns may result in:

- a focus on quantity, rather than meaning
- homogenisation of qualitative data analysis approaches
- pressures or expectations that all qualitative researchers will use them
- a privileging of coding and retrieval methods
- distancing of the researcher from the data
- inappropriate use of technology
- time consumed in learning to use computer packages (St John & Johnson, 2000, p. 395).

A major concern with the use of CAQDAS is the potential for qualitative researchers to feel obligated to obtain large amounts of data, which can result in focusing on breadth, instead of depth (St John & Johnson, 2000). Generally, the larger the qualitative data size, the more time is needed to explore these complex data for the ‘meanings, conceptual understanding, and discourses’ that shape research findings (St John & Johnson, 2000, p. 395). This may lead to unintended consequences, with not all qualitative data receiving the same level of attention when in-depth analyses are performed. One way to overcome this concern is to check data saturation while qualitative data are collected, so that no further collection of data reveals anything new. The current researcher transcribed each interview recording before conducting the next interview. In this way, the researcher gained a good understanding of when data saturation was reached; thus, the researcher was able to make an informed decision about whether further interviews were required.

Another concern raised with the use of CAQDAS is the increased likelihood of homogenised qualitative data analysis approaches (St John & Johnson, 2000). Indeed, there is a definite trend that the majority of qualitative research uses such computer programs for data analyses, and the unintended consequences include findings ‘becoming homogeneous, deterministic and limited’ (St John & Johnson, 2000, p. 396). In general, CAQDAS is better suited to the research area of grounded theory, where coding and retrieval are conventional; however, this may not be the case for other domains of qualitative research, such as discourse analysis, when in-depth analysis of short data extracts and their meanings is conducted (Seale, 2008). Neuman (2011) asserted that data collection in qualitative research is ‘an interactive process in which particular researchers operate in an evolving setting’; thus, the findings from a similar research project by different researchers, or researchers using alternative measures, will vary from one another. Therefore, this concern relating to homogenised qualitative data analysis approaches may be minimised.

The majority of CAQDAS is based on the assumption that qualitative data analysis mainly involves coding and retrieval, as many of these software packages are developed by researchers whose research background is in grounded theory, which uses coding and retrieval methods to engage in theory building (St John & Johnson, 2000). However, other branches of qualitative research (e.g., phenomenology) do not necessarily focus on theory building. Therefore, excessive coding—which refers to the disintegration of data into many unnecessarily smaller pieces—may result in data being decontextualised, so that the essence of meanings may no longer be visible (St John & Johnson, 2000). This was carefully considered when the researcher analysed and coded interview data using thematic analysis to ensure that each code contained its relevant context and meaning.

Similarly, reduction of data through coding may separate researchers from the rich qualitative data, which may lead to 'loss of meaning and context and creating sterile and dehumanised data' (St John & Johnson, 2000, p. 396). This suboptimal outcome may be exacerbated if qualitative researchers/data analysts—particularly in a large project with multiple researchers/data analysts involved—merely focus on analysing pre-coded words and phrases, without underlying unspoken, yet critical, cues, such as facial expressions. In fact, St John and Johnson (2000) cautioned that 'using technology in qualitative research may strip away the meanings with which inflection and body language imbue words' (p. 396). Therefore, technology must be used appropriately and sensibly, with some parts of the recording and/or notes taken attached to certain pre-coded texts to minimise any possible uncertainties, while preserving the richness of the qualitative data as much as possible. As previously mentioned, the current researcher maintained a close distance to the interview data from the beginning of the study to the completion of writing the thesis so that rich information within the interview data was not lost.

Another concern raised with the use of CAQDAS is the large investment of time and effort required to initially learn and keep up with updates for the particular software that researchers decide to use. However, this is not limited to qualitative research. Quantitative researchers also need to invest such time and effort to learn statistical programs, such as SPSS. This investment in time and effort to learn to use software can be a challenge, especially for early career researchers who need to learn particular software, in addition to the learning required to understand the procedures of research. Despite this, use of computer programs is advisable and beneficial, particularly, but not necessarily, when dealing with large databases (Creswell, 2013). It is noteworthy that there is no pressure and expectation that all qualitative researchers/data analysts need to use CAQDAS. The researcher considers himself an early career researcher who has a preference for mixed methods. Therefore, the investment in time and effort to learn CAQDAS, NVivo, was considered worthwhile because this is the platform to be used when analysing qualitative data for future research.

There are alternative options to CAQDAS, such as the Microsoft Word package. Seale (2008) explained that many features in the Microsoft Word package can perform similar functions to CAQDAS. For instance, a function called a 'spike' allows users to copy multiple words and/or paragraphs while examining a document, and paste all 'spiked' words/paragraphs in a separate document, instead of copying and pasting each word/paragraph. However, in the long term, it is worth the effort to learn and become familiar with preferred computer software, particularly for those individuals who plan to become career researchers, as there are many more powerful functions—such as visualisation of conceptual maps—that can be easily created with the use of CAQDAS.

After carefully considering the advantages and concerns regarding the use of CAQDAS, it was decided to use CAQDAS for the qualitative data analysis for this part of the study because the benefits outweighed disadvantages. The question remaining was which particular CAQDAS would be used for this study. There were various qualitative analysis software packages available and after a period of competing, advancing and consolidating many different types of software programs, three major CAQDAS packages came to dominate: ATLAS.ti, MAXQDA and NVivo (Davidson & di Gregorio, 2011). There are some similarities and differences in terms of the functionalities, strengths and weaknesses each software package presents. Evers, Sliver, Mruck, and Peeters (2011) conducted an investigation comparing these three major software packages, and produced similar conclusions when the same dataset was independently analysed. Additionally, the interpretative process of qualitative data analysis remains with researchers/data analysts, regardless of which CAQDAS package is used. Therefore, the current study used NVivo (Version 11 for Apple Macintosh computers), as this was made available, free of charge, to the researcher by the university.

5.3.3 Ethics.

Ethics is defined as 'moral principles that govern a person's behaviour or the conducting of an activity' (Oxford University Press, 2017) and the ethical conduct of research is a central part of the research process. Morris (2015, p. 19) provided six fundamental considerations when conducting research:

1. Research should be designed, reviewed and undertaken to ensure integrity, quality and transparency.
2. Research staff and participants must usually be fully informed of the purpose, methods and intended possible uses of the research; what their participation in the research entails; and what risks, if any, are involved. Some variation is allowed in very specific research contexts.
3. The confidentiality of information supplied by research participants and the anonymity of respondents must be respected.
4. Research participants must take part voluntarily, free from any coercion.
5. Harm to research participants and researchers must be avoided in all instances.
6. The independence of research must be clear, and any conflicts of interest or partiality must be explicit.

The above ethical principles were considered prior to and during the process of obtaining human ethics clearance from the university Human Research Ethics Committee, which was established and conducted in compliance with the current National Health and Medical Research Council National Statement on Ethical Conduct in Human Research. The required ethical clearance was obtained on 10 October 2016 (Appendix 8).

5.4 Results and Interim Discussion

This section describes the main themes under TEM implementation that arose from thematic analysis, followed by separately discussing each of the main themes with supporting quotations from the transcripts. The main themes identified were as follows: impracticality, lack of guidance and support, TEM implementation and TEM in practice. Although there were four separate themes, these themes were somewhat interwoven. For instance, one of the six key considerations for successful implementation of a policy initiative is ensuring the appropriate mix of required expertise and quantity of resources available (ANAO, 2016). However, because of insufficient resources from CASA, there was a lack of support and guidance provided to Australian general aviation pilots. Before discussing the themes, it is appropriate to provide a brief historical background of CASA, given that CASA is responsible for the successful implementation of TEM in Australian general aviation.

5.4.1 Brief historical background of the Civil Aviation Safety Authority (CASA).

CASA is a regulatory body for Australian civil aviation. The birth of CASA can be traced back to May 1921, when a civil aviation branch was established within the Defence Department to administer legislation through the *Air Navigation Act 1920* (Bartsch, 2013). The civil aviation branch evolved into the Department of Civil Aviation in 1938, when the regulation of civil aviation and military activities was separated (Bartsch, 2013). The department then transformed into the Civil Aviation Authority in 1988, and the authority was expected to handle both 'commercial viability' and the 'enhancement of an air safety regulatory regime' (Bartsch, 2013, p. 100). However, these two competing expectations and a consequent series of unfortunate accidents led to the disestablishment of the authority, transferring the commercial and operational functions to Airservices Australia, and the establishment of CASA in 1995 as an independent statutory authority under the *Civil Aviation Act 1988* to constitutionally focus on the air safety regulatory regime (Bartsch, 2013; Colmar Brunton, 2016). Therefore, the primary function of CASA (2018) is to provide safety regulations to Australia's civil air operations, nationally and internationally, as well as providing comprehensive safety education and training programs.

5.4.2 Impracticality.

Impracticality is the first theme that arose from this study's thematic analysis. The Oxford Online Dictionary (2018) defined practicality as 'an aspect of a situation that involves the actual doing or experience of something rather than theories or ideas' (para. 2). General aviation is primarily a vocationally based industry; thus, it is important to consider practical aspects when new safety initiatives are proposed and introduced in general aviation, such as TEM implementation. However, all this study's participants expressed concerns regarding a lack of practicality when it came to understanding the concept of TEM and implementing it in practice, whether in daily routine flying or flight training. Participant 3 stated that *'CASA's tried hard ... they* *did it very consciously, very formally and completely impractically'*. Participant 4 provided a similar observation: *'Did we actually need to get someone to put together a paper 90% of aviation community could not understand because they don't have degrees? … We probably should've put it on a practical level'*. Participant 5 shared a similar observation regarding CASA's theory examination on TEM:

> *'I got a sample … Is this a handling error or procedural error or operational decision error or intentional non-compliance error? So, to me, I thought, well, I'm not really worried if students couldn't put a tag on what sort of error it is. I'm more concerned with, did he spot the error? Uh, I would prefer to see … questions on more scenario focused, you know. Put up a scenario. Is there a threat or error situation likely to be developed here? What would be the countermeasure or what could be the possible undesired aircraft state? That's how I like to see it focused, not is this a thing?' (Participant 5)*

This was consistent with a finding from a stakeholder survey (Comar Brunton, 2016). The online survey involved a total of 1,217 participant stakeholders in Australian aviation, and indicated that approximately 58% of the stakeholders expressed disagreement when asked whether CASA did a good job of translating participants' legal obligations into practical guidance (Colmar Brunton, 2016).

Possessing a thorough understanding of theory is important; however, because general aviation is a more vocationally oriented industry, it is also important to place an equal, if not greater, emphasis on practical aspects. Clearly, these two approaches to theoretical knowledge and practical skills are complementary to each other, yet create difficulties when they need to be integrated. Based on the comments from the participants, it could be well established that a lack of practicality was evident when TEM was mandated in Australian general aviation, and the probable reasons for this may be twofold. The first reason is likely a result of senior management and decision makers within CASA who may be inadequately familiar with the general aviation industry. Participant 2 stated, *'You've got all these managers who do not understand their core business'*. Participant 1 supported this view:

'CASA particularly … but if you look at the profile of a lot of senior management within CASA, ones we got last probably 10 to 15 years has very little expertise in aviation there. So, they are driving the show and they are not too much focusing into the operational aspect, uh, what's going on in the operational field, be it in airlines or be it in the training industry? They are not going to put their focus there because, by large, they don't understand it.' (Participant 1)

Participant 1 further added that it is important to have a good understanding of the industry, so that effective communication methods can be established:

'I mean they did the usual thing of producing some CAAP information, Civil Aviation Advisory Publications or something, but that was just a document, which is then reliant on people getting hold of that document, reading it, understanding it, taking it apart, then setting up the implementation plan, and that really doesn't fit the style of industry.' (Participant1)

Again, this was consistent with the findings from the stakeholder survey (Colman Brunton, 2016). The results indicated that almost 50% of the stakeholders were dissatisfied (28% very dissatisfied and 20% dissatisfied) with CASA's ongoing dialogue with the industry. Again, almost 50% did not agree that CASA valued input from the industry (Colmar Brunton, 2016).

The second probable reason for the lack of practicality when TEM was mandated seemed to occur more at the operational level. Participant 4 felt that CASA did not have the right people in teams responsible when the implementation of TEM was planned:

'I believe at that time, um, CASA had engaged in a consultant, um, who I don't think was from an aviation background at that time, uh. A lot of what came out made no sense whatsoever, um, and it proved, at that time, quite difficult to find anything aviation specific with threat and error

management because it had only been introduced, um, for its integration into a training program.' (Participant 4)

This was similar when CRM was first introduced in the US in 1981. The criticisms of the first generation of CRM stated that it was largely psychological in nature, and that the associated games and exercises to illustrate concepts were unrelated to aviation, which negatively affected overall acceptance among pilots (Helmreich et al., 1999). Similarly, yet outside the aviation context, Zhang (2009) identified that the majority of teachers in vocational-technical schools come from other disciplines, with little practical experience and a consequent lack of disciplinary understanding, together with inadequate work experience, which contributed to a poorer learning experience. These highlighted the importance of having adequate industry understanding so that successful implementation of new initiatives is achieved as intended.

The stakeholder survey also confirmed the current study's findings. When asked to rate participants' level of satisfaction with the ways that the CASA developed aviation safety regulations, over half of the participants (38% very dissatisfied and 20% dissatisfied) indicated dissatisfaction with the way the CASA developed aviation safety regulations (Colmar Brunton, 2016).

Participant 4 further added that his experience of other CASA initiatives was that they were not aimed at the right level for the industry:

> *'I think it's not just threat and error management, but, in a lot of ways, right down to, um, 142[3] and you know, expositions and, um, training manuals, I think we are overcomplicating the delivery of training. I personally believe we'll actually end up with worse outcomes. If you keep something simple, you have a much higher level of compliance than what you do when you are overcomplicating things.' (Participant 4)*

[3] This refers to Part 142 integrated flight training in aeroplanes.

Participant 4 continued that *'the more academic approach we have today, um, I don't believe this is more effective'*, and that the so-called academic approach and over-complication has resulted in the following: *'That's what, how it got happened to TEM. It was overcomplicated to the point that some people just refuse to even look at it'* (Participant 4). This was consistent with a finding from the stakeholder survey, where only 28% of the stakeholders considered it 'easy' or 'very easy' to comply with their aviation safety regulations (Colmar Brunton, 2016). Outside aviation, Zhang (2009) observed a similar trend in vocational-technical schools in China, where theory was over-emphasised, while practical training was placed in a secondary position. This trend underestimated the added advantages of further enhancing students' knowledge and appreciation (Zhang, 2009).

It can be well established from the participants' accounts that, when TEM was first implemented, it was not pitched at a level that was appropriate to the general aviation industry. Consequently, it was viewed by many stakeholders as an additional requirement with which they needed to comply, without them appreciating the value of TEM. In addition, the findings from the stakeholder survey were similar to the participants' recounts. Thus, the lack of practicality was the first barrier identified as hindering the successful implementation of TEM in Australian general aviation.

5.4.3 Lack of guidance and support.

All participants were more in agreement that there was inadequate guidance and support provided by CASA when TEM was implemented in Australian general aviation. It was established in Study 1 that the participants' responses were neutral (M = 3.44, SD = 1.25) when asked to rate whether the CASA produced adequate training and guidance materials for TEM training for general aviation.

The Guild of Air Pilots and Air Navigators (GAPAN), now known as the Honourable Company of Air Pilots, in collaboration with Australian Transport Safety Bureau (ATSB), developed a 'TEM train-the-trainer' course for general aviation and low-capacity air transport operations, and conducted the course in 10 different locations around Australia between August and October 2007 in preparation for the regulatory changes (Cheng et al, 2009). Participant 4 recalled:

> *'The GAPAN material is actually quite good, arguably, um, miles ahead of what was actually produced by the regulator, but how many people actually attended that GAPAN roadshow? I guess that was the question—what were their positions, were they actually syllabus designers or CFIs, um, or the ground school instructors or the flight instructors? There were, they did have good attendance, but I would probably say less than 20% of the industry.' (Participant 4)*

According to Cheng et al. (2009), approximately 312 people attended the training. This was a significant under-representation, considering there were over 20,000 licence holders (e.g., ATPL = 6,825, CPL = 4,189 and PPL = 10,563) in the aeroplane category alone during the 2009 to 2010 periods (CASA, 2010). In addition, the course was only conducted in major cities; thus, organisations in rural and remote areas would have encountered difficulty attending the courses.

In addition to the GAPAN roadshow, CASA has been running a range of seminars and workshops, such as AvSafety seminars, to assist general aviation pilots to stay up to date and to promote aviation safety. Given that TEM was a new and important safety initiative, there should have been extensive coverage to promote TEM. Participant 4 recalled that, *'They have touched on TEM in seminars, yeah, um. Not recently, not for a very long time, um, but they have touched on them'*. Participant 3 also shared his observation:

> *'so, was there any special training on it? No. CASA had run courses for private pilots where they could talk about those things, but I think I only attended one and mostly they were dealing with questions and they don't—there was no overall philosophy, except they've got to cover and demonstrate that you're considering threats and how you manage them.' (Participant 3)*

Further, Participant 1 added:

> *'Industry needed, certainly from the outset, an education roadshow to go around the regional centres, set up workshops and explain what it was,*

what they wanted, what outcomes they wanted out of TEM implementation … It left the industry pretty much on their own devices.' (Participant 1)

Similar findings were noted in the stakeholder survey. The survey used a 10-point Likert scale, where 0 represented 'very dissatisfied' and 10 represented 'very satisfied'. When stakeholders rated whether the CASA had clearly and succinctly explained the regulations and how they would affect industry stakeholders, the mean agreement was merely 2.9 out of 10 (Colmar Brunton, 2016). These results suggest that more thoughtful planning and provision of ongoing guidance are prerequisites for the successful implementation of any new initiative.

All the participants had been actively involved in flight testing before and after TEM was implemented, and one of the requirements to maintain Flight Examiner (FE) rating[4] was to attend a biennial professional development program (PDP). Participant 1 described the PDP as follows:

'Those PDPs primarily focused on taking us through, during a two-day program, the elements of a flight test or a series of flight tests, looking at components and elements, making sure we met regulatory requirements, making sure we covered all aspects, discuss any grey areas about what was in the flight test requirements. That was primarily the focus.' (Participant 1)

Participant 4 shared his experience with scepticism when he attended PDPs:

'What we tended to get in, um, in PDPs was more discussion points. And it was open to the room and generally you would find that every PDP I remember up till, um, probably 2012 was just, um, examiners were left to deliberate and debate, um, and you would generally come away no wiser. I mean, it really depends on who was a more, greater influential speaker in the room as to which view people end up taking, but because examiners

[4] A Flight Examiner (FE) is defined as a person who holds a FE rating with associated FE endorsements, or a person approved to conduct flight tests and/or proficiency checks (CASA, 2017).

in their egos naturally serves as pilots, um, has everybody generally thinking they know the best, the way they've done this, different or right way, you would usually arrive at these events, um, with your own opinions. You may have your eyes opened to opinions of others, but I think the majority of people still left with their own opinions.' (Participant 4)

The introduction of TEM would have affected FEs in terms of the way they examined flight test candidates and new TEM-related flight test items. However, despite its rightly deserved attention regarding how TEM items should be accurately and adequately examined, Participant 1's recount was somewhat different:

'…it [PDP] didn't focus on TEM. It didn't mention it and there was no separate PDP program to cover it to say 'this is a new thing we, CASA, have introduced to the industry … we really want to focus on it and we want you to focus on it and therefore we will give you separate training'. No, they didn't do that.' (Participant 1)

A lack of guidance and education opportunities on TEM caused Participant 5 to have difficulty in terms of examining and completing TEM elements on a flight testing form:

'Especially on the test form for the flight tests. It's just got three very basic items, uh, but, uh, like, when you are doing a flight test, they just say, 'recognise and manage threats', 'recognise and manage errors', 'recognise and manage undesired aircraft state'. Now, that is probably too broad and too general, I think, to be even meaningful, unless you really had a thorough knowledge of what threat and error management is.' (Participant 5)

Participant 5 further added that:

'My main concern right from the outset is, as an examiner, um, checking that the students are managing threats and errors and getting back to, you know, having a test sheet that has a bit more, more prompts on it, rather than, you know, you can probably go back and say, well, refer to the

Manual of Standards[5]*, but the Manual of Standards, while it elaborates a little bit more, it doesn't elaborate a lot and, uh, I unfortunately don't know what a quick fix would be to trying to fix that problem.' (Participant 5)*

A similar observation was made by Participant 4 when assessing TEM during a flight test:

> *'There was no 'this is threat and error management. This is how to actually build it into flight tests and this is what you are looking for', remembering the Day VFR [Visual Flight Rules] Syllabus didn't have a section specifically for testing, nor it did actually provide any real clear guidance on performance criteria, what it was we were actually looking to see.' (Participant 4)*

Participant 4 further added that:

> *'We were provided with no training ourselves, and so myself and the CFI at that time had never been trained threat and error management ourselves, and we were trying to read a highly technical, um, advice, materials from CASA, um, which, I guess, left all of us to probably feel dazed and confused.' (Participant 4)*

The comments above shed some light on the finding from Study 1 that the trainer group did not strongly feel that TEM was appropriately assessed during flight tests (M = 3.67, SD = 1.2).

Based on the participants' recounts, there was a lack of adequate support and definitive guidance when TEM was first introduced. Three participants offered possible reasons regarding why there was a lack of guidance provided by CASA. Participant 4 stated: *'It was a lack of understanding due to a lack of education, lack of education because lack of understanding of threat and error management within the regulator*

[5] The Part 61 Manual of Standards (MOS) is a legislative document that provides information relating to flight crew licensing regulations (CASA, 2018f)

themselves'. Participant 4's recount about the lack of understanding of TEM among the regulators themselves was supported by Participant 3:

> *'When they were introducing it, they were sort of saying things like 'we're only still trying to find out how to do it'. Even later, I spoke to a senior manager at CASA, and he said, 'we have no idea what it's about', so we were enforcing something that we didn't understand.' (Participant 3)*

Consequently, it was regulated that TEM was required to be assessed during flight tests, yet there was no adequate guidance provided to FEs due to a lack of understanding of TEM within the CASA. Participant 5 recalled that:

> *'It may well be that CASA themselves might have been struggling, you know, 'how the hell are we going to assess it?', so I'll just put the headings in there and leave it up to the examiner, let him work it out himself.' (Participant 5)*

Participant 4 supported Participant 5's observation: *'To this day, I would be fairly confident in saying that you could ask 15 inspectors at CASA and you'll have 15 different opinions on threat and error management and their understanding'*. The lack of support and guidance from CASA affected how successfully TEM was implemented in flight schools. Participant 1 described this as follows:

> *'Lower end of the scale to probably middle part of the scale probably didn't implement it very well at all, mainly because of lack of understanding and, uh, the big end of town, who have got more resources and probably more expertise and so they probably did it reasonably well.' (Participant 1)*

Participant 1 continued that:

> *'Places like Flight Training Adelaide, CAE Oxford get feedback from airline clients and/or their parent company. CAE Oxford down at Moorabbin—they used to be one of the QANTAS providers, previous*

named General Flying Services … I used to go down there. They would've drawn a lot of their expertise on TEM down here from their parent company in UK and Europe because EASA [European Aviation Safety Agency] brought TEM in a number of years before we did, so, you are right. They had resources and they have contacts in the big end of town to bring the stuff in and probably did it quite well. The lower end, they don't.' (Participant 1)

However, Participant 3 argued that large flight schools do not always do a good job, particularly in terms of mentoring junior Flight Instructors (FI). These junior FIs at large flight schools conduct the majority of flight training:

'The big schools, and, is in fact less likely to give that expertise because the only ones are—there are much fewer senior instructors who, because it's a big school, you only get to see the others from time to time. The other thing is those big schools are so busy training that they don't tend—they don't mentor their junior instructors. They all tell you they do, but having observed them, um, they don't mentor the junior instructors. All they do is if, if their instructors are continually having trouble producing students at standards, he will get criticised. He won't get an assistance to do this. They'll say 'you're not doing a good job'. They may lose their job and they might get another cheap instructor. So, I think they are less likely to, to give extra, they will have more resources, but spread thin, uh, but resources are being run by a much higher percentage of less experienced instructors. They are barely out of commercial licence.' (Participant 3)

Based on awareness of the different levels of TEM implementation and lack of mentoring opportunities for junior FIs, Participant 1 highlighted the importance of producing a FI manual:

'You can embed TEM in a comprehensive instructor manual covering teaching basic exercises, but also, as you work through the stages of training how to introduce TEM, the concepts of it, practical scenarios etcetera, and that would be the basis on which to say, 'we train our

> *instructors, we standardise our instructors'. Things would flow from there ... If you embed TEM in the instructor manual, it's got all the details in there then you go—okay, we train to this, it is expected to be done like this at the flight schools.' (Participant 1)*

Participants made a few suggestions that the lack of support and guidance from CASA could have been due to resource constraints and disproportionate budget allocations within the regulator, which would have affected adequate allocation of staff and their required initial and ongoing training. Participant 1 shared his experience in this regard:

> *'The industry grew too much. In fact, with a budget cut, uh, CASA expanded in all sorts of ways, but not with on-the-ground people. They've got fewer FOIs [Flying Operations Inspectors], much more management and their administrative staff, so they didn't give resources to do it.' (Participant 1)*

Participant 2 added that: *'So someone upstairs and the standard's been written all of these beautiful stuffs, right? But they really haven't trained their staff. So, their staff can't, don't understand it, can't deliver it ... it's across the board'.* All participants shared very similar views that there was inadequate guidance provided by CASA when TEM was implemented in Australian general aviation. The findings from Study 1 also supported this view. Although the participants were very positive about GAPAN's roadshow, the number of pilots who actually attended the roadshow was very low compared with the total number of licenced pilots in Australia at that time. In addition, the participants identified that there was no specific training provided to FEs regarding how TEM was required to be assessed. Further, based on participants' observations, there were very limited educational opportunities for general aviation pilots to learn TEM principles and their proper application. This lack of adequate guidance and support from CASA was seemingly due to financial constraints and the consequent absence of adequately trained staff. Therefore, it was then a logical step to examine how TEM was implemented, which is discussed in the next section.

5.4.4 Threat and error management implementation.

Oxford (2018) defined implementation as 'the process of putting a decision or plan into effect' (para. 2), and any implementation certainly requires a thoughtful process. The general consensus among all participants was that the implementation of TEM was problematic. Participant 1 recalled that:

> *'It was mandated by CASA to be inserted into the VFR syllabus and flight schools were then required to teach TEM as it was known, but implementation was the problem … Nobody really understood what it was CASA wanted and nobody really clearly defined it and it was not very well implemented.' (Participant 1)*

Participant 2 added that: *'I think that they've introduced threat and error management which's a fantastic idea, but I don't think it's been … it's been poorly introduced and therefore there's no value, you know'*. The Australian National Audit Office (ANAO) (2014) provided a guide that identified six key considerations 'when implementing a policy initiative—the act of translating policy into reality—so that intended benefits are realised' (p. 3). These considerations were governance, managing risk, engaging stakeholders, planning, resources and monitoring, and review and evaluation (ANAO, 2014).

The first key consideration, governance, refers to the arrangements and practices that enable the achievement of expected outcomes (ANAO, 2014). Sound governance arrangements require a committed executive, supported by a senior responsible officer who is accountable for the success of policy implementation (ANAO, 2014). The senior responsible officer is required to either possess the appropriate skills or have access to such skills to oversee the implementation process, as well as organising a suitably skilled implementation team to assist him or her (ANAO, 2014). However, based on the participants' recounts, as discussed below, it was evident that a sound governance arrangement was not well established:

'CASA particularly ... if you look at the profile of a lot of senior management within CASA, ones we got last probably 10 to 15 years has very little expertise in aviation there. So, they are driving the show and they are not ... too much focusing into the operational aspect, uh, what's going on in the operational field, be it in airlines or be it in the training industry. They are not going to put their focus there because, by large, they don't understand it.' (Participant 1)

'When they were introducing it, they were sort of saying things like 'we're only still trying to find out how to do it'. Even later, I spoke to a senior manager at CASA, and he said, 'we have no idea what it's about', so we were enforcing something that we didn't understand.' (Participant 3)

'I believe at that time, um, CASA had engaged in a consultant, um, who I don't think was from an aviation background at that time, uh. A lot of what came out made no sense whatsoever, um, and it proved, at that time, quite difficult to find anything aviation specific with threat and error management because it had only been introduced, um, for its integration into a training program ... I'll be brutally honest, during those PDPs and in that era, was, I think a lot of ... there was no technical specialist actually employed by CASA and threat and error management.' (Participant 4)

The findings from a previous stakeholder survey suggested a similar view, in which the participants strongly believed that CASA was disconnected from the aviation industry, which resulted in a perceived lack of understanding of the operational effects that regulatory changes could bring (Colman Brunton, 2016). In addition, the survey indicated that only 26% of respondents were very satisfied (9%) or satisfied (17%) when asked whether CASA had provided competent and capable staff (Colman Brunton, 2016).

The second key consideration for successful implementation is risk management (ANAO, 2014). Risk in this context refers to dispersion around the expected outcomes. CASA (2008) introduced TEM in response to the endorsement and recommendation to include TEM training in all pilot training by ICAO, with a view to further improve flight safety in general aviation. As such, the risk management aspect of successful TEM implementation was to ensure that TEM was properly introduced, taught, practised and assessed to achieve the intended outcome of further improving flight safety in general aviation. However, as established in the previous sections and in the participants' accounts below, poor teaching and inconsistent assessment of TEM among FEs were evident. Therefore, successful implementation of TEM was hindered because of inadequate risk management by CASA:

> *'It was mandated by CASA to be inserted into the VFR syllabus and flight schools were then required to teach TEM as it was known but ... Nobody really understood what it was CASA wanted and nobody really clearly defined it ... There was not an education roadshow set up by the regulator, which should have happened to go around and do workshops with flight schools.' (Participant 1)*

> *'There was no 'this is threat and error management. This is how to actually build it into flight tests and this is what you are looking for', remembering the Day VFR Syllabus didn't have a section specifically for testing, nor it did actually provide any real clear guidance on performance criteria, what it was we were actually looking to see.' (Participant 4)*

Similar observations were made in the stakeholder survey, where only 14% of respondents strongly agreed (4%) or agreed (10%) when asked whether CASA clearly and succinctly explained the regulations and how they would affect industry stakeholders (Colman Brunton, 2016).

The third key consideration for successful implementation is engaging stakeholders (ANAO, 2014). Figure 5.6 below depicts the importance of establishing a high level of engagement with relevant stakeholders. The ANAO (2014) asserted that effective stakeholder engagement requires a clear objective for consultation and the identification of key stakeholders, while maintaining clear and timely communication to wider communities.

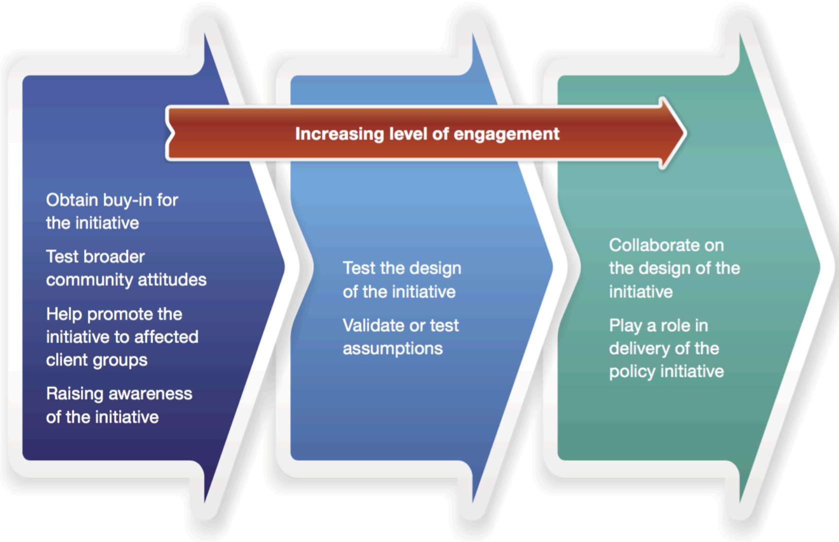

Figure 5.6. Reasons for engaging with stakeholders. Source: adapted from ANAO (2014, p. 38).

However, the observations by participants were divergent regarding the establishment of effective stakeholder engagement. Participant 1 recalled:

> *'As I already said, TEM was a buzzword. It was pushed out there and it was required by CASA to be introduced to flight schools, and nobody really had a clear understanding of what it was, its components or elements.' (Participant 1)*

Participant 1 further added that:

> *'Industry needed, certainly from the outset, an education roadshow to go around the regional centres, set up workshops and explain what it was, what they wanted, what outcomes they wanted out of TEM implementation … It left the industry pretty much on their own devices.' (Participant 1)*

Participant 4 shared a similar view that, before TEM was formally introduced, a practical and thoughtful approach was required, offering a series of:

> *'Practically focused training session first and then providing the guidance materials on how to actually roll it out into a school environment. If you*

did it that way, you probably would've had a greater pick up and a greater success.' (Participant 4)

Participant 4 also highlighted a lack of close engagement between the CASA and the industry:

> *'If they actually came in and lived, like walked into flight training organisations and actually saw what happens in a day-to-day basis and saw how flying schools operated, how long briefs and short briefs are conducted and how flight lessons were conducted today, what was actually going on, that probably would've been better, uh, prepared or better armed, uh, for the development of training materials for the industry and to have a better roll-out plan.' (Participant 4)*

This suboptimal level of effective stakeholder engagement was also identified in the stakeholder survey. A finding from the survey indicated that almost half of the survey participants were very dissatisfied (28%) or dissatisfied (20%) with the CASA's ongoing dialogue with the industry (Colmar Brunton, 2016). Another finding from the survey suggested that only 11% of participants strongly agreed (4%) or agreed (7%) when asked whether CASA always consulted with the most appropriate people in the industry when developing and reforming aviation safety regulations (Colmar Brunton, 2016). Almost half of the survey participants either strongly disagreed (31%) or disagreed (18%) that CASA would value input from industry. The participants indicated their disconnect from participating in consultation, and the survey indicated that the top two reasons for this disconnect were as follows: 'was not aware I could' (36%) and 'won't make a difference/waste of time' (28%) (Colmar Brunton, 2016). Participant 4 concurred that: *'They might call for, uh, you know, they might put out, um, a Notice of Proposed Rule Making, [and] people may respond, but their response rate is about 1% in the community'.*

Participant 1 observed varying degrees in the ways TEM was implemented in different flight training organisations, citing a combination of low attendance at the GAPAN roadshow, lack of a separate roadshow for flight schools, unclear conduct of the

AvSafety seminar, and lack of clear and timely communications and support to the wider general aviation community:

> *'So, I guess implementation in the flight schools sort of varied from small ones—small country ones, who probably didn't have and still haven't got a clue about TEM, probably didn't even believe in it and therefore really did not implement it at all—to the major flying schools, like Flight Training Adelaide—an organisation that, yeah, would've embraced TEM. They would've had a clear understanding about what it was and, uh, would've implemented it into their procedures. So, clearly it depends on which part of the industry you are looking at.' (Participant 1)*

It was interesting to note that, despite the varying degrees in the ways TEM was implemented, Participant 4 recalled only very small differences when assessing candidates from different flight schools:

> *'From an organisation that understands academics of it and does the ground work, um, put some show on the ground, um, work through threat and error management. There're massive differences what you see there, but there's very little difference between what I have experienced in flight with applicants from both ends of spectrum. I think everybody has always done a fairly good job of in-flight training students to identify threats and errors.' (Participant 4)*

The fourth key consideration for successful implementation is planning, which ANAO (2014) described as providing 'a map of how an initiative will be implemented, addressing matters such as timeframe, dependencies with other policies or activities, program logic, phases of implementation, roles and responsibilities, resourcing, and compliance with legal and policy requirements' (p. 43). A number of participant comments above suggested that more thorough planning would have led to more successful implementation of TEM. A comment from Participant 4 also highlighted the importance of in-depth planning before TEM was implemented:

'I think, the first question should've been asked is: What is this? Is this actually something new or is this something we're already doing and we've just really created, um, a new term for it because all of a sudden, we've got academics to do, uh, come up with something new? Um, have we just created something for the sake of creating something? I think it already exists. If it did already exist, what was actually wrong with the way it's been, what was it actually problem here? Did we actually need to get someone to put together a paper 90% of aviation community could not understand because they don't have degrees, um, or interest, perhaps? Um, we probably should've put it on a practical level, so the first thing we should've done would've been to actually run a series of training sessions, practically focused training session, showing threat and error management in practice, um, teaching CFIs syllabus design, how and examiners, um, how—what is actually a threat, what is actually an error, what are plausible management strategies, um, help them to actually understand the system first of all, and then providing the guidance materials on how to actually roll it out into a school environment. If you did that way, you probably would've had a greater pick up and a greater success.' (Participant 4)

Participant 1 shared his recent experience with CASA that highlights poor planning when Part 61[6] was introduced:

'What happened was they introduced Part 61, and the PDP programs had not been developed. As one of three guinea pigs CASA used to train their licencing staff, I move across from ATO [Approved Testing Officer] to flight examiner in September '14. It was to see what the transition process was like. CASA needed a couple of guinea pigs, so from September '14, my proficiency check—i.e., my renewal as a flight examiner qualification, if you like—was due up in two years—i.e., was due at the end of August, thirty-first of August that year. When I applied for that proficiency check

[6] Part 61 refers to a section in the *Civil Aviation Safety Regulations 1998* that prescribes the requirements and standards for the issue of flight crew licences (CASA, 2017).

> *to be done back in June, just before I was going on leave, I didn't get an answer. I still didn't get an answer when I returned from leave because, one, Part 61 said that you had to do a PDP before you could do your proficiency check. They didn't have a PDP … CASA hasn't developed one and when I asked about my proficiency check, they said, 'well, we don't know what we are going to do with proficiency checks because we haven't worked it out yet'. So, on the thirty-first August, the day my examiner rating technically ran out, there was an exemption that was mailed to me that said examiner proficiency checks were delayed for 18 months, until CASA sort the whole thing out because we don't know what we are going to do yet.' (Participant 1)*

These two comments, as well as other previous comments, highlighted the importance of thorough planning well before any initiative was going to be implemented.

The fifth key consideration for successful implementation is ensuring an appropriate mix of required expertise and quantity of resources available (ANAO, 2014). As aforementioned, assessment of the required skills and resources is carefully considered during the planning stage; however, ANAO (2014) equally identified that the unavailability of personnel with required expertise to implement a planned initiative is one of the most common implementation problems. This was observed by Participant 4, who did not believe the CASA had an appropriate consultant when TEM was implemented:

> *'I believe at that time, um, CASA had engaged in a consultant, um, who I don't think was from an aviation background at that time, uh. A lot of what came out made no sense whatsoever, um, and it proved, at that time, quite difficult to find anything aviation specific with threat and error management because it had only been introduced, um, for its integration into a training program.' (Participant 4)*

The sixth key consideration for successful implementation is the establishment of ongoing active management that comprises well-coordinated monitoring, review and evaluation processes (ANAO, 2014). Ongoing active management aims to appropriately inform and facilitate other relevant key building blocks for decision-making and to support activities for successful implementation of an initiative (ANAO, 2014). According to Participant 3, this active management did not seem to have been well established within CASA:

> *'When they were introducing it, they were sort of saying things like 'we're only still trying to find out how to do it'. Even later, I spoke to a senior manager at CASA, and he said, 'we have no idea what it's about', so we were enforcing something that we didn't understand.' (Participant 3)*

Participant 1 offered one of the ways to establish ongoing active management.

> *'Well, when we ... the first part of our conversation ... implementation, they had, you know, safety education section in there for since forever because one of the mandated CASA functions ... is educating. So, there should've been an opening roadshow to roll out TEM to those regional centres to for them to implement first of all. Then they can reinforce that every 12 to 18 months or whatever. If you embed TEM in the instructor manual, it's got all the details in there, then you go—okay, we train to this, it is expected to be done like this at the flight schools. Their annual or 18-month roadshow would go around and saying, 'we've done some appropriate sampling around the industry and the feedback is that there were still some weak areas—this, this, this and this ... this is how we recommend to fix them'. It is the normal close loop.' (Participant 1)*

This review of the six key considerations for successful implementation of an initiative, alongside the relevant recounts from the participants, suggests that TEM implementation in Australian general aviation was perceived by many as laissez-faire. This could also be identified from a finding in the stakeholder survey that Colmar Brunton (2016) conducted. When asked how satisfied participants felt with the CASA's development of aviation safety regulations, over half of the participants were very dissatisfied (38%) or dissatisfied (20%) with the CASA's development of aviation safety regulations (Colmar Brunton, 2016).

As a summary of this section, the following paragraphs discuss research comparing hospitals that displayed varying degrees of implementation of the World Health Organisation's (WHO) surgical safety checklist. This checklist incorporated a CRM-style briefing and debriefing with a surgical pause (Wakeman & Langham, 2018). Conley, Singer, Edmondson, Berry, and Gawande (2011) conducted qualitative research examining implementation of the WHO surgical safety checklist in five Washington State hospitals to identify certain characteristics that led to effective implementation. The study suggested that certain factors led to highly efficient and successful implementation of the surgical safety checklist use, such as 'active leadership … extensive discussion and training, piloting, multidisciplinary communications, real time coaching, and ongoing feedback' (Conley et al., 2011, p. 877).

A hospital with highly effective implementation of the surgical safety checklist exhibited thorough preparation prior to implementation, promoted open and regular dialogue between the implementation team and stakeholders (e.g., surgeons and nurse managers), provided extensive discipline-specific training and ongoing real-time coaching and feedback throughout the implementation period, and closely monitored progress (Conley et al., 2011). In short, the implementation was highly successful.

In contrast, a hospital that experienced less effective implementation of the surgical safety checklist exhibited largely opposite characteristics to the hospital with highly effective implementation (Conley et al., 2011). The hospital with less effective implementation did not provide dedicated education or training before implementation, and did not establish real-time coaching, feedback or an ongoing monitoring mechanism (Conley et al., 2011). Consequently, the frequency and quality of checklist use were poor. One of the participants remarked, 'different people do it in different ways' (Conley et al., 2011, p. 877), thereby suggesting there was no standardised approach in terms of the use of the surgical safety checklist, as a result of poor implementation.

The above two cases of the hospitals highlight the importance of a thoughtful process to enable the successful implementation of an initiative, and it was unfortunate that CASA's TEM implementation more closely resembled the hospital that experienced less effective implementation of the checklist. Despite the poor implementation of TEM, it became a mandatory licence issue flight test item in 2009; thus, it was required to be taught, assessed and practised, as discussed in the following section.

5.4.5 Threat and Error Management in practice.

This section discusses how TEM has been taught and practised in Australia since its implementation in 2009. One of the findings from Study 1 indicated that TEM was incorporated in ground schools, pre-/post-flight briefings and instructional flights. Ground school[7] refers to a theory delivery course. The syllabus for the ground school is based on Part 61 of MOS (CASA, 2018f). This manual comprises eight schedules, and Schedule 3 prescribes the aeronautical knowledge standards for all Part 61 licences (e.g., PPL and CPL), as well as ratings and endorsements. Table 5.6 provides a list of TEM-specific items that are required to be covered during the ground school. It is evident that almost all items are identical between the PPL syllabus and CPL syllabus.

Table 5.6 *Syllabus Items for TEM from Part 61 of the Manual of Standards: adapted from CASA (2018e).*

PPL	CPL
• Describe the basic principles of TEM	
• Explain the principles of TEM and detail a process to identify and manage threats and errors during single-pilot operations	• Explain the principles of TEM and detail a process to identify and manage threats and errors during single-pilot operations
• Define ‘threat’ and give examples of threats	• Explain the meaning of ‘threat’ and give examples of threats
• Give an example of a committed error and how action could be taken to ensure safe flight	• Give an example of a committed error and how action could be taken to ensure safe flight
• Explain how the use of checklists and SOPs can prevent errors	• Explain how the use of checklists and SOPs can prevent errors
• Give examples of how a UAS can develop from an unmanaged threat or error	• Describe how a UAS can develop from an unmanaged threat or error
• Explain what resources a pilot could identify and use to avoid or	• Explain what resources a pilot could identify and use to avoid or

[7] Ground school is similar to scheduled lectures/classes where theoretical contents for relevant licences (e.g., PPL and CPL) are taught.

manage a UAS, such as being lost or entering adverse weather	manage a UAS, such as being lost or entering adverse weather
• Explain the importance of ensuring that tasks are prioritised to manage a UAS	• Explain the importance of ensuring that tasks are prioritised to manage a UAS
• Give examples of how establishing and maintaining interpersonal relationships can promote safe flight	• Describe how establishing and maintaining interpersonal relationships can promote safe flight

TEM is also covered in pre-/post-flight briefings. The pre-flight briefing is an abbreviated version of a long briefing, and typically takes around 15 minutes. Although TEM can be discussed during a long briefing, it is more appropriate to be covered as part of a pre-flight briefing prior to flight because:

> *'Threat and error management better lives in a pre-flight briefing where we are able to actually consider environmental factors of the day and the two organic beings on board the aircraft and the actual condition of the aircraft in which we are flying on that day—none of which is known at a long briefing stage. So I prefer that it will create a discussion threat and error management in a pre-flight briefing stage, despite the fact that, um, you know, in our long briefing format, there is a requirement for TEM.' (Participant 4)*

During the pre-flight briefing, a FI and trainee pilot usually go through the flight sequences. According to Participant 5, the pre-flight briefing concludes with a TEM consideration:

> *'We modified all our briefings and, um, we just put TEM as an addition to the briefings. Okay, so the actual pre-flight briefing didn't change, but we incorporated TEM at the end of the briefings ... Now, the last thing before you walk out the door, let's see what can turn to worms.'*

There are other forms of briefings immediately before or during flight, such as before-take-off briefings, emergency briefings and approach briefings. These briefings are typically verbalised at the appropriate phase of flight. For example, a typical emergency briefing for a single-engine aeroplane occurs before lining up on an active runway, and proceeds as follows:

> *'Any malfunctions before airborne, close the throttle, keep the aeroplane straight and stop within the remaining runway. Engine failure after airborne with sufficient runway remaining, check forward and re-land. Engine failure after airborne without sufficient runway remaining, check forward, look for a suitable paddock within 30 degrees*[8] *either side of the nose and glide to the paddock.' (Provided by the researcher)*

Verbalising the emergency briefing serves as a clear action plan in the unlikely event of engine failure occurring, so that pilots are better prepared and can react immediately. However, Participant 3 described these briefings as *'the ritual chant to please the gods'*, and pointed out that these briefings are not completed in the way they are intended to be:

> *'The other ritual chant to please the gods was the before-take-off and emergency briefs … So we're sitting, at a run-up point, and they say, 'okay, I'm taking off from Runway 10, left or right, and I'm going to, my, abort speed is whatever it was and then once I'm, if I'm short of that, then I'm going to stop on the runway', and if it's passed that point, and again, everybody uses exactly the same practised spiels, and said, 'after that, I'm going to lower the nose, look 30 degrees either side of the nose for a suitable place to land and I'll land on that suitable point'. And so, at the end of that … I would say—I don't know how many times I said this—so, 'okay, we're taking off on 10 Archerfield, off the end of the runway, there is a whole industrial area, full of factories.' (Participant 3)*

Participant 3 continued that:

> *'Instead of repeating the datum, actions after take-off, let's consider what your options are after take-off because what you say, the briefing, is*

[8] The actual angle varies depending on surroundings from the runway end of departure.

completely ineffectual [fictitious], as soon as you have an engine failure and you point it to the factories, you're going to do something else anyway … so now you're making a decision under pressure at the wrong time, you're very likely to make a wrong decision … that's an example I had given it to the students and said, 'don't do it that way.' (Participant 3)

The above comment illustrates that the briefings were not given the deserved attention and consideration, and were not what they were intended to be. Participant 5 added that:

'Just the old parrot fashion, you know, pre-take-off checks, you know, some of them just rote, just rattle it off and, uh, I absolutely hate that, you know, it's rehearsed … Um, when things become rote, unfortunately, they are not practised conscientiously … they're not actually sitting there, scratching their head: 'now, what could go wrong, you know, in the next five minutes when I was going to take off? What could possibly go wrong?' That's what we're trying to do with this one.' (Participant 5)

After TEM demonstration became a flight test requirement, TEM briefings were incorporated into the relevant briefings. For instance, the approach briefing involves verbalising a rejoining procedure based on an active runway and TEM considerations, such as a large number of traffic departing, arriving and operating in a circuit pattern. Participant 1 described a good example of TEM demonstration in this phase of flight:

'A classic example is coming into one of, say, one of the VFR entry points back to Archerfield here, which is, you know, notoriously busy around those locations … You have traffic coming outbound, you are coming inbound and you get a lot of traffic, uh … chopper, helicopter at low level, helicopter traffic around there, so naturally it is a quite busy area and everything. A good student will prepare himself to get his ATIS[9] *information well out of the way and all copied up and everything well in*

[9] Automatic Terminal Information Service (ATIS) is a continuous broadcast of relevant operational information, such as a runway in use and meteorological conditions for a specific aerodrome.

> *advance, so when actually hitting the hot spot, if you like, he will get whatever checklist is done out of the way, make broadcast calls early to alert people that he is in the area and everything. He makes sure he is on the right frequency so he is able to listen out and deal with those potential threats, having recognised in the first place the threats in the area, error management is covered by good focus on SOPs, say, getting his calls out the way, getting his checklist done, those sorts of things, so, yeah, that's what the good student does.' (Participant 1)*

In contrast, Participant 2 described TEM demonstration as: *'They mouth this off, but they don't really understand what they are actually telling you'*. Participant 3 provided the following example of a poor TEM demonstration:

> *'… to demonstrate it that they were using threat and error management, you're coming in from a training area, uh, at Archerfield, and, at some stage, they've got to show that they are aware of threats and how they need to be managed, so I would get this ritual chant—it was all learnt off by heart and that was, that was practically the same, everybody from a school would say virtually exactly the same words. That is, 'I'm coming into Park Ridge … it's going to be traffic likely from the area converging here … so I'll have to keep my eyes on that traffic and be aware of what they're doing, and so they're letting me know that they're aware and thinking about it … Once we passed here, I will make a call, I will be given rejoin instruction. I'm expecting to land on Runway 10 or 28 and so I will go to this point and then I will come down and then I'll come around to land on the left or right runway and then I'll turn off at a taxi-way', whatever, because … [indiscernible] … taxi via, taxi to this and that and this comes back to my whole point. So, they've given me whole this spiel and I'm sitting there and I'm watching this aircraft, converging on us from my right … But he was so busy with saying his spiels, you know, he wasn't actually doing threat and error management by keeping his eyes open … So the whole briefing meant nothing, which means they were wasting their time, trying to impress me with their threat and error management, and the whole thing was pointless because, concentrating on that, they didn't*

> *see the aeroplane and telling me what they were going to do after they land and the runway was changed and everything is out the window.' (Participant 3)*

Participant 3 explained that flight test candidates had to verbalise these briefings because:

> *'… the way of doing all that was, unless I can hear you say something that shows me that you are doing it, I can't tick that box. Whereas I used to say, I can see whether you are managing or not, you know.' (Participant 3)*

Participant 3 expressed concern with the prevalence of this practice:

> *'They were told to give this spiel and this happened in Archerfield at different schools, uh, Townsville, Rockhampton, at Toowoomba and the similar process for whole time because they felt that they had to produce by telling me all about it.' (Participant 3)*

As Participants 3 and 5 stated, these briefings were rote learnt, and usually derived from FIs who provided written scripts for their students to learn. Participant 4's comment highlighted the inappropriateness of completely artificial written scripts for students to verbalise:

> *'Which is rubbish. It's a completely waste of time. The issue with scripting anything, particularly if that particular scenario or external factor does not actually present themselves on that day, is that now you've got a student trying to assess simulated conditions, taking them away from actually using their brains and assessing real conditions. Um, in fact, they may think that by working through their scenario or the script, that their job is actually done, and make you have no further thought to actually considering threats and errors for the day.' (Participant 4)*

Participant 4 offered a better way to use a script to prompt which actions need to be completed:

> *'… we are simply providing them with place holders to identify for themselves threats and errors and put them into the script, so if you had a script that sort of said, you know, the fact that's affecting today's flying are: blank, blank and blank, and the wind is blank, and therefore it is or it is not, you know, blank, um, I think that encourages students to think and it directs their thought process to consideration of threats and errors.' (Participant 4)*

Based on the participants' comments, it was evident that poor TEM implementation has negatively affected the way TEM is taught and practised in Australian general aviation. The participants' main criticism involved individuals verbalising a series of briefings without conscientiously thinking about the real-life situation. Use of completely written scripts that may not resemble conditions on a particular day was another area that needs immediate attention.

5.5 Chapter Summary

This chapter has discussed the qualitative phase of sequential mixed methods research, with the central focus of better understanding how TEM was implemented in Australian general aviation through interviewing five highly experienced FEs. All participants shared the same view that TEM was not implemented well, and four themes arose from the analysis: impracticality, lack of support and guidance, TEM implementation and TEM in practice. Based on the findings, a further study was designed to verify and generalise the findings from the present study. The third study (quantitative study) is presented in the next chapter.

Chapter 6: Study 3 (Quantitative Study)

6.1 Purpose

This study involved a quantitative (second) phase of sequential mixed methods research approach using an exploratory sequential design (Figure 3.6), and its central purpose was to verify and generalise the findings from the previous study (Study 2). To fulfil this purpose, an online survey was devised based on the four themes that arose from the thematic analysis in Study 2 (qualitative study): impracticality, lack of guidance and support, TEM implementation and TEM in practice. The combined findings from studies 2 and 3 are expected to provide a better understanding of how TEM was implemented and this implementation's effects on the way TEM has been viewed and practised in Australian general aviation.

6.2 Methods

6.2.1 Design overview.

The primary purpose of this study was to collect information to verify and generalise the findings from Study 2 (qualitative study). To collect the required data, an online survey was devised and administered. The survey contained six questions for each theme, as well as group-specific questions for FIs and FEs, based on comments made by the participants in Study 2. The survey also collected demographic data, such as the highest CASA-issued licence each participant either held or had held, and whether the participants were actively flying, either commercially or privately, before TEM training became mandated. The entire survey is presented in Appendix 9 so that readers have a better understanding of the contents and structure of the survey.

6.2.2 Trustworthiness in research.

Lincoln and Guba (1985) proposed that trustworthiness is one of the ways to persuade the intended audiences that the findings are 'worth paying attention to' (p. 290). According to Lincoln and Guba (1985), there are four assessment criteria used to achieve trustworthiness: truth value, applicability, consistency and neutrality. It was established in Section 5.2.3 that these four assessment criteria appear differently in quantitative and qualitative research. In Section 5.2.3.1, trustworthiness in qualitative research was described in detail. In this section, a brief description of the four assessment criteria applicable to quantitative research (Table 6.1) is provided, followed by more detailed descriptions in Section 6.2.2.1.

Table 6.1 *Comparison of Assessment Criteria for Trustworthiness in Quantitative Research.*

Criterion	Quantitative approach
Truth value	Internal validity
Applicability	External validity
Consistency	Reliability
Neutrality	Objectivity

Source: adapted from Krefting (1991, p. 217).

Truth value refers to researchers' confidence in the truth of the findings, based on the research design, participants and context (Krefting, 1991). In quantitative research, internal validity is the alternative criterion for truth value, and is supported when the controlled variation in an independent (cause) variable accounts for changes in a dependent (effect) variable (Lincoln & Guba, 1985).

Applicability refers to how well findings can be generalised to alternative contexts, settings or groups (Krefting, 1991). In quantitative research, applicability is termed external validity, and refers to generalisation of research findings from small samples to a larger population.

Consistency is the third criterion of trustworthiness, and considers how consistent the findings of research are if repeated involving the same samples of participants or in a similar context (Krefting, 1991). In quantitative research, reliability is deemed to be equivalent to consistency, which emphasises 'the value of repeatability', whereby replication of the same study would produce the same or very similar findings (Krefting, 1991, p. 216), even if the methods of study were altered (e.g., different groups of participants).

Neutrality is the fourth and final criterion of trustworthiness, and determines the legitimacy of findings without the 'biases, motivation, interests or perspectives' of researchers (Lincoln & Guba, 1985, p. 290). In quantitative research, neutrality is termed objectivity, and is generally achieved through a high level of intersubjective agreements and the use of appropriate methodology that minimises such violations (Lincoln & Guba, 1985). To this end, quantitative researchers are encouraged to maintain appropriate distance from their subjects to avoid unintended influences, such as biases. The following subsection further discusses in-depth the four assessment criteria to achieve trustworthiness in quantitative research.

6.2.2.1 Trustworthiness in quantitative research.

Lincoln and Guba (1985) suggested four questions that researchers should ask themselves to ensure the trustworthiness of their research and its findings and interpretations. According to Lincoln and Guba (1985), the first question that researchers can pose to ensure 'truth value' is: 'How can one establish confidence in the "truth" of the findings of a particular inquiry for the subjects (respondents) with which and the context in which the inquiry was carried out?' (p. 290). This assessment criterion for trustworthiness is more commonly known as internal validity in quantitative research, which refers to the best approximation of 'validity of the cause-and-effect inference linking the independent variable and the dependent variable' (Wiersma & Jurs, 2009, p. 139). This is achieved by a robust research design that minimises false conclusions (Neuman, 2011). However, it is worth noting that this is an approximation, rather than an absolute truth, because the ultimate test of internal validity for quantitative researchers is to show 'an isomorphism (a one-to-one relationship) with that reality' that cannot be proved (Lincoln & Guba, 1985, p. 294). For quantitative researchers to achieve the best approximation, eight threats to internal validity, summarised by Campbell and Stanley (1963), should be considered, managed and mitigated. These eight threats are presented in Table 6.2.

Table 6.2 *Threats to Internal Validity.*

Threat	Example
1. History—unanticipated events occurring while the experiment is in progress that affect the dependent variable	During a relatively short instructional experiment, one group of subjects misses instruction because of a power failure at the school
2. Maturation—processes operating within the subject as a function of time	In a learning experiment, subject performance begins decreasing after about 50 minutes because of fatigue
3. Testing—the effect of taking one test on the scores of a subsequent test	In an experiment with the dependent variable of performance on a logical reasoning test, a pre-test cues the subjects about the post-test
4. Instrumentation—an effect because of inconsistent use of the measuring instruments	Two examiners in an instructional experiment administer the post-test with different instructions and procedures
5. Statistical regression—an effect caused by a tendency for subjects selected based on extreme scores to regress towards an average performance on a subsequent test	In an experiment involving reading instruction, subjects grouped because of poor pre-test reading scores show considerably greater gains than the average and high readers
6. Selection—an effect because of the groups of subjects not being randomly assigned to groups; a selection factor is operating such that the groups are not equivalent	The experimental group in an instructional experiment consists of a high-ability class, while the control group is an average-ability class
7. Mortality—an effect because of subjects dropping out of the experiment on a non-random basis	In a health experiment designed to determine the effects of various exercises, those subjects finding exercise most difficult stop participating
8. Selection–maturation interaction—an effect of maturation not being consistent across the groups because of some selection factor	In a problem-solving experiment, intact groups of junior and senior high school students are involved, and the junior students tire of the task sooner than the older students

Source: adapted from Wiersma and Jurs (2009, p. 141).

The present study used a short online survey to gather the required data. When the survey was designed and reviewed careful consideration was given to ensure that no threat listed in Table 6.2 adversely influenced the survey findings. For instance, all participants were presented with the same survey questions (to satisfy 'instrumentation' in table 6.2) and the participants in the survey varied in terms of their involvement in general aviation (to satisfy 'selection' in table 6.2). Further, the online survey was relatively short (maturation) and participants were able to save parts of their responses and complete the rest later (history).

According to Lincoln and Guba (1985), the second question that quantitative researchers should ask themselves to satisfy the assessment criterion for trustworthiness and to ensure 'applicability' is: 'How can one determine the extent to which the findings of a particular inquiry have applicability in other contexts or with other subjects (respondents)?' (p. 290). This assessment criterion for trustworthiness is more commonly known as external validity in quantitative research, and refers to generalisation of the research findings to a large population (Neuman, 2011). It is noteworthy that there is an inverse relationship between internal validity and external validity. If a study is designed for stronger internal validity, then external validity would be weaker, and vice versa. Therefore, there should be an optimum balance between these two types of validity. Quantitative researchers aim for their findings to be comparable and transferable (Wiersma & Jurs, 2009). To achieve this, four threats should be considered, managed and mitigated, as summarised by Campbell and Stanley (1963). These four threats are presented in Table 6.3.

Table 6.3 *Threats to External Validity.*

Threat	Example
1. Interaction effect of testing—pretesting interacts with the experimental treatment and causes some effect, such that the results will not generalise to an un-pretested population	In a physical performance experiment, the pre-test cues the subjects to respond in a certain way to the experimental treatment that would not be the case if there were no pre-test
2. Interaction effects of selection biases and the experimental treatment—an effect of some selection factor of intact groups interacting with the experimental treatment that would not be the case if the groups were formed randomly	The results of an experiment in which teaching method is the experimental treatment is effective with low achievers, yet not as effective with high achievers
3. Reactive effects of experimental arrangements—an effect that is due to the artificial or novel experimental settings (note that this can also threaten internal validity)	An experiment in remedial reading instruction has an effect that does not occur when the remedial reading program is implemented in the regular classroom

4. Multiple-treatment interference—when the same subjects receive two or more treatments (as in a repeated measure design), there may be a carryover effect between treatments, such that the results cannot be generalised to single treatments	In a drug experiment, the same animals are administered four different drug doses in some sequence; the effects of the second and fourth doses cannot be separated from possible (delayed) effects of the preceding doses

Source: Adapted from Wiersma and Jurs (2009, p. 141).

Again, this study used an online survey and carefully examined each threat to ensure they did not appear to adversely affect the external validity of the findings from the online survey. For instance, a closer examination of the survey response data showed no multiple attempts (satisfying multiple-treatment interface in table 6.3). Further, the participants in the survey varied in terms of their involvement in general aviation (interaction effects of selection biases and the experimental treatment).

The third question that Lincoln and Guba (1985) suggested for quantitative researchers to consider to ensure consistent research and findings is: 'How can one determine whether the findings of an inquiry would be repeated if the inquiry were replicated with the same (or similar) subjects (respondents) in the same (or similar) context?' (p. 290). This assessment criterion for trustworthiness refers to reliability in quantitative research. It suggests that, if an experiment is repeated over and over under the same (or very similar) conditions, the results will be the same (or very similar). Neuman (2011) offered four ways to improve reliability: clearly conceptualise constructs, use a precise level of measurement, use multiple indicators and use pilot tests.

The first way to improve reliability—clearly conceptualise constructs—refers to developing clear and specific theoretical definitions, so that each measure indicates one concept (Neuman, 2011). For instance, measuring a quantity of red paint in a glass beaker is more reliable than attempting to measure a quantity of red paint that is mixed with blue paint in the same glass beaker. The current study's survey questions were based on the themes identified in the previous study, and each question was worded in a way that specifically addressed one particular theme only. The appropriateness of the survey questions was reviewed by two of the researcher's supervisors. In addition, the questions were peer-reviewed by a PhD candidate who was independent from this study.

The second way to increase reliability—use of a precise level of measurement—refers to an idea that the higher the level of measurement, the more accurate and reliable the measurement will be (Neuman, 2011). Using the glass beaker example again, a glass beaker with 10 mm increments will measure a quantity of red paint more accurately, and so the measures will be more reliable, than a glass beaker with 10 cm increments. Surveys typically collect ordinal data using Likert scales with various numerical scales (e.g., five- and seven-point scales). The online survey used in this study adopted a five-point Likert scale because Dawes (2008) found that there were no significant variances in data characteristics from five-, seven- or 10-point scaled surveys, yet five-point scales provided better quality data than did seven-point scales (Revilla, Saris, & Krosnick, 2013).

The third way to enhance the level of reliability—using multiple indicators—encourages researchers to collect several measures for a construct under investigation, so that hypothesis testing and its interpretations are more dependable and consistent through triangulation of measures (Newman, 2011). For instance, Koglbauer, Kallus, Braunstingl and Boucsein (2011) measured several dependent measures such as instructor ratings of participant's performance, subjective workload using a German version of the NASA Task Load Index (NASA-TLX) and physiological measures (e.g., chest electrocardiogram [ECG] and electrodermal activity [EDA]) to test training effects of anticipation-based training in real and simulated flight. The present study is a quantitative (second) phase of sequential mixed methods research (exploratory sequential design) that collects responses from the survey questions based on findings from a qualitative (first) phase of research. Therefore, this research design is expected to further enhance the level of reliability of the findings by analysing both quantitative and qualitative data.

The fourth way to increase the level of reliability—use of pilot tests—involves developing one or more preliminary versions and testing these before applying the final version for data collection and analysis (Neuman, 2011). This process is time consuming, yet can save time and effort in the long term, as the final version is more reliable and researchers can have more confidence in their findings. In the current research, the use of an exploratory study (Study 1) and the qualitative (first) phase of sequential research design (Study 2) were deemed to satisfy the increased level of reliability.

The final question that Lincoln and Guba (1985) suggested when considering assessment criteria is: 'How can one establish the degree to which the findings of an inquiry are determined by the subjects (respondents) and conditions of the inquiry and not by the biases, motivations, interests, or perspectives of the inquirer?' (p. 290). This assessment criterion for 'trustworthiness' refers to objectivity in quantitative research. Lincoln and Guba (1985) offered two criteria to ensure objectivity: establishing intersubjective agreement and adopting appropriate methodology and methods. The former is based on the notion that, if multiple observers agree on a certain phenomenon, their shared view will be claimed to be objective (Lincoln & Guba, 1985). In the case of the latter, the objectivity of the research may be compromised if inappropriate or lax methodology and methods are used that may introduce unplanned bias and preconceived results (Lincoln & Guba, 1985). It is generally accepted that, if the first three criteria (internal and external validity and reliability) are appropriately addressed, then objectivity is naturally ensured.

6.2.3 Participants.

A total of 102 participants completed this study's online survey. However, five participants indicated that they did not primarily fly VH-registered aircraft[10], which suggested that their involvement in general aviation was in recreational aviation. Given that TEM was not formally introduced by CASA in the recreational aviation sector, these five responses were excluded from analysis. Therefore, a total of 97 responses were analysed for this study. Table 6.4 presents a summary of the participants' demographic information.

[10] Aircraft registered in Australia start with 'VH' which is the nationality mark allocated to Australia. However, aircraft registered under Recreational Aviation Australia (RA-Aus) do not have 'VH' in their registration number.

Table 6.4 *Demographic Information of Participants.*

Demographic features	Frequency	Per cent
Involvement in general aviation		
• Private operation	14	14.4
• Flight training	39	40.2
• Agricultural	9	9.3
• Small charter operation	13	13.4
• Other[11]	22	22.7
Actively flying before TEM was mandated in July 2009?		
• Yes	79	81.4
• No	18	18.6
Average flying hours per month		
• Less than 10 hours	23	23.7
• 10~50 hours	49	50.5
• 51~100 hours	20	20.6
• More than 100 hours	5	5.2
Highest CASA licence hold/have held		
• No licence (student pilot)	2	2.1
• PPL	11	11.3
• CPL	52	53.6
• ATPL	32	33
Holder of a FI rating?		
• Yes	46	47.4
• No	51	52.6
Holder of a FE rating?		
• Yes	30	30.9
• No	67	69.1

6.2.4 Materials.

The main part of the survey contained a number of questions based on the four themes identified in Study 2. Impracticality was the first theme that arose from the thematic analysis in Study 2. General aviation is primarily a vocationally based industry; thus, it is important to consider practical aspects when new initiatives are proposed and implemented in general aviation, such as mandating TEM. However, all participants from Study 2 expressed concerns about a lack of practicality in terms of understanding the concept of TEM and implementing TEM in practice, whether in daily routine

[11] The majority of others encompassed Royal Flying Doctor Service (RFDS)pilots, mustering pilots, rescue pilots and firefighting pilots.

flying or flight training. The six questions in Table 6.5 were formulated to verify and generalise the first theme.

Table 6.5 *Six Questions on Impracticality.*

Themes	Questions
Impracticality	1. CASA has appropriately translated TEM concepts into practical guidance to general aviation. 2. There are adequate practical examples relating to TEM principles and concepts for general aviation. 3. When it comes to flying operations within general aviation, CASA places greater emphasis on the theoretical aspects of flying than the practical aspects. 4. Senior executives and decision makers within CASA have a good understanding of how general aviation operates. 5. I find it easy to keep up to date with changes to aviation safety regulations/initiatives. 6. I am satisfied with the way CASA develops aviation safety regulations/initiatives.

Lack of guidance and support was the second theme that arose from the thematic analysis in Study 2. All participants in Study 2 agreed that there was inadequate guidance provided by CASA when TEM was implemented in Australian general aviation. The six questions in Table 6.6 were formulated to verify and generalise the second theme.

Table 6.6 *Six Questions on Lack of Guidance and Support.*

Themes	Questions
Lack of support and guidance	1. CASA provides adequate training and guidance material on TEM training for general aviation. 2. The concept of TEM is easy to understand. 3. CASA explains TEM and how it affects my role and/or activities in a clear and succinct manner. 4. I regularly attend CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives. 5. I find it useful to attend seminars, such as AvSafety and/or others organised by CASA. 6. I find the number of safety-related publications and resources from CASA useful.

TEM implementation was the third theme that arose from the thematic analysis in Study 2. All participants in Study 2 shared their observations and experience that TEM implementation in general aviation was poor. The six questions in Table 6.7 were formulated to verify and generalise the third theme.

Table 6.7 *Six Questions on TEM Implementation.*

Themes	Questions
Poor implementation	1. CASA consults with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM. 2. CASA aviation safety advisors have a consistent understanding of TEM principles. 3. I am satisfied with the way CASA implements new aviation safety regulations/initiatives. 4. The number of CASA staff with appropriate expertise in TEM is adequate. 5. CASA maintains an effective ongoing dialogue with the general aviation industry. 6. CASA explains new safety regulations/initiatives and how they affect the general aviation industry in a clear and succinct manner.

TEM in practice was the fourth theme that arose from the thematic analysis in Study 2. The findings from Study 2 suggested that poor TEM implementation, as well as impracticality and lack of guidance and support, had adversely affected the way TEM was taught and practised in Australian general aviation. The six questions in Table 6.8 were formulated to verify and generalise the fourth theme.

Table 6.8 *Six Questions on TEM in Practice.*

Themes	Questions
TEM in practice	1. I use TEM principles when I am undertaking flying activities. 2. I have a clear understanding of what I need to do when applying TEM principles. 3. Consideration of TEM is an important part of my flight preparation. 4. I always perform a series of briefings (e.g., emergency and take-off briefings), taking into account the prevailing conditions and situations. 5. I find TEM briefings an unnecessary additional task. 6. I can see the value in TEM when I am undertaking flying activities.

In addition to the questions for the main themes, there were two additional questions, as well as group-specific questions, based on various comments from the participants in Study 2. The list of these additional questions is presented in Table 6.9. The group-specific questions were for current or former FEs and FIs.

Table 6.9 *Additional and Group-specific Questions.*

Category	Questions
Additional questions	1. The benefits of TEM training are overrated. 2. The use of TEM training/principles should be expanded to other sectors, such as recreational aviation and RPAS sectors.
FEs	1. CASA provided adequate guidance to examiners before TEM became a mandatory flight test item. 2. During flight tests, assessment of TEM is standardised among FEs. 3. During flight tests, candidates verbalise their briefings (e.g., emergency briefing) without implementing conscientious thought.
FIs	1. During instructional flights, my students verbalise their briefings (e.g., emergency briefing) without implementing conscientious thought. 2. I find the CASA's FI manual useful. 3. I have a clear understanding of what I need to do when teaching TEM principles to my students. 4. My organisation provides adequate mentoring opportunities for junior FIs.

All the above survey questions had five response options: strongly disagree, disagree, neither agree nor disagree, agree and strongly agree. For the purpose of statistical analysis, numerical values were assigned to each of the potential responses as follows: 'strongly disagree' = 1, 'disagree' = 2, 'neither agree nor disagree' = 3, 'agree' = 4 and 'strongly agree' = 5.

6.2.5 Procedure.

An email invitation with a web link for the online survey was sent to organisations that were listed on the Aeroclub directory website (www.aeroclub.com.au/directory/flying-clubs), where directory tags contained 'Aeroclub' and 'aeroplane' and/or 'helicopter'. In addition, the same email invitation was sent to CASA-approved FEs found on the CASA website. Personal contacts on Facebook and LinkedIn were also used to increase the number of participations.

The front page of the survey contained participant information, and stated the purpose of the survey, expected benefits of the survey, risk of involvement in the survey, and privacy and confidentiality issues (Appendix 9). The survey used an implied consent method to obtain consent from the participants, whereby submission of the completed survey was an indication of consent to participate in the study. No remuneration was offered as an incentive to participate in the survey. Given that this study was considered to involve the minimal level of risk and ethical issues, in accordance with the National Statement of Ethical Conduct in Research Involving Humans, the online Expedited Ethical Review Checklist was completed when applying for ethics approval. Acknowledgement of ethics approval is presented in Appendix 10.

6.2.6 Hypotheses.

Based on the findings from Studies 1 and 2, the following hypotheses were tested in this study:

H_1: Survey participants were more in agreement that TEM lacks practicality.

H_2: Survey participants were more in agreement that there was a lack of support and guidance from CASA.

H_3: Survey participants were more in agreement that TEM implementation in Australian general aviation was poor.

H_4: Survey participants were more in agreement that TEM was poorly practised in Australian general aviation.

H_5: Impracticality negatively affected the way TEM was viewed and practised.

H_6: Impracticality resulted in survey participants' view of poor TEM implementation.

H_7: Lack of guidance and support from CASA negatively affected the way TEM was viewed and practised.

H_8: Lack of guidance and support from CASA resulted in participants' view of poor TEM implementation.

H_9: Poor TEM implementation mediated the effect of impracticality on the way TEM was viewed and practised.

H_{10}: Poor TEM implementation mediated the effect of lack of guidance and support from CASA on the way TEM was viewed and practised.

6.3 Results

Statistical analyses in this section were performed in two phases. The first phase was a basic level statistical analysis to give the researcher a high-level overview of the survey data. The second phase was a more advanced level to provide a holistic view of the data and the effects of themes and their interrelationships on the way TEM was viewed, taught and practised.

6.3.1 Results (first phase).

SPSS (Version 25 for Apple Macintosh computers) was used with the level of significance, alpha, set to be $p < .05$ for all statistical analyses. The results were presented in a combination of descriptive and inferential statistics.

6.3.1.1 Theme 1: Impracticality.

All five participants from Study 2 expressed concerns of a lack of practicality in terms of understanding the concept of TEM and implementing it in practice, whether in daily routine flying or flight training. To test whether this view was supported by the survey participants in this study, both descriptive and inferential statistics are provided below. For inferential statistics, a one-sample Wilcoxon signed rank test with a hypothesised median of 3 (neither agree nor disagree) was performed for each of the six questions for the first theme—impracticality. This was undertaken to compare the mean score for each of the six questions with the mean score of the scale of 3. In addition, a Mann-Whitney test was performed for each of the six questions to determine whether there were statistical differences in responses between: (i) the survey participants who were actively flying before TEM was introduced in 2009 and the survey participants who were not; (ii) the survey participants who held or had held a FI rating and the non-FI rating holders; and (iii) the survey participants who held or had held a FE rating and the non-FE rating holders. A Kruskal-Wallis test was also performed for each of the six questions to determine whether there were statistical differences in the responses among survey participants with different flying licences[12] (PPL, CPL and ATPL).

For the first question on impracticality, a total of 63 (64.9%) survey participants disagreed or strongly disagreed that CASA had appropriately translated TEM concepts into practical guidance for general aviation (Figure 6.1). The result from the one-sample Wilcoxon signed rank test was statistically significant ($p < .001$), thereby suggesting that a statistically significant number of survey participants disagreed that CASA had appropriately translated TEM concepts into practical guidance for general aviation.

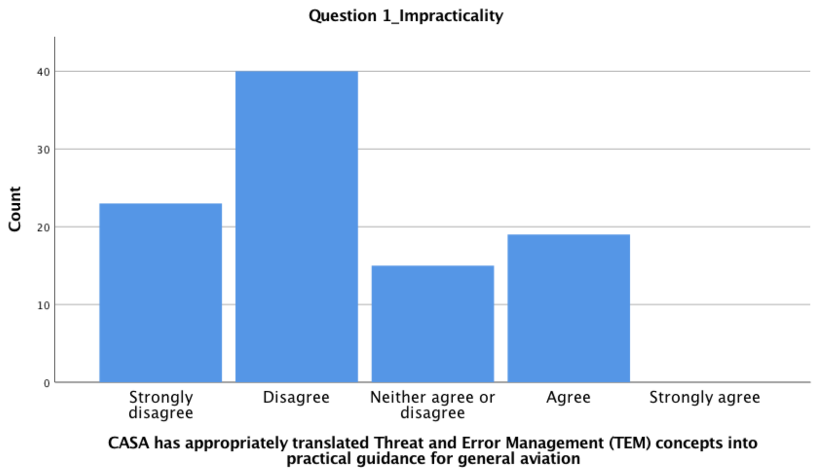

Figure 6.1. Responses to the first question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between survey participants who were actively flying before TEM was formally introduced in 2009 (pre-TEM) and those who were not (post-TEM). The result showed that the pre-TEM group did not significantly differ from the post-TEM group in their responses ($U = 885.5$, $z = 1.7$, $p = .09$, $r = .17$), thereby suggesting that both groups shared a similar view that CASA did not appropriately translate TEM concepts into practical guidance for general aviation.

[12] Student pilots were excluded from the analysis because there were only two students involved, and their level and length of involvement in general aviation may not have been sufficiently lengthy to gauge meaningful insights.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the survey participants who held or had held a FI rating (FI) and those who did not (non-FI). The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}098$, $z = -.57$, $p = .57$, $r = -.02$), thereby suggesting that both groups shared similar opinions that CASA did not appropriately translate TEM concepts into practical guidance for general aviation.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the survey participants who held or had held a FE rating (FE) and those who did not (non-FE). The result again showed that the FE group did not significantly differ from the non-FE group in their responses ($U = 991$, $z = -.12$, $p = .91$, $r = -.01$), thereby suggesting that both groups were more in disagreement that CASA had appropriately translated TEM concepts into practical guidance for general aviation.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the survey participants with different flying licences (PPL, CPL and ATPL). The result from the test suggested that the overall responses for Question 1 varied according to groups: $H(2) = 7.94$, $p = .02$. A post hoc pairwise comparison with adjusted p-values showed no significant differences between CPL and ATPL holders ($p = 1.00$, $r = -.05$) and between PPL and ATPL holders ($p = .06$, r =. 24). However, there was a significant difference in the overall responses between PPL and CPL holders ($p = .02$, $r = -.29$), thereby suggesting that a relatively greater number of CPL holders did not agree that CASA had appropriately translated TEM concepts into practical guidance for general aviation.

For the second question on impracticality, a total of 52 (53.6%) participants strongly disagreed or disagreed that there were adequate practical examples relating to TEM principles and concepts for general aviation (Figure 6.2). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that relatively more participants disagreed that there were adequate practical examples relating to TEM principles and concepts for general aviation.

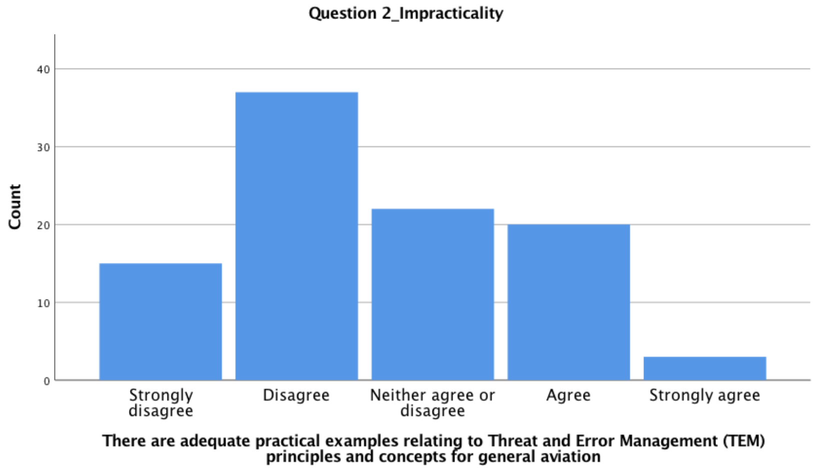

Figure 6.2. Responses to the second question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result indicated that the pre-TEM group did not significantly differ from the post-TEM group in their responses (U = 712.5, z = .02, p = .99, r = .00), thereby suggesting that both groups shared a similar view that there were not adequate practical examples relating to TEM principles and concepts for general aviation.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses (U = 1,039.5, z = -1.01, p = .32, r = -.10), thereby suggesting that both groups shared similar opinions that there were not adequate practical examples relating to TEM principles and concepts for general aviation.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses (U = 908, z = -.79, p = .43, r = -.08), thereby suggesting that both groups were more in agreement that there were not adequate practical examples relating to TEM principles and concepts for general aviation.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 2 did not vary according to groups ($H(2) = 2.78$, $p = .25$), thereby suggesting that all three groups shared a similar view that there were not adequate practical examples relating to TEM principles and concepts for general aviation.

For the third question on impracticality, a total of 71 (73.2%) survey participants agreed or strongly agreed that CASA placed greater emphasis on the theoretical aspects of flying than the practical aspects (Figure 6.3). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that the participants were more in disagreement that CASA placed greater emphasis on the practical aspects of flying in terms of flying operations in general aviation.

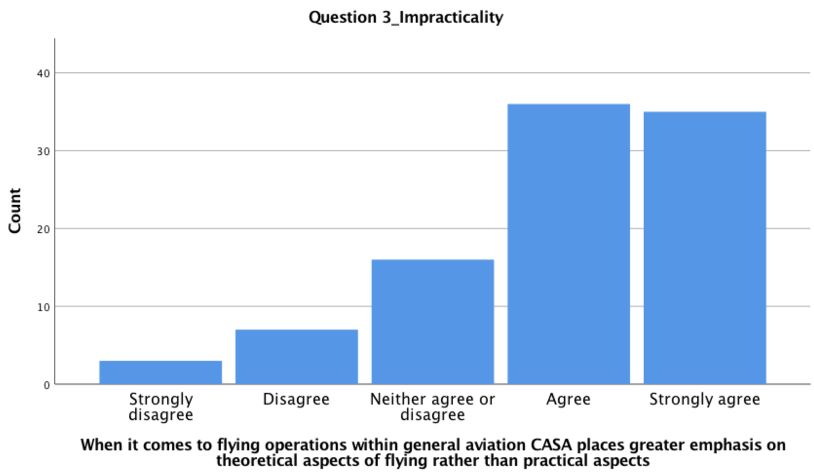

Figure 6.3. Responses to the third question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group did not significantly differ from the post-TEM group in their responses ($U = 746.5$, $z = .35$, $p = .73$, $r = .04$), thereby suggesting that both groups were more in agreement that, instead of practical aspects, CASA placed greater emphasis on the theoretical aspects of flying in terms of flying operations in general aviation.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}359.5$, $z = 1.42$, $p = .16$, $r = .14$), thereby suggesting that both groups shared a similar view that CASA placed greater emphasis on the theoretical aspects of flying in terms of flying operations in general aviation.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses ($U = 1{,}120$, $z = .95$, $p = .34$, $r = .10$), thereby suggesting that both groups shared similar opinions that CASA placed greater emphasis on the theoretical aspects of flying in terms of flying operations in general aviation.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 3 did not vary according to groups ($H(2) = 4.65$, $p = .10$), thereby suggesting that all three groups were more in agreement that CASA placed greater emphasis on the theoretical aspects of flying in terms of flying operations in general aviation.

For the fourth question on impracticality, a total of 86 (88.7%) survey participants disagreed or strongly disagreed that senior executives and decision makers in the CASA had a good understanding of how general aviation operated (Figure 6.4). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that participants were more in disagreement that senior executives and decision makers within CASA had a good understanding of how general aviation operated.

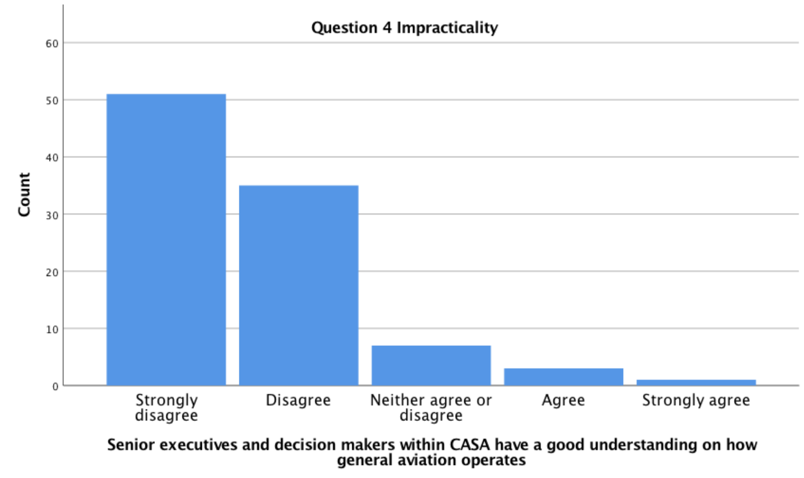

Figure 6.4. Responses to the fourth question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group did not significantly differ from the post-TEM group in their responses (U = 854, z = 1.48, p = .14, r = .15), thereby suggesting that both groups were more in agreement that senior executives and decision makers in CASA had a poor understanding of how general aviation operated.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses (U = 1,172, z = -.01, p = .99, r = .00), thereby suggesting that both groups were more in agreement that senior executives and decision makers in CASA had a poor understanding of how general aviation operated.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses (U = 937, z = -.59, p = .56, r = -.06), thereby suggesting that both groups were more in agreement that senior executives and decision makers in CASA had a poor understanding of how general aviation operated.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 4 did not vary according to groups ($H(2) = 2.13$, $p = .35$), thereby suggesting that all three groups were more in agreement that senior executives and decision makers within CASA had a poor understanding of how general aviation operated.

For the fifth question on impracticality, a total of 65 (67%) survey participants disagreed or strongly disagreed that they found it easy to keep up to date with changes to aviation safety regulations/initiatives (Figure 6.5). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that participants were more in agreement that it was difficult to keep up to date with changes to aviation safety regulations/initiatives.

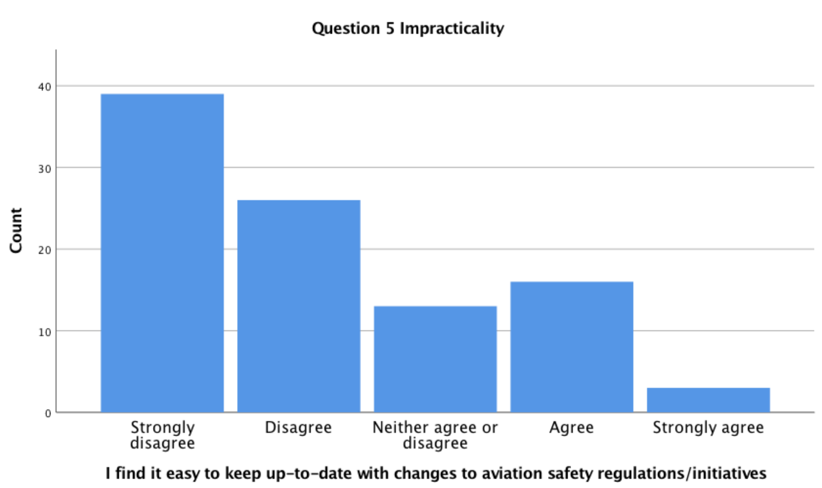

Figure 6.5. Responses to the fifth question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group were significantly different from the post-TEM group in their responses ($U = 963$, $z = 2.45$, $p = .01$, $r = .25$), thereby suggesting that the pre-TEM group had relatively more difficulty in keeping up to date with the changes to aviation safety regulations/initiatives than the post-TEM group. The majority of the pre-TEM group either held or had held CPL ($n = 39$) or ATPL ($n = 32$).

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}123.5$, $z = .38$, $p = .71$, $r = .04$), thereby suggesting that both groups similarly had difficulties keeping up to date with the changes to aviation safety regulations/initiatives.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses ($U = 842.5$, $z = -1.33$, $p = .18$, $r = -.14$), thereby suggesting that both groups similarly had difficulties keeping up to date with the changes to aviation safety regulations/initiatives.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that overall responses for Question 5 did not vary according to groups ($H(2) = 3.36$, $p = .19$), thereby suggesting that all three groups were more in agreement that keeping up to date with the changes to aviation safety regulations/initiatives was difficult.

For the sixth question on impracticality, a total of 74 (76.3%) survey participants disagreed or strongly disagreed that they were satisfied with the way CASA developed aviation safety regulations/initiatives (Figure 6.6). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that relatively more participants were dissatisfied with the way CASA developed aviation safety regulations/initiatives.

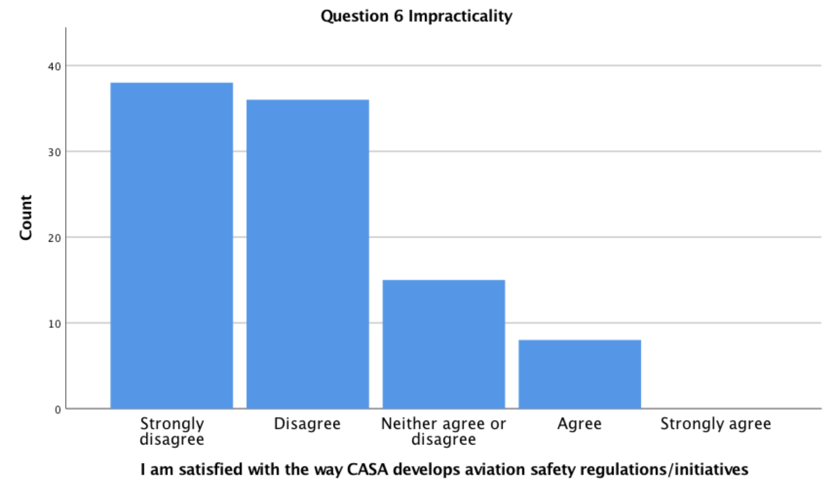

Figure 6.6. Responses to the sixth question to test impracticality.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group were significantly different from the post-TEM group in their responses ($U = 962$, $z = 2.48$, $p = .01$, $r = .25$), suggesting that the pre-TEM group were more dissatisfied with the way CASA developed aviation safety regulations/initiatives than the post-TEM group.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}105$, $z = -.52$, $p = .60$, $r = -.05$), thereby suggesting that both groups were more dissatisfied with the way CASA developed aviation safety regulations/initiatives.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses ($U = 874.5$, $z = -1.08$, $p = .28$, $r = -.11$), thereby suggesting that both groups were more dissatisfied with the way CASA developed aviation safety regulations/initiatives.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 6 did not vary according to groups ($H(2) = 5.28$, $p = .07$), thereby suggesting that all three groups were more dissatisfied with the way CASA developed aviation safety regulations/initiatives.

Statistical analysis of the six questions on impracticality was performed using both descriptive and inferential statistics. The overall results supported the findings from Study 2, where all five participants expressed concerns regarding a lack of practicality in terms of understanding the concept of TEM and implementing TEM in practice, whether in daily routine flying or flight training. The results also supported the first hypothesis, with survey participants more in agreement that TEM lacked practicality. Impracticality would have high correlation with the way the survey participants viewed the level of guidance and support provided by CASA before and after TEM was implemented. Its analysis and results are presented in the next section.

6.3.1.2 Theme 2: Lack of guidance and support.

All participants from Study 2 were in agreement that there was not adequate guidance and support provided by CASA when TEM was implemented in Australian general aviation. Similarly, a finding from Study 1 indicated that participants' response was neutral ($M = 3.44$, $SD = 1.25$) when asked to rate whether CASA produced adequate training and guidance materials on TEM training for general aviation. To test whether these findings were supported by the survey participants in this study, both descriptive and inferential statistics were generated for each survey question, and the results are presented below.

For the first question on the lack of guidance and support, a total of 62 (63.9%) survey participants disagreed or strongly disagreed that CASA provided adequate training and guidance material on TEM training for general aviation (Figure 6.7). The result from the one-sample Wilcoxon signed rank test was statistically significant ($p < .001$), thereby suggesting that a large number of participants disagreed that CASA provided adequate training and guidance material on TEM training for general aviation.

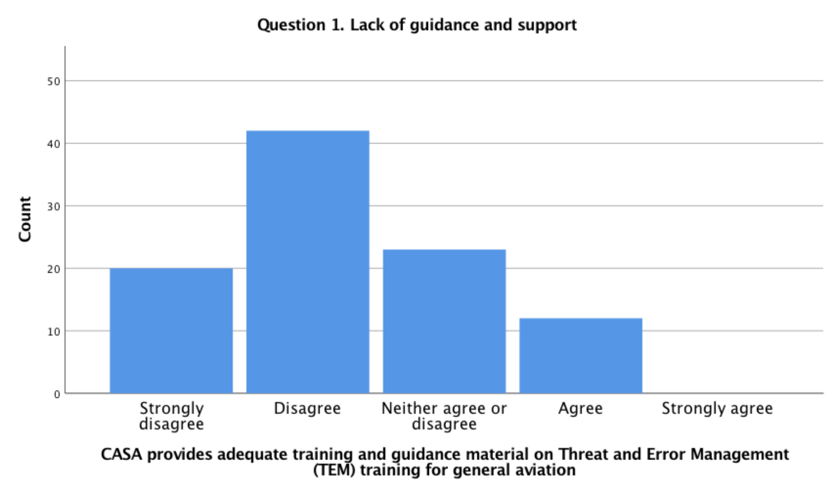

Figure 6.7. Responses to the first question to test lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group did not significantly differ from the post-TEM group in their responses (U = 863, z = 1.49, p = .14, r = .15), thereby suggesting that both groups were more in agreement that CASA did not provide adequate training and guidance material on TEM training for general aviation.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses (U = 1,114, z = -.45, p = .65, r = -.05), thereby suggesting that both groups shared a similar view that CASA did not provide adequate training and guidance material on TEM training for general aviation.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses (U = 944, z = -.50, p = .62, r = -.05), thereby suggesting that both groups had similar opinions that CASA did not provide adequate training and guidance material on TEM training for general aviation.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that overall responses for Question 1 did not vary according to groups ($H(2) = .52$, $p = .77$), thereby suggesting that all three groups shared a similar view that CASA did not provide adequate training and guidance material on TEM training for general aviation.

For the second question on the lack of guidance and support, a total of 51 (52.6%) survey participants disagreed or strongly disagreed that the concept of TEM was easy to understand (Figure 6.8). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that a large number of participants disagreed that the concept of TEM was easy to understand, and this would have been contributed to by the lack of guidance and support from CASA.

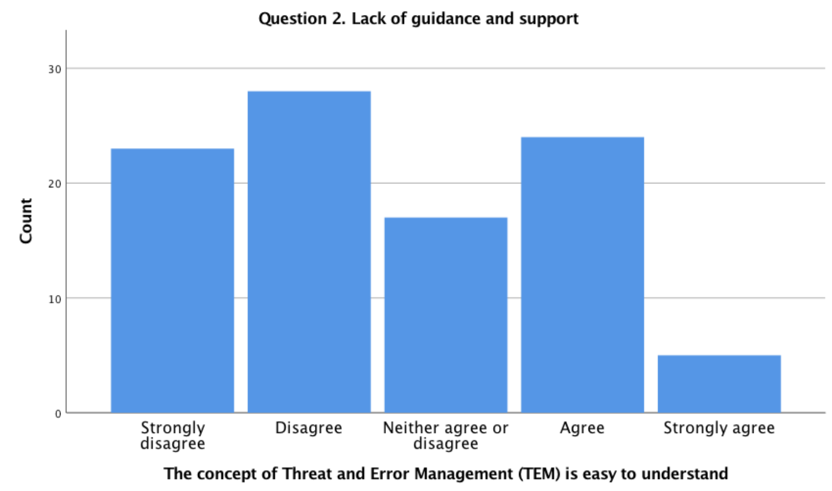

Figure 6.8. Responses to the second question to test lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group did not significantly differ from the post-TEM group in their responses ($U = 729$, $z = .17$, $p = .86$, $r = .02$), thereby suggesting that both groups shared a similar view that the concept of TEM was not easy to understand.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1,061$, $z = -.83$, $p = .40$, $r = -.08$), thereby suggesting that both groups had similar opinions that the concept of TEM was difficult to understand.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses ($U = 1,022.5$, $z = .14$, $p = .88$, $r = .01$), thereby suggesting that both groups were more in agreement that the concept of TEM was not easy to understand.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that overall responses for Question 2 did not vary according to groups ($H(2) = 2.16$, $p = .34$), thereby suggesting that all three groups shared a similar view that the concept of TEM was difficult to understand.

For the third question on the lack of guidance and support, a total of 71 (73.2%) survey participants disagreed or strongly disagreed that CASA explained TEM and the ways it affected participants' role and/or activities in a clear and succinct manner (Figure 6.9). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that a large number of participants disagreed that CASA explained TEM and the way it affected participants' role and/or activities in a clear and succinct manner.

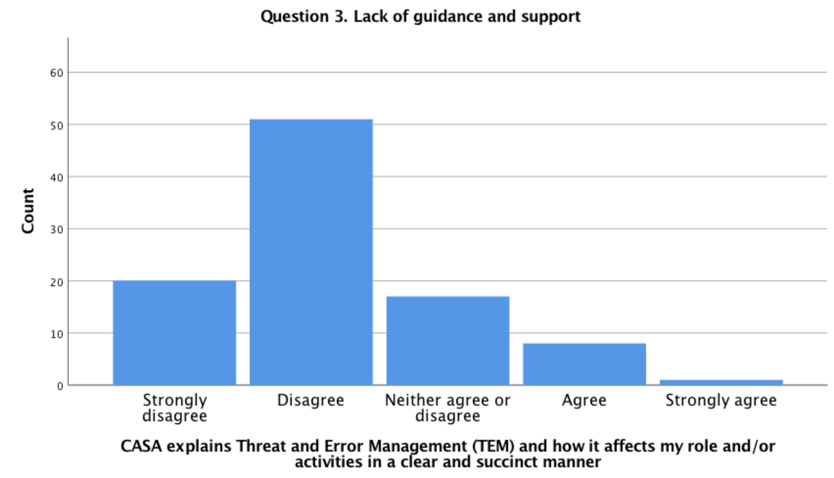

Figure 6.9. Responses to the third question to test lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses (U = 905, z = .1.96, p = .05, r = .20), thereby suggesting that both groups shared a similar view regarding whether CASA explained TEM and the way it affected their role and/or activities in a clear and succinct manner.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses (U = 1,108, z = -.51, p = .61, r = -.05), thereby suggesting that both groups were more in agreement that CASA did not explain TEM and the way it affected their role and/or activities in a clear and succinct manner.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group did not significantly differ from the non-FE group in their responses (U = 1,088.5, z = .71, p = .48, r = .07), thereby suggesting that both groups had similar opinions that CASA did not explain TEM and the way it affected their role and/or activities in a clear and succinct manner.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 3 did not vary according to groups ($H(2) = 1.13$, $p = .57$), thereby suggesting that all three groups shared a similar view that CASA did not explain TEM and the way it affected their role and/or activities in a clear and succinct manner.

For the fourth question on lack of guidance and support, less than half of the survey participants (40.2%) did not regularly attend CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives (Figure 6.10). The result from the one-sample Wilcoxon signed rank test was not statistically significant ($p = .535$), thereby suggesting mixed views among the survey participants on regular attendance at CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives.

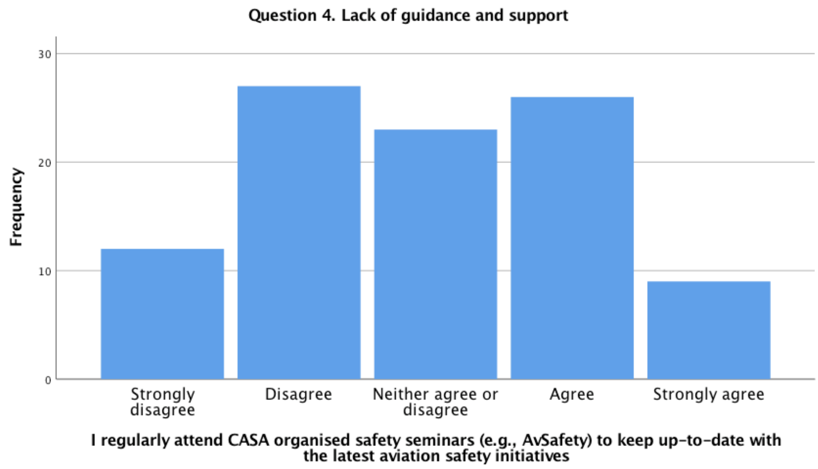

Figure 6.10. Responses to the fourth question to test lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 669.5$, $z = -.40$, $p = .69$, $r = -.04$), thereby suggesting that both groups shared mixed views on attendance at CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses (U = 971, z = -.1.5, p = .13, r = -.15), thereby suggesting that both groups had mixed opinions on attendance at CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result showed that the FE group was significantly different from the non-FE group in their responses (U = 667.5, z = -2.71, p = .007, r = -.28), thereby suggesting that a statistically significant number of participants from the non-FE group did not regularly attend CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 4 did not vary according to groups ($H(2)$ = 1.55, p = .46), thereby suggesting that all three groups shared mixed views on attendance at CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives.

For the fifth question on the lack of guidance and support, less than half of the survey participants (23.7%) did not consider it useful to attend seminars, such as AvSafety and/or others, organised by CASA, while 32% of participants answered the question with neutral responses (Figure 6.11). The result from the one-sample Wilcoxon signed rank test was again not statistically significant (p = .098), showing mixed views on the usefulness of CASA-organised seminars, such as AvSafety.

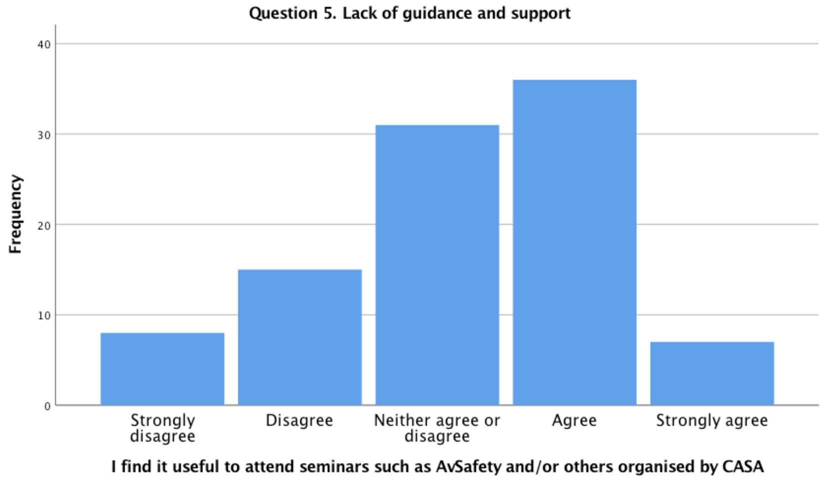

Figure 6.11. Responses to the fifth question to test the lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 774.5$, $z = .62$, $p = .54$, $r = .06$), thereby suggesting that both groups similarly shared mixed views on the usefulness of attendance at CASA-organised seminars, such as AvSafety.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}118.5$, $z = -.41$, $p = .68$, $r = .06$), thereby suggesting that both groups again shared varied views on the usefulness of attendance at CASA-organised seminars, such as AvSafety.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 1{,}022$, $z = .14$, $p = .89$, $r = .01$), thereby suggesting that both groups again similarly shared mixed views on the usefulness of attendance at CASA-organised seminars, such as AvSafety.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 5 did not vary according to groups ($H(2) = 3.4$, $p = .14$), thereby suggesting that all three groups similarly shared mixed views on the usefulness of attendance at CASA-organised seminars, such as AvSafety.

For the sixth question on lack of guidance and support, less than half of the survey participants (32%) did not consider the number of safety-related publications and resources from CASA useful, and a similar number of participants were neutral about this question, while a slightly greater number of participants indicated agreement (Figure 6.12). The result from the one-sample Wilcoxon signed rank test was again not statistically significant (p = .523), showing mixed views on the usefulness of safety-related publications and resources from CASA.

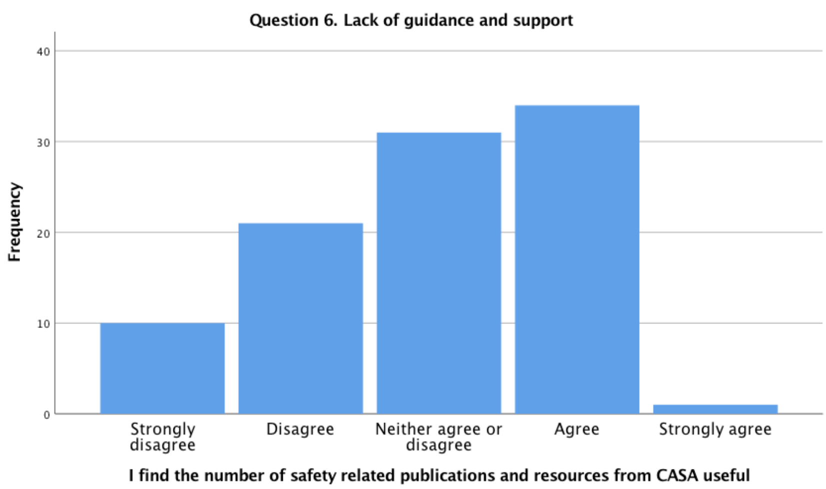

Figure 6.12. Responses to the sixth question to test the lack of guidance and support.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 832.5$, $z = 1.18$, $p = .24$, $r = .12$), thereby suggesting that both groups similarly shared mixed views on the usefulness of safety-related publications and resources from CASA.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}127.5$, $z = -.34$, $p = .73$, $r = -.03$), thereby suggesting that both groups had mixed opinions on the usefulness of safety-related publications and resources from CASA.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 933$, $z = -.59$, $p = .56$, $r = -.06$), thereby suggesting that both groups similarly shared mixed views on the usefulness of safety-related publications and resources from CASA.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 6 did not vary according to groups ($H(2) = 1.09$, $p = .58$), thereby suggesting that all three groups had mixed opinions on the usefulness of safety-related publications and resources from the CASA.

Analyses of six questions on the lack of guidance and support were undertaken using both descriptive and inferential statistics. The overall results partially supported the findings from Study 2, where all five participants were more in agreement that there was not adequate guidance provided by CASA when TEM was implemented in Australian general aviation. While the results from the first three questions supported the hypothesis, the other three questions did not support the hypothesis. Therefore, the results partially supported the second hypothesis, with survey participants more in agreement that there was a lack of guidance and support provided by CASA.

Impracticality and the lack of support and guidance from CASA would have high correlation with the way the survey participants viewed TEM implementation. The TEM implementation analyses and results are presented in the next section.

6.3.1.3 *Theme 3: Threat and error management implementation.*

All participants from Study 2 shared similar views that TEM implementation was poor. To verify whether this was supported by the survey participants in this study, both descriptive and inferential statistics were generated for each survey question, and the results are presented below.

For the first question on TEM implementation, a total of 79 (81.4%) survey participants disagreed or strongly disagreed that CASA consulted with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM (Figure 6.13). The result from the one-sample Wilcoxon signed rank test was statistically significant ($p < .001$), suggesting that the majority of survey participants did not agree that CASA consulted with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

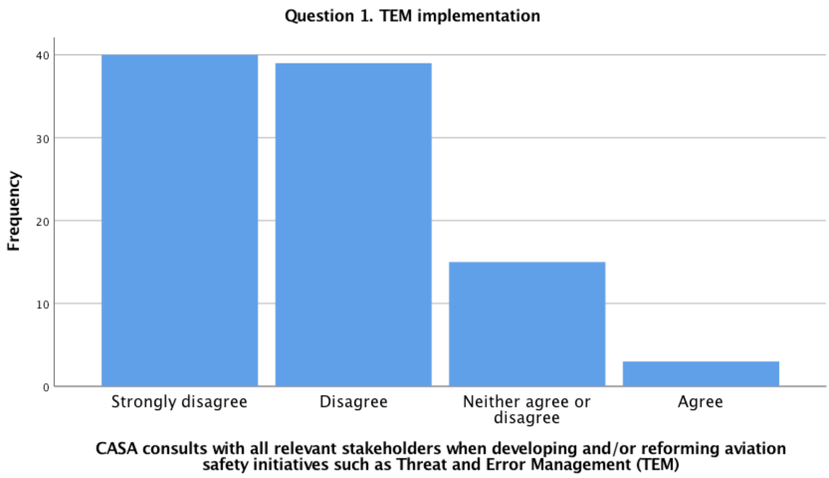

Figure 6.13. Responses to the first question on TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was significantly different from the post-TEM group in their responses (U = 1,001, z = 2.9, p = .004, r = .29), thereby suggesting that relatively more participants in the pre-TEM group disagreed that CASA consulted with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group did not significantly differ from the non-FI group in their responses (U = 1,104.5, z = -.53, p = .59, r = -.05), thereby suggesting that both groups were more in agreement that CASA did not consult with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses (U = 974.5, z = -.26, p = .80, r = -.03), thereby suggesting that both groups shared a similar view that CASA did not consult with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 1 did not vary according to groups ($H(2)$ = 2.14, p = .34), thereby suggesting that all three groups had similar opinions that CASA did not consult with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

For the second question on TEM implementation, a total of 49 (50.5%) survey participants disagreed or strongly disagreed that CASA Aviation Safety Advisors (ASAs) had a consistent understanding of TEM principles (Figure 6.14). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), suggesting that only a small number of survey participants agreed that CASA ASAs had a consistent understanding of TEM principles.

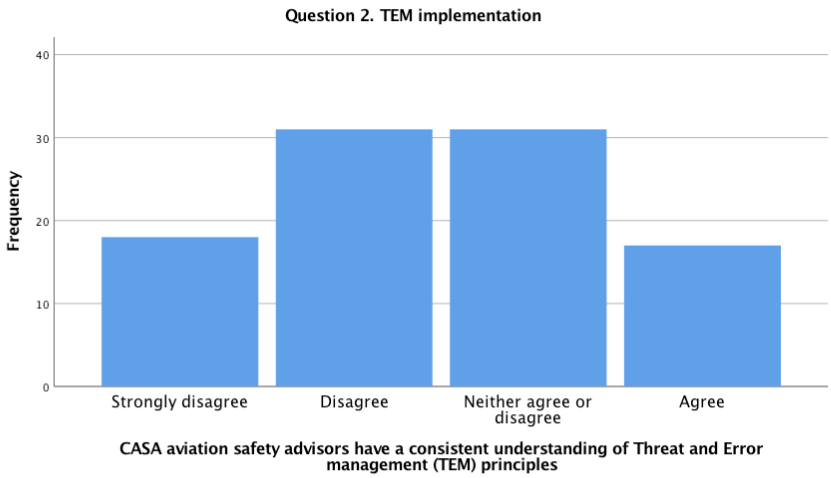

Figure 6.14. Responses to the second question on TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 829$, $z = 1.14$, $p = .25$, $r = .12$), thereby suggesting that both groups shared a similar view on the inconsistent understanding of TEM principles among CASA ASAs.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 995$, $z = -1.34$, $p = .18$, $r = -.14$), thereby suggesting that both groups were more in agreement that CASA ASAs did not have a consistent understanding of TEM principles.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 864.4$, $z = -1.14$, $p = .25$, $r = -.12$), thereby suggesting that both groups shared similar opinions on a consistent understanding of TEM principles among CASA ASAs.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 2 did not vary according to groups ($H(2) = 2.64$, $p = .27$), thereby suggesting that all three groups were more in agreement that CASA ASAs did not have a consistent understanding of TEM principles.

For the third question on TEM implementation, a total of 78 (80.4%) survey participants were not satisfied with the way CASA implemented new aviation safety regulations/initiatives (Figure 6.15). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby supporting most survey participants' dissatisfaction with the way CASA implemented new aviation safety regulations/initiatives.

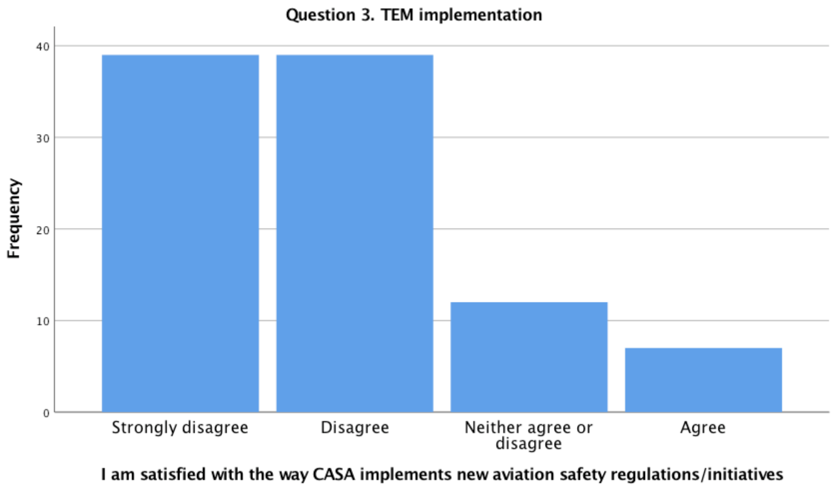

Figure 6.15. Responses to the third question on TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 893.5$, $z = 1.82$, $p = .07$, $r = .18$), thereby suggesting that both groups similarly shared dissatisfaction with the way CASA implemented new aviation safety regulations/initiatives.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1,098$, $z = -.58$, $p = .56$, $r = -.06$), thereby suggesting that both groups were similarly dissatisfied with the way CASA implemented new aviation safety regulations/initiatives.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 928.5$, $z = -.64$, $p = .52$, $r = -.06$), thereby suggesting that both groups shared similar dissatisfaction with the way CASA implemented new aviation safety regulations/initiatives.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 3 did not vary according to groups ($H(2) = 4.01$, $p = .14$), thereby suggesting that all three groups were more in agreement that the CASA's approach to implementing new aviation safety regulations/initiatives was not satisfactory.

For the fourth question on TEM implementation, only seven (7.2%) survey participants agreed that the number of CASA staff with appropriate expertise in TEM was adequate (Figure 6.16). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby supporting a large number of survey participants' responses regarding the inadequate number of CASA staff with appropriate expertise in TEM.

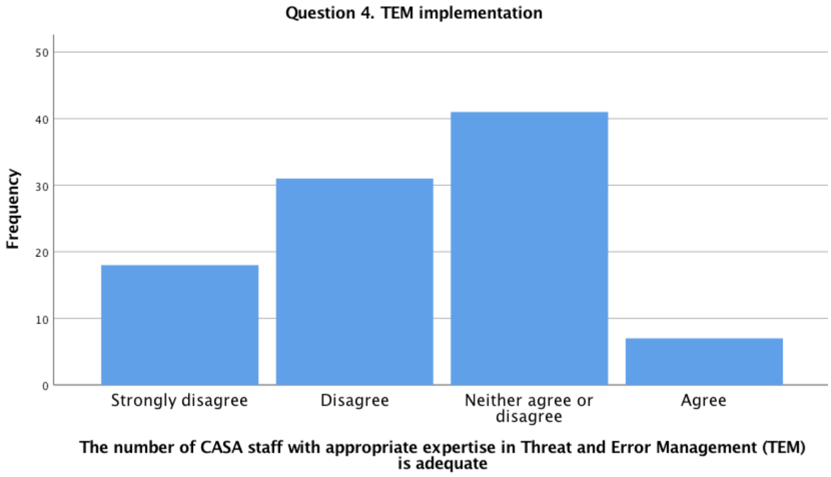

Figure 6.16. Responses to the fourth question on poor TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 799.5$, $z = .87$, $p = .38$, $r = .09$), thereby suggesting that both groups similarly considered that the number of CASA staff with appropriate expertise in TEM was inadequate.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}255$, $z = .63$, $p = .53$, $r = .06$), thereby suggesting that both groups were more in agreement that there were inadequate numbers of CASA staff with appropriate expertise in TEM.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 1{,}182$, $z = 1.47$, $p = .14$, $r = .15$), thereby suggesting that both groups shared a similar view that the number of CASA staff with appropriate expertise in TEM was inadequate.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 4 did not vary according to groups ($H(2) = 4.29$, $p = .12$), thereby suggesting that all three groups similarly considered that greater numbers of CASA staff with appropriate expertise in TEM were required.

For the fifth question on TEM implementation, only eight (8.2%) survey participants agreed that CASA maintained an effective ongoing dialogue with the general aviation industry (Figure 6.17). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), thereby suggesting that the majority of survey participants disagreed that CASA maintained an effective ongoing dialogue with the general aviation industry.

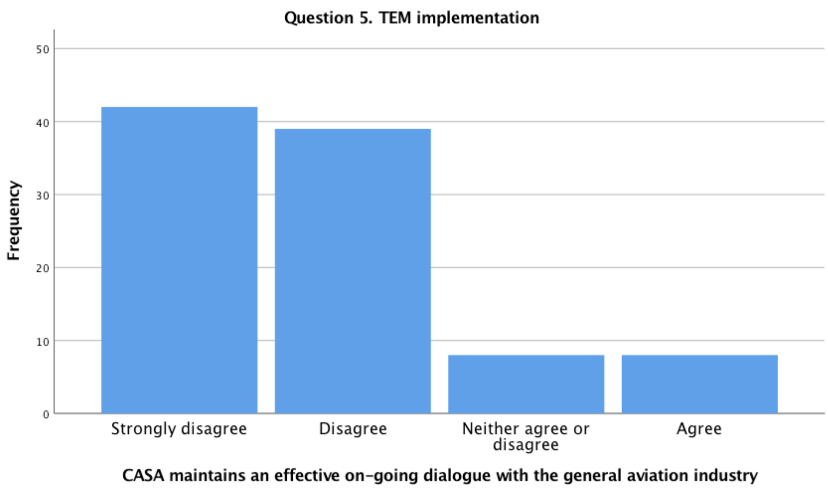

Figure 6.17. Responses to the fifth question on TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 835.5$, $z = 1.25$, $p = .21$, $r = .13$), thereby suggesting that both groups similarly showed dissatisfaction with CASA not maintaining an effective ongoing dialogue with the general aviation industry.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 979.5$, $z = -1.51$, $p = .13$, $r = -.15$), thereby suggesting that both groups were more in agreement that CASA did not maintain an effective ongoing dialogue with the general aviation industry.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 872$, $z = -1.12$, $p = .26$, $r = -.11$), thereby suggesting that both groups shared a similar view that CASA did not maintain an effective ongoing dialogue with the general aviation industry.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 5 did not vary according to groups ($H(2) = .78$, $p = .68$), thereby suggesting that all three groups had similar opinions that CASA did not maintain an effective ongoing dialogue with the general aviation industry.

For the sixth question on TEM implementation, only five (5.2%) survey participants agreed that CASA clearly and succinctly explained new safety regulations/initiatives and the way they would affect the general aviation industry (Figure 6.18). The result from the one-sample Wilcoxon signed rank test was again statistically significant ($p < .001$), suggesting that CASA did not clearly and succinctly explain new safety regulations/initiatives and their effects on the general aviation industry.

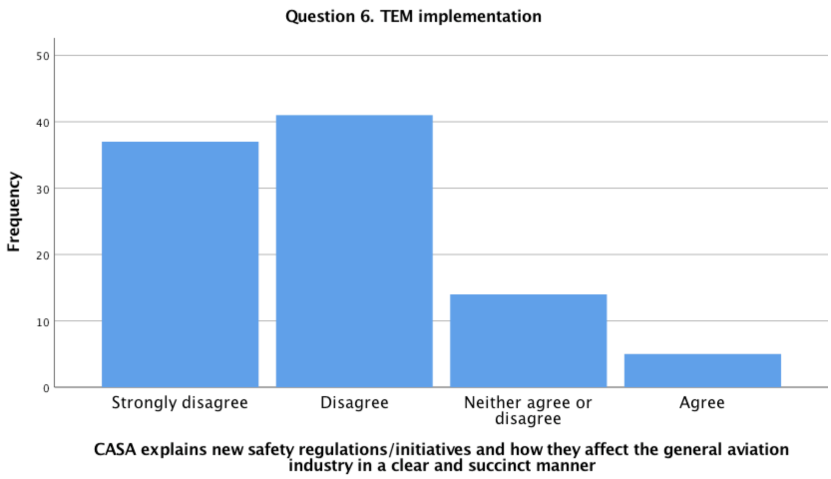

Figure 6.18. Responses to the sixth question on TEM implementation.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 815.5$, $z = 1.04$, $p = .30$, $r = .11$), thereby suggesting that both groups similarly disagreed that CASA clearly and succinctly explained new safety regulations/initiatives and the way they would affect the general aviation industry.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result again showed that the FI group did not significantly differ from the non-FI group in their responses ($U = 1{,}123$, $z = -.39$, $p = .70$, $r = -.04$), thereby suggesting that both groups shared a similar view that CASA did not clearly and succinctly explain new safety regulations/initiatives and their effects on the general aviation industry.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 966.5$, $z = -.32$, $p = .75$, $r = -.03$), thereby suggesting that both groups similarly disagreed that CASA clearly and succinctly explained new safety regulations/initiatives and how they affected the general aviation industry.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 6 did not vary according to groups ($H(2) = 1.39$, $p = .50$), thereby suggesting that all three groups had similar opinions regarding whether CASA clearly and succinctly explained new safety regulations/initiatives and their effects on the general aviation industry.

Analyses of the six questions on poor implementation were presented using both descriptive and inferential statistics. The overall results supported the findings from Study 2, where all five participants shared a similar view that the implementation of TEM was poor. The result supported the third hypothesis, with survey participants more in agreement that TEM implementation in Australian general aviation was poor.

Analyses of the survey questions thus far have suggested that the approach to TEM was impractical and there was a lack of guidance and support from CASA. These findings, together with the other factors identified, led the survey participants to hold negative opinions on TEM implementation, which would likely have negatively affected the way TEM was taught and practised. Its analyses and results are presented in the next section.

6.3.1.4 Theme 4: Threat and error management in practice.

All participants from Study 2 shared similar views that poor TEM implementation had negatively affected the way TEM was taught and practised in Australian general aviation. To verify this, both descriptive and inferential statistics were generated for each survey question, and the results are presented below.

For the first question on TEM in practice, a total of 57 (58.7%) survey participants agreed or strongly agreed that they used TEM principles when they were undertaking flying activities (Figure 6.19). The result from the one-sample Wilcoxon signed rank test was statistically significant ($p < .001$), suggesting that the majority of survey participants used TEM principles when undertaking flying activities. This is an encouraging finding, despite the consensus among the survey participants about TEM being impractical, lacking guidance and support from CASA, and being poorly implemented.

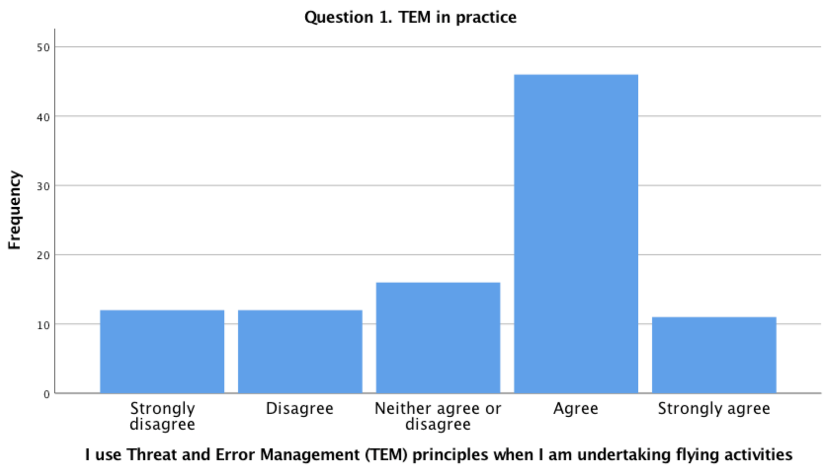

Figure 6.19. Responses to the first question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 744$, $z = .33$, $p = .75$, $r = .03$), thereby suggesting that both groups were more in agreement that they used TEM principles when undertaking flying activities.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the non-FI group was significantly different from the FI group in their responses ($U = 884$, $z = -2.22$, $p = .03$, $r = -.23$), thereby suggesting that relatively more survey participants from the non-FI group did not use TEM principles when undertaking flying activities.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the non-FE group was significantly different from the FE group in their responses ($U = 755$, $z = -2.08$, $p = .04$, $r = -.21$), thereby suggesting that relatively more survey participants from the non-FE group did not use TEM principles when undertaking flying activities.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 1 did not vary according to groups ($H(2) = 1.09$, $p = .58$), thereby suggesting that all three groups were more in agreement that they used TEM principles when undertaking flying activities.

For the second question on TEM in practice, a total of 47 (48.4%) survey participants agreed or strongly agreed that they had a clear understanding of what they needed to do when applying TEM principles (Figure 6.20). The result from the one-sample Wilcoxon signed rank test was not statistically significant ($p = .249$), thereby suggesting that there were mixed views on whether the survey participants had a clear understanding of what was needed when applying TEM principles.

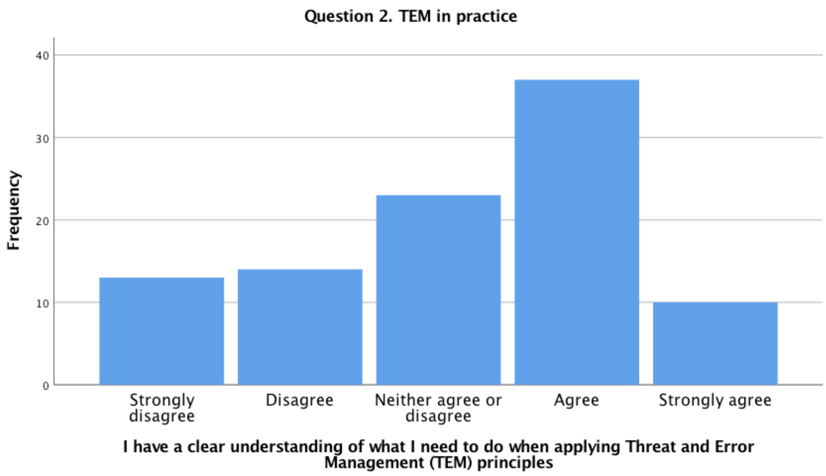

Figure 6.20. Responses to the second question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 688$, $z = -.22$, $p = .82$, $r = -.02$), thereby suggesting that both groups similarly had a clear understanding of what was required when applying TEM principles.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the non-FI group was significantly different from the FI group in their responses ($U = 748.5$, $z = -3.19$, $p = .001$, $r = -.32$), thereby suggesting that the non-FI group was less clear about what was required when applying TEM principles.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the non-FE group was significantly different from the FE group in their responses ($U = 712$, $z = -2.38$, $p = .02$, $r = -.24$), thereby suggesting that the non-FE group was less clear about what was required when applying TEM principles.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 2 did not vary according to groups ($H(2) = 2.87$, $p = .24$), thereby suggesting that all three groups had a similar level of understanding about what was required when applying TEM principles.

For the third question on TEM in practice, a total of 58 (59.8%) survey participants agreed or strongly agreed that consideration of TEM was an important part of their flight preparation (Figure 6.21). The result from the one-sample Wilcoxon signed rank test was statistically significant ($p = .002$), suggesting that the majority of survey participants considered TEM an important part of their flight preparation.

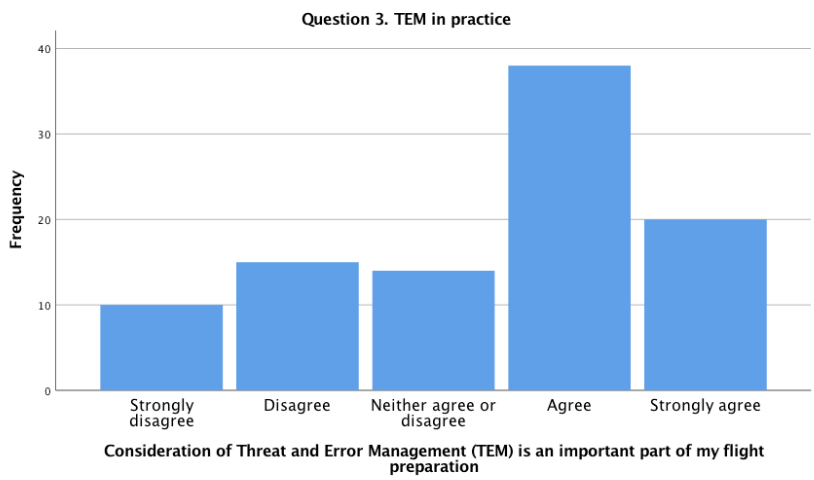

Figure 6.21. Responses to the third question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 771.5$, $z = .58$, $p = .56$, $r = .06$), thereby suggesting that both groups were more in agreement that consideration of TEM was an important part of their flight preparation.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the non-FI group was significantly different from the FI group in their responses ($U = 862$, $z = -2.34$, $p = .002$, $r = -.24$), thereby suggesting that consideration of TEM was a less important part of the non-FI group's flight preparation.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result showed that the FE group was not significantly different from the non-FE group in their responses ($U = 815.5$, $z = -1.54$, $p = .12$, $r = -.16$), thereby suggesting that both groups shared a similar view that consideration of TEM was an important part of their flight preparation.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 3 did not vary according to groups ($H(2) = 3.15$, $p = .21$), thereby suggesting that all three groups were more in agreement that an important part of flight preparation was consideration of TEM.

For the fourth question on TEM in practice, a total of 79 (81.5%) survey participants agreed or strongly agreed that they always performed a series of briefings (e.g., emergency and take-off briefings), taking into account the prevailing conditions and situations (Figure 6.22). The result from the one-sample Wilcoxon signed rank test supported this ($p < .001$).

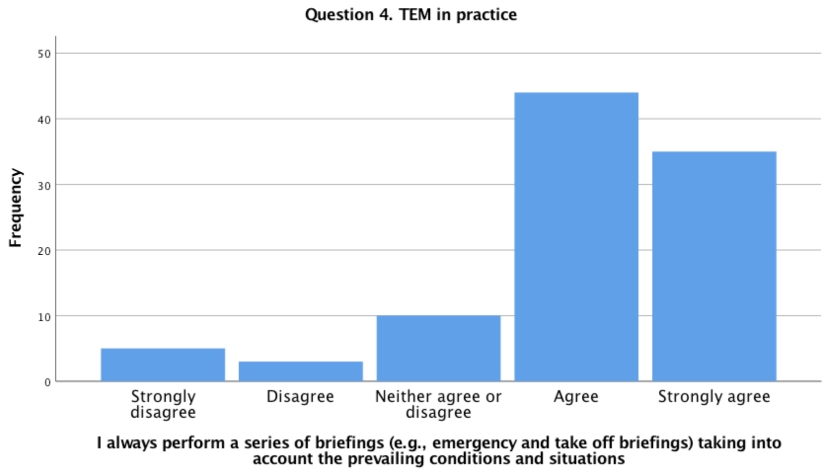

Figure 6.22. Responses to the fourth question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses (U = 601.5, z = -1.10, p = .27, r = -.11), thereby suggesting that both groups similarly considered the prevailing conditions and situations prior to performing a series of briefings (e.g., emergency and take-off briefings).

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the non-FI group was significantly different from the FI group in their responses (U = 811.5, z = -2.82, p = .005, r = -.29), thereby suggesting that the survey participants from the non-FI group did not always perform a series of briefings (e.g., emergency and take-off briefings) considering the prevailing conditions and situations.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the non-FE group was significantly different from the FE group in their responses (U = 653, z = -2.97, p = .003, r = -.30), thereby suggesting that the non-FE group did not always consider the prevailing conditions and situations prior to performing a series of briefings (e.g., emergency and take-off briefings).

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 4 varied according to groups ($H(2) = 10.14$, $p = .006$). A post hoc pairwise comparisons with adjusted p-values showed that there were no significant differences between the PPL and CPL holders ($p = 1.00$, $r = .03$) and between the PPL and ATPL holders ($p = .23$, $r = .18$). However, there was a significant difference in the overall responses between the CPL and ATPL holders ($p = .005$, $r = -.32$), thereby suggesting that a relatively greater number of CPL holders did not always consider the prevailing conditions and situations prior to performing a series of briefings (e.g., emergency and take-off briefings).

For the fifth question on TEM in practice, less than half of the survey participants (30.9%) strongly agreed or agreed that TEM briefing was an unnecessary additional task (Figure 6.23). The result from the one-sample Wilcoxon signed rank test was not statistically significant ($p = .181$), suggesting mixed views on this question among the survey participants.

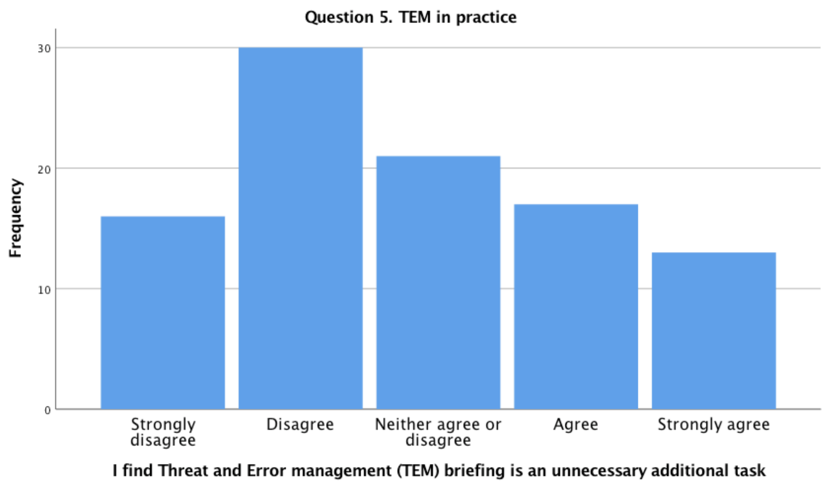

Figure 6.23. Responses to the fifth question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 660$, $z = -.49$, $p = .63$, $r = -.05$), thereby suggesting that both groups similarly considered TEM briefing an unnecessary additional task.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group was not significantly different from the non-FI group in their responses ($U = 1{,}396.5$, $z = 1.66$, $p = .10$, $r = -.05$), thereby suggesting that both groups did not consider TEM briefing an unnecessary additional task.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result again showed that the FE group was not significantly different from the non-FE group in their responses ($U = 1{,}217$, $z = 1.70$, $p = .09$, $r = .17$), thereby suggesting that both groups did not consider TEM briefing an unnecessary additional task.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 5 did not vary according to groups ($H(2) = 2.31$, $p = .31$), thereby suggesting that all three groups similarly agreed that TEM briefing was not considered an unnecessary additional task.

For the sixth question on TEM in practice, a total of 61 (62.9%) survey participants saw value in TEM (Figure 6.24) when undertaking flying activities. The result from the one-sample Wilcoxon signed rank test supported this ($p < .001$).

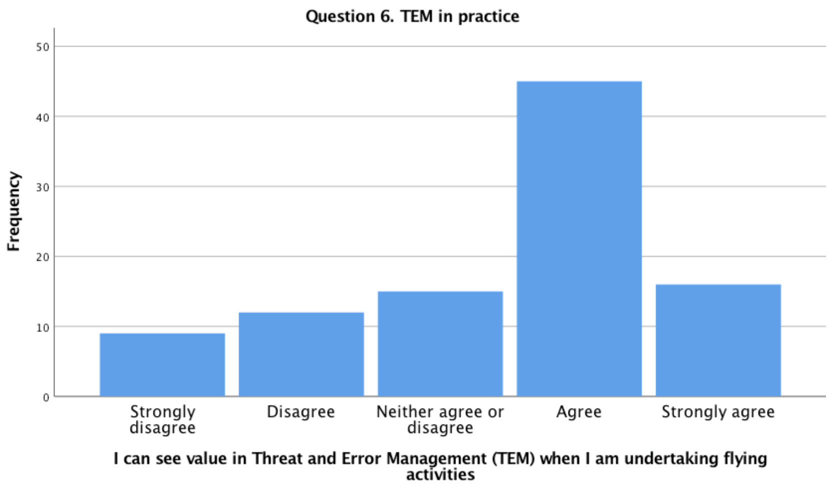

Figure 6.24. Responses to the sixth question on TEM in practice.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 852.5$, $z = 1.39$, $p = .16$, $r = .14$), thereby suggesting that both groups similarly saw value in TEM when undertaking flying activities.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group was not significantly different from the non-FI group in their responses ($U = 945$, $z = -1.75$, $p = .08$, $r = -.18$), thereby suggesting that both groups shared a similar view that TEM was valuable when undertaking flying activities.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result showed that the non-FE group was significantly different from the FE group in their responses ($U = 745.5$, $z = -2.15$, $p = .03$, $r = -.22$), thereby suggesting that the non-FE group saw less value in TEM when undertaking flying activities.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 6 did not vary according to groups ($H(2) = 2.22$, $p = .33$), thereby suggesting that all three groups shared similar views on the value of TEM when undertaking flying activities.

Analyses of six questions on TEM in practice were presented using both descriptive and inferential statistics. The overall results partially supported the findings from Study 2, where all five participants were more in agreement that there was not adequate guidance provided by CASA when TEM was implemented in Australian general aviation. The result also partially supported the second hypothesis, with survey participants more in agreement that there was a lack of guidance and support provided by CASA. While the results from the first three questions supported the hypothesis, the other three questions did not support the hypothesis.

A total of 24 questions on the four main themes were analysed. In addition to these questions, there were further questions included in the survey. The first two additional questions were for all the survey participants, while the remaining questions were group specific (three questions for current or former FEs, and four questions for current or former FIs). The following two subsections present the results of these questions.

6.3.1.5 *Additional questions.*

The first additional question examined whether the benefits of TEM training were overrated, and was included because, when this particular question was asked in Study 1, the result showed the highest Kruskal-Wallis value by licence type, and mixed views on the perceived benefit of TEM training (Lee et al., 2016). The responses from the survey participants again indicated mixed views on the benefits of TEM training (Figure 6.25), and the result from the one-sample Wilcoxon signed rank test supported this (p = .702).

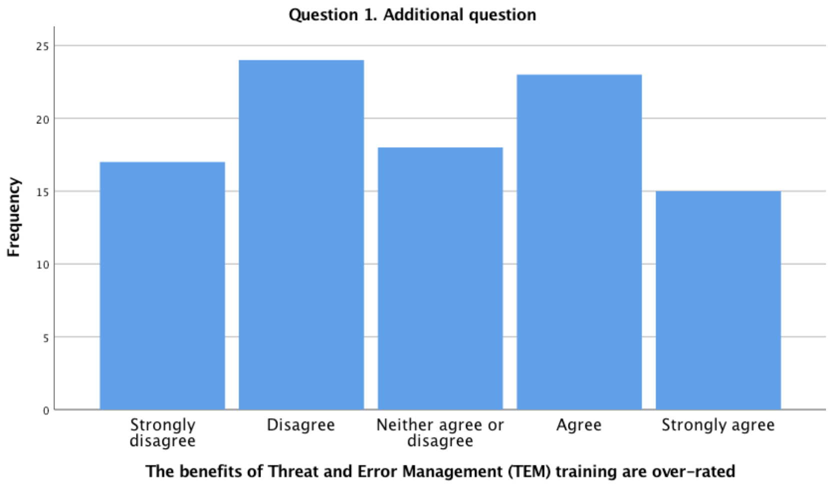

Figure 6.25. First additional question on the benefits of TEM.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses (U = 588, z = -1.17, p = .24, r = -.12), thereby suggesting that both groups shared similar views on the benefits of TEM.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the non-FI group was significantly different from the FI group in their responses ($U = 1{,}453$, $z = 2.07$, $p = .04$, $r = 21$), thereby suggesting that the non-FI group was more in agreement that the benefits of TEM were overrated.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result showed that the FE group was not significantly different from the non-FE group in their responses ($U = 1{,}139$, $z = 1.07$, $p = .29$, $r = .11$), thereby suggesting that both groups shared similar views on the benefits of TEM.

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for this question did not vary according to groups ($H(2) = .64$, $p = .73$), thereby suggesting that all three groups shared similar views on the benefits of TEM.

The second additional question considered whether the use of TEM training/principles should be expanded to other sectors, such as recreational aviation and RPAS sectors. This question was included because there were mixed comments from the participants in Study 2, as follows:

> *'Well, my immediate reaction to that would be: why they would be any different? I mean, they are operating in the same environment, uh, essentially general aviation, the threats are by and large the same, errors that people are going to make and management of those ... same principles are applied.' (Participant 1)*

> *'I would ... my reaction would be to say, yes, it would be helpful everywhere, but I would say the caution to that is, if it's done the way it's been done under my observation in aviation, general aviation, it would just be a distraction.' (Participant 3)*

Almost half of the survey participants (49.5%) strongly agreed or agreed that the use of TEM training/principles should be expanded to other sectors, such as recreational aviation and RPAS sectors (Figure 6.26). The result from the one-sample Wilcoxon signed rank test indicated mixed views as to whether TEM training/principles should be expanded to other sectors, such as recreational aviation and RPAS sectors ($p = .110$).

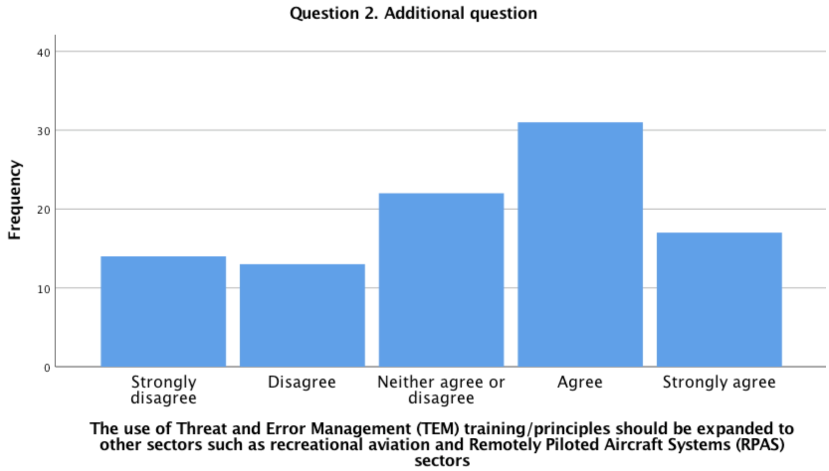

Figure 6.26. Second additional question on the expansion of TEM to other sectors.

The Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the pre-TEM and post-TEM groups. The result showed that the pre-TEM group was not significantly different from the post-TEM group in their responses ($U = 833$, $z = 1.17$, $p = .24$, $r = .12$), thereby suggesting that both groups shared similar views on the expansion of TEM training/principles to other general aviation sectors, such as recreational aviation and RPAS.

Another Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FI and non-FI groups. The result showed that the FI group was not significantly different from the non-FI group in their responses ($U = 914$, $z = -1.92$, $p = .06$, $r = -.19$), thereby suggesting that both groups shared similar views on the expansion of TEM training/principles to other general aviation sectors.

Similarly, the Mann-Whitney test was performed to determine whether there was a statistical difference in responses between the FE and non-FE groups. The result showed that the non-FE group was significantly different from the FE group in their responses (U = 748, z = -2.06, p = .04, r = -.21), thereby suggesting that relatively fewer survey participants in the non-FE group were more in agreement about expanding TEM training/principles to other general aviation sectors (e.g., recreation and RPAS).

The Kruskal-Wallis test was performed to determine whether there was a statistical difference in responses among the PPL, CPL and ATPL groups. The result from the test suggested that the overall responses for Question 2 did not vary according to groups (H(2) = .2.22, p = .33), thereby suggesting that all three groups shared similar views on the expansion of TEM training/principles to other general aviation sectors.

6.3.1.6 Specific questions for flight examiners.

There were three specific questions, based on comments from Study 2, for survey participants who held or had held FE ratings to determine their experience with CASA's support and guidance for FEs prior to TEM implementation, and the level of standardisation on TEM assessment after TEM implementation among FEs during flight tests. In addition, a question was included to gauge whether a series of briefings were completed without involving conscientious thought, which was one of main criticisms from participants in Study 2.

The first question asked whether 'CASA provided adequate guidance to examiners before TEM became a mandatory flight test item'. The majority (83.4%) of respondents disagreed or strongly disagreed that survey participants received adequate guidance from CASA before TEM became a mandatory flight test item (Figure 6.27). The result from the one-sample Wilcoxon signed rank test supported this strong disagreement (p < .001).

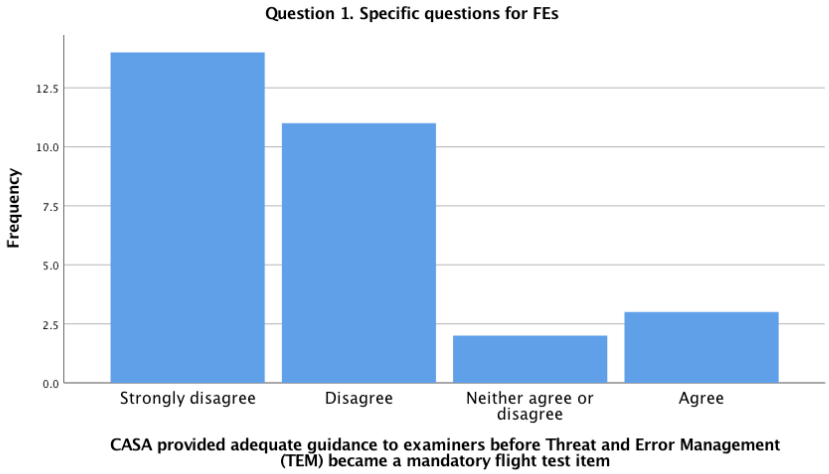

Figure 6.27. First specific question for FEs.

The second question asked whether 'during flight tests, assessment of TEM is standardised among FEs'. More than half (56.7%) of the FEs disagreed or strongly disagreed that there was standardisation among FEs with regard to TEM assessment (Figure 6.28). The result from the one-sample Wilcoxon signed rank test again supported this strong disagreement (p = .002).

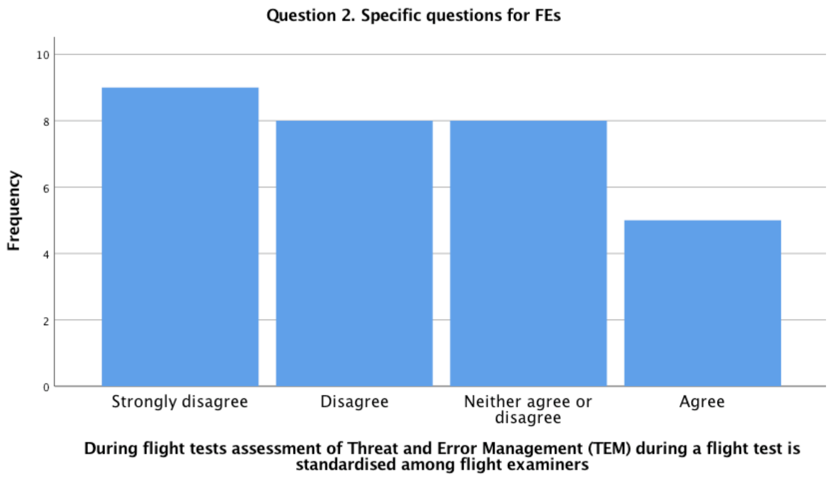

Figure 6.28. Second specific question for FEs.

The third question asked whether 'during flight tests, candidates verbalise their briefings (e.g., emergency briefing) without putting conscientious thought into them'. More than half (54.3%) of the FEs agreed or strongly agreed with this question (Figure 6.29). The result from the one-sample Wilcoxon signed rank test again supported this strong disagreement (p = .026).

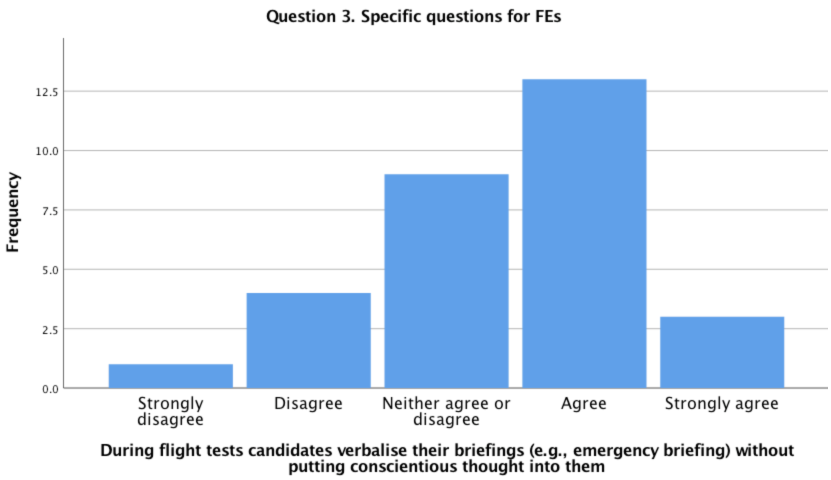

Figure 6.29. Third specific question for FEs.

The findings clearly suggested that there was a lack of support and guidance from CASA to FEs before TEM was implemented as an assessment item for licence issue flight tests. This has likely led to a lack of standardisation among FEs regarding assessment of TEM during flight tests. In addition, the FEs were more in agreement that flight test candidates were observed to artificially complete a series of safety briefings without conscientiously considering the prevailing conditions and situations.

6.3.1.7 Specific questions for flight instructors.

There were four specific questions, based on comments from Study 2, for those survey participants who held or had held a FI rating to gauge whether a series of briefings were completed without using conscientious thought, which was one of main criticism from participants in Study 2. Other questions were asked to determine the survey participants' understanding of TEM, as well as organisational supports to improve their instructing performance.

The first question asked whether 'during instructional flights, my students verbalise their briefings (e.g., emergency briefing) without putting conscientious thought into them'. There were mixed views on this question (Figure 6.30). The result from the one-sample Wilcoxon signed rank test supported this mixed view (p = .99).

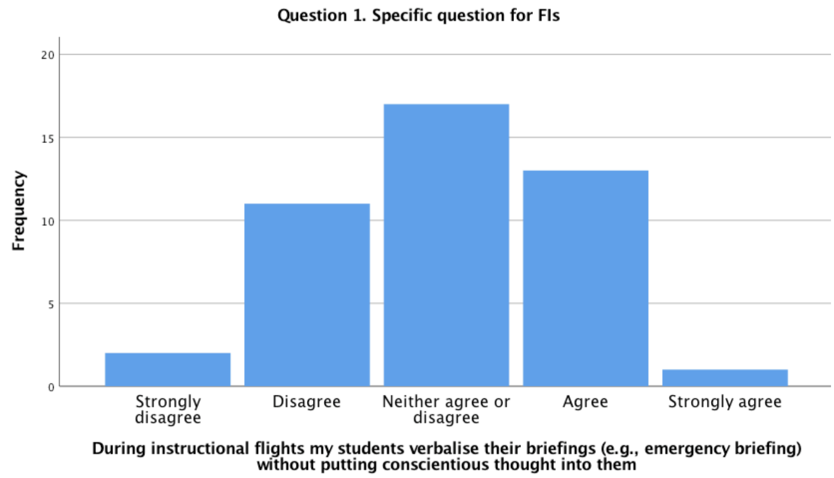

Figure 6.30. First FI-specific question.

The second question asked whether 'I find CASA's FI manual useful'. Again, there were mixed views on this question, with 10 survey participants indicating disagreement, while over 40% (n = 19) of survey participants recorded neutral responses (Figure 6.31). The result from the one-sample Wilcoxon signed rank test also supported this mixed view (p = .888).

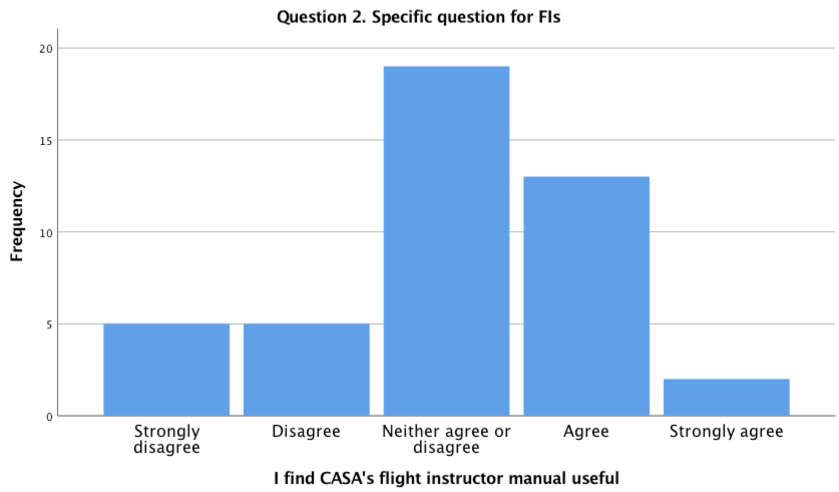

Figure 6.31. Second FI-specific question.

The third question asked whether 'I have a clear understanding of what I need to do when teaching TEM principles to my students'. More than half of the survey participants (60.9%) agreed or strongly agreed that they clearly understood what needed to be done when TEM principles were taught (Figure 6.32). The result from the one-sample Wilcoxon signed rank test supported this (p = .003).

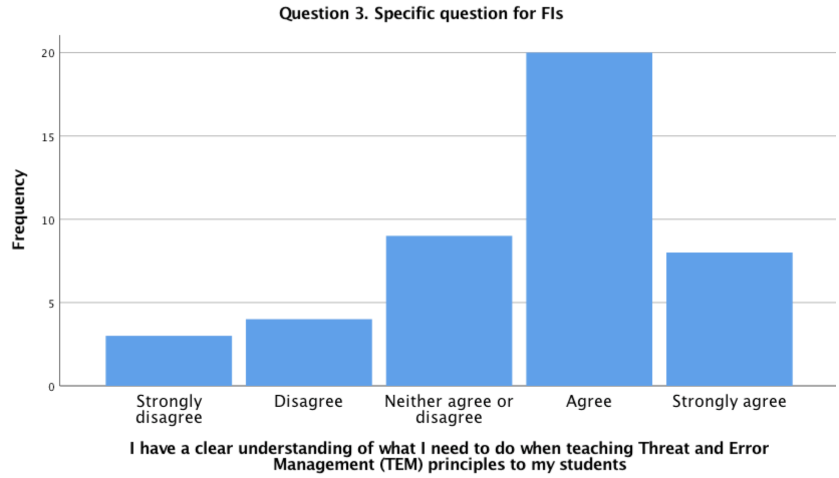

Figure 6.32. Third FI-specific question.

The fourth question asked whether 'my organisation provides adequate mentoring opportunities for junior FIs'. More than half of the survey participants (67.4%) agreed or strongly agreed that there were adequate mentoring opportunities for junior FIs within their organisations (Figure 6.33). The result from the one-sample Wilcoxon signed rank test supported this ($p < .001$).

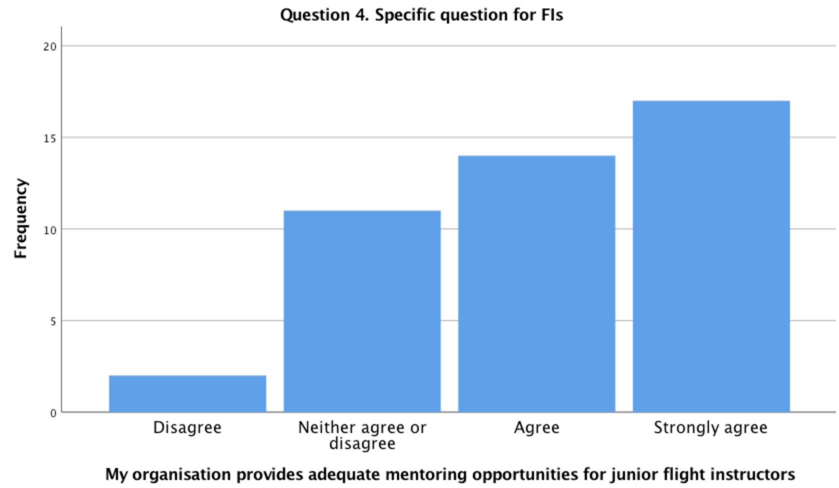

Figure 6.33. Fourth FI-specific question.

6.3.2 Results (second phase).

This section presents the second phase of the data analysis to create a model that displays the interrelationships among the four themes from Study 2 to test the hypotheses listed in Section 6.2.6. To achieve this, SPSS Amos (Version 25 for Windows) was used to perform Structural Equation Modelling (SEM).

6.3.2.1 Brief description of statistical procedures used.

SEM is a statistical modelling technique that takes a confirmatory (hypothesis-testing) approach to help researchers conceptualise a theoretical model (Byrne, 2016; Hox & Bechger, 1998). SEM has two parts: the first part is Confirmatory Factor Analysis (CFA) and the second part is path analysis (Hox & Bechger, 1998). CFA is used when a researcher has some knowledge of the underlying structure (Byrne, 2016). The current study was primarily based on the findings from Study 2, and the researcher had several a priori hypotheses: impracticality (first theme), lack of guidance and support (second theme) and TEM implementation (third theme). These themes were correlated and each affected the way TEM was taught and exercised (fourth theme). The second part involved pictorially indicating theoretical structural inter-relationship between the variables (themes and survey questions) in a model via path analysis, which is described as a statistical approach to estimate relationships and determine the strength of the paths in a SEM model (Hair, Black, Babin, Anderson, & Tatham, 2006).

6.3.2.2 Data screening processes.

Prior to examining the suitability of data for SEM, two data screening processes were completed. The first process aimed to ensure there were no missing responses. In this analysis, only completed survey responses were used; thus, there were no missing responses. The second process involved identifying any unengaged responses. Unengaged responses refer to the same responses for every survey question. These unengaged responses do not provide additional useful information about the variation within survey data; thus, they need to be deleted. The standard deviation calculation for each participant's responses was performed. The standard deviations for each participant's responses were above zero, which suggested that there were no unengaged responses from this survey dataset.

6.3.2.3 Sample validation processes.

Prior to performing SEM, there are several sample validation processes to be completed. The first step is to check multivariate normality within the dataset. Multivariate normality refers to a normal distribution to multiple variables (Field, 2013). A linear regression model was created based on 24 survey questions (six questions per theme), and any *p*-values less than 0.01 indicated multivariate outliers. Three participants' responses fell in this case and were removed from analysis.

The second step is to check homoscedasticity. Field (2013) described homoscedasticity as 'an assumption in regression analysis that the residuals at each level of the predictor variable(s) have similar variances' (p. 876). The scatter plot (Figure 6.34) below indicated that residuals were reasonably spread around the central zero line, thereby suggesting that homoscedasticity was adequately met.

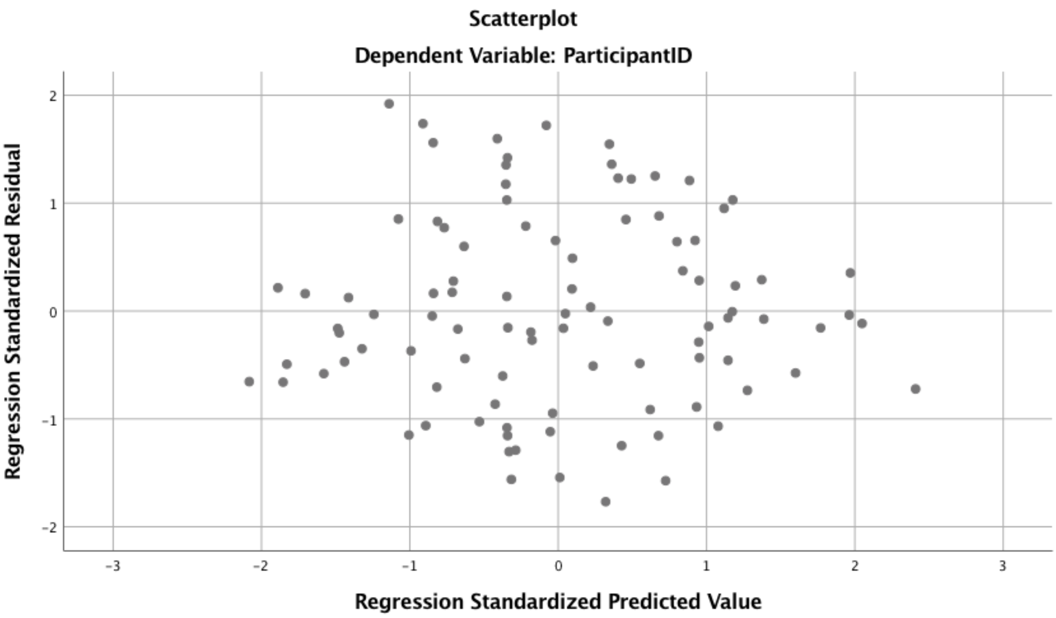

Figure 6.34. Scatter plot to check homoscedasticity.

The third step is to check for multicollinearity for each survey question. Field (2013) described multicollinearity as ‘a situation in which two or more variables are very closely linearly related’ (p. 879). Grewal, Cote, and Baumgartner (2004) cautioned that high multicollinearity, together with low measure reliability and small sample size, can lead to a high incidence of Type II errors; thus, survey questions with high multicollinearity should be removed from analysis. Linear regression with collinearity statistics indicated there was no high multicollinearity present in any of the current survey questions.

The fourth step is to determine the required sample size. Prior to calculating the sample size, it is necessary to determine how many latent variables (i.e., four themes) and observed variables (i.e., 19 survey questions from four themes) along with other values, as shown in Figure 6.35. The latent variables refer to the four themes identified in Study 2. The observed variables refer to the number of survey question for each them (six survey questions per theme). To finalise the total number of questions (i.e., observed variables) to determine the required sample size, Cronbach’s alpha was used to check the reliability of questions for each theme, and, if required, remove questions from a respective theme so that the Cronbach’s alpha was at least 0.7 (Hair et al., 2006).

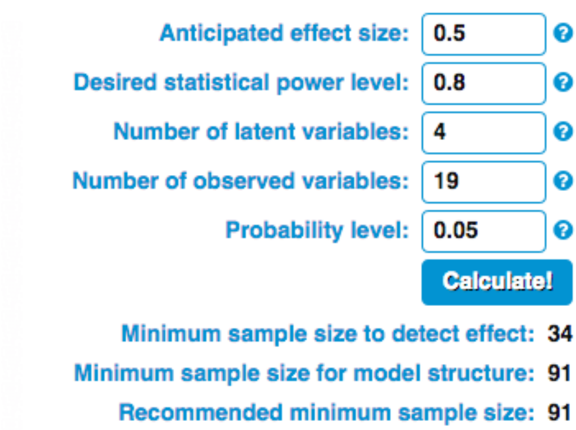

Figure 6.35. Estimation of sample size Source: accessed https://www.danielsoper.com/statcalc/calculator.aspx?id=89 on 2 August 2018.

The first reliability test on six survey questions on the first theme—impracticality—resulted in an overall Cronbach's alpha of .577. Closer inspection of the tabled item-total statistics showed that the value in the column labelled 'Cronbach's alpha if item deleted' for Question 3 was greater than the overall alpha; hence, the question was deleted to improve reliability. After removing the question 'when it comes to flying operations in general aviation, CASA places greater emphasis on the theoretical aspects of flying than the practical aspects', the overall Cronbach's alpha was increased to .797. Further inspection of the table showed that the value in the column labelled 'Cronbach's alpha if item deleted' for Question 2 was very slightly greater, yet very similar (i.e., .799). The researcher decided not to remove Question 2 because it would not affect the overall Cronbach's alpha. Therefore, based on the reliability test, five questions from the first theme were retained for later analyses.

The first reliability test on six survey questions on the second theme—lack of guidance and support—resulted in an overall Cronbach's alpha of .736. Although this satisfied the minimum Cronbach's value of 0.7, it was decided to remove the fourth question: 'I regularly attend CASA-organised safety seminars (e.g., AvSafety) to keep up to date with the latest aviation safety initiatives'. This question was removed because it had the value of .751 in the column labelled 'Cronbach's alpha if item deleted'. After removing the question, the overall Cronbach's alpha slightly increased to .751. Closer inspection of the column indicated that the value for the fifth question was still greater than the overall Cronbach's alpha. Consequently, the fifth question was removed: 'I find it useful to attend seminars, such as AvSafety and/or others organised by CASA'. After re-running the reliability test, the overall Cronbach's alpha was further increased to .791 and the values in the column (Cronbach's alpha if item deleted) were lower than this. Therefore, it was decided to use four questions from the second theme.

The first reliability test on six survey questions on the third theme—TEM implementation—resulted in an overall Cronbach's alpha of .875, and all the values in the column 'Cronbach's alpha if item deleted' were lower than the overall alpha. Therefore, it was decided to use all six questions from the third theme.

The first reliability test on six survey questions on the fourth theme—TEM in practice—resulted in an overall Cronbach's alpha of .581. Closer inspection of the tabled item-total statistics showed that the value in the column labelled 'Cronbach's alpha if item deleted' for the fifth question was .877. The question 'I find TEM briefing an unnecessary additional task' was removed, and a second reliability test was performed. Overall, the Cronbach's alpha increased to .877, but the value of the column (Cronbach's alpha if item deleted) for the fourth question was .905. The question 'I always perform a series of briefings (e.g., emergency and take-off briefings), taking into account the prevailing conditions and situations' was removed, and another reliability test was performed. Overall, the Cronbach's alpha further increased to .905, and all the values in the column (Cronbach's alpha if item deleted) were lower than the overall alpha. Therefore, it was decided to use four questions from the fourth theme.

After checking the overall Cronbach's alpha for each theme and its associated survey questions, a total of 19 survey questions from four themes (minimum four to maximum six questions per theme) were used in the following analyses, which satisfied the minimum of three, preferably four, survey questions per theme (Hair et al., 2006). In terms of sample size, an online a priori sample size calculator for SEM (https://www.danielsoper.com/statcalc/calculator.aspx?id=89) with 19 survey questions and other values (Figure 6.35) suggested the recommended minimum sample size of 91. Given that there were 94 completed survey responses, this satisfied the recommended minimum sample size required to analyse data using SEM.

6.3.2.4 Factor analysis.

Factor analysis is a class of statistical techniques to define the underlying structure among variables (Hair et al., 2006). The three main uses of factor analysis are as follows: (i) to understand the structure of a set of variables, (ii) to construct a questionnaire to measure an underlying variable and (iii) to reduce a dataset to a smaller manageable size without losing the key original information (Field, 2014). Broadly, there are two types of factor analysis: exploratory factor analysis (EFA) and CFA.

As the name suggests, EFA is used in situations where links between the observed and latent variables are unknown or uncertain (Byrne, 2016). This is a hypothesis-generating process that does not have a priori assumptions about the variables, factors and their interrelationships. The questions used in this study were based on the previous study (Study 2) to verify and generalise the findings from Study 2. In this study, there were several hypotheses already established; thus, EFA was not considered necessary.

6.3.2.5 Confirmatory Factor Analysis (CFA).

Bryne (2016) described CFA as a statistical technique used when a researcher has some a priori knowledge of the interrelationships between the observed variables and underlying factors. Performing CFA is the first of two steps involved in SEM (Hox & Bechger, 1998). In this study, after completing the sample validation process, an individual model for each theme was developed and checked using CFA to determine whether it was a good fit to the data, based on the following criteria (Table 6.10).

Table 6.10 *SEM Fit Indices.*

Fit measures	Full name	Acceptable scale	Note
CMIN/DF	Minimum discrepancy and its degree of freedom	≤ 2	
RMSEA	Root mean square error of approximation	$\leq .08$	Range between .08 and .10 is acceptable (mediocre fit)
IFI	Incremental fit index	> .90	> .95 is a good fit
NFI	Normed fit index	> .90	> .95 is a good fit
CFI	Comparative fit index	> .90	> .95 is a good fit

Source: adapted from Byrne (2016) and Hox and Bechger (1998).

For a model with inadequate fit, Modification Indices (MI) values are examined to obtain guidance regarding how the model needs to be modified (Hox & Bechger, 1998). For example, the model fit can be improved by identifying a pair of survey questions with the highest MI value and pairing the two using a double-headed arrow in the model.

The first theme—impracticality—was measured using five questions (Table 6.11). The initial standardised loadings were well above the minimum of .5 (Hair et al., 2006), thereby suggesting that all questions were highly correlated with each other.

Table 6.11 *Summary of Initial Findings (CFA) on the First Theme (Impracticality).*

Survey questions	Initial loadings	Final loadings
Q1. The CASA has appropriately translated TEM concepts into practical guidance for general aviation.	.83	.69
Q2. There are adequate practical examples relating to TEM principles and concepts for general aviation.	.66	Deleted
Q4. Senior executives and decision makers within the CASA have a good understanding of how general aviation operates.	.65	.71
Q5. I find it easy to keep up to date with changes to aviation safety regulations/initiatives.	.63	.72
Q6. I am satisfied with the way the CASA develops aviation safety regulations/initiatives.	.72	.79

The questions were subjected to CFA to check whether the model was a good fit to the data. The initial CFA analysis indicated that all the fit measures in Table 6.11 were outside the acceptable level of model fit statistics above (Table 6.10), thereby suggesting that the model was not a good fit. Therefore, modification of the model was required.

Table 6.12 *Achieved Fit Indices for First Theme (Impracticality).*

	CMIN/DF	RMSEA	IFI	Tucker-Lewis index (TLI)	CFI
Initial	6.726	.248	.844	.679	.840
Final	1.681	.086	.989	.966	.989

Upon inspection of the MI values, it was identified that e1 and e2 had the highest MI value (17.908). Therefore, an ad hoc attempt was made to retest the model after pairing the errors (Figure 6.36). This modification improved all the fit measures to the acceptable scales: CMIN/DF = 1.194, RMSEA = .046, IFI = .996, TLI = .989 and CFI = .996.

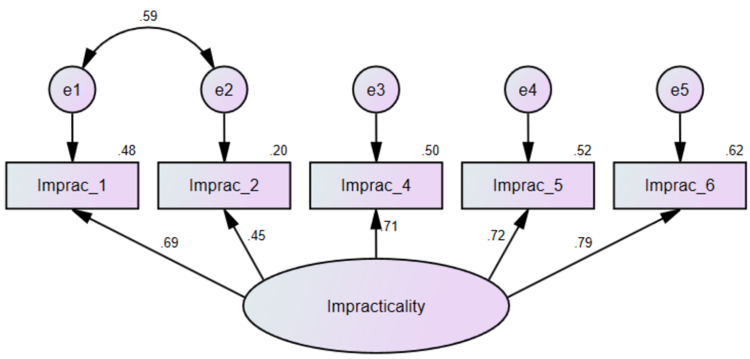

Figure 6.36. Five-item single model on impracticality.

However, the standardised loading for the second question was reduced to .45 (Figure 6.36), which suggested that it poorly correlated with all other questions in the model. Consequently, another ad hoc attempt was made to retest the model after deleting the second question (Figure 6.37). Deletion of the second question improved all fit measures to the acceptable scales (Table 6.12), as well as high standardised loadings.

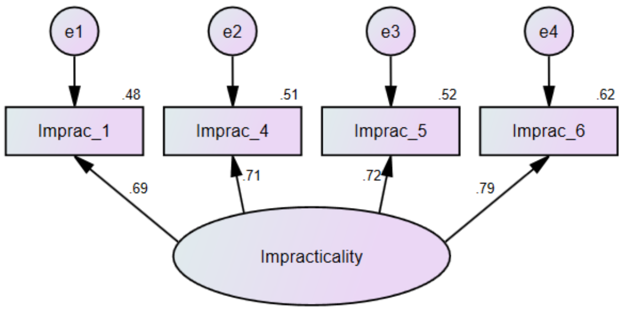

Figure 6.37. Revised four-item single model on impracticality.

Based on the findings generated with an online calculator (http://www.thestatisticalmind.com/calculators/comprel/composite_reliability.htm), the composite reliability for these four survey questions was .818. This was well above the acceptable level of .7 to indicate adequate internal consistency (Hair et al., 2006).

The second theme—lack of guidance and support—was measured using four questions (Table 6.13). The initial standardised loadings suggested that all the questions were highly correlated with each other (all above .50).

Table 6.13 *Summary of Initial Findings (CFA) on Second Theme (Lack of Guidance and Support).*

Survey questions	Initial loadings
Q1. The CASA provides adequate training and guidance material on TEM for general aviation.	.76
Q2. The concept of TEM is easy to understand.	.71
Q3. The CASA explains TEM and how it affects my role and/or activities in a clear and succinct manner.	.75
Q6. I find the number of safety-related publications and resources from the CASA useful.	.55

The questions were subjected to a CFA to check whether the model was a good fit to the data. The initial CFA analysis showed that all the fit measures in Table 6.14 were within the acceptable level of model fit statistics above (Table 6.10), thereby suggesting that the model was a good fit. Therefore, no modification to the model was required. Figure 6.38 displays the model for the second theme. The composite reliability for the four questions was .789, which was again above the acceptable level of .7 to indicate adequate internal consistency (Hair et al., 2006).

Table 6.14 *Achieved Fit Indices for Second Theme (Lack of Guidance and Support).*

	CMIN/DF	RMSEA	IFI	TLI	CFI
Final	.569	.000	1.008	1.026	1.000

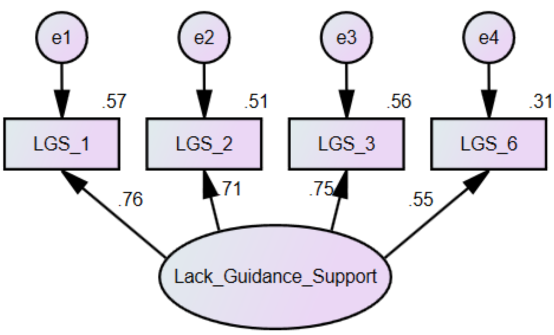

Figure 6.38. Four-item single model on lack of guidance and support.

The third theme—poor implementation—was measured using six questions (Table 6.15). The initial standardised loadings suggested that all the questions were highly correlated with each other (all above .50).

Table 6.15 *Summary of Initial Findings (CFA) on Third Theme (Poor Implementation).*

Survey questions	Initial loadings	Final loadings
CASA consults with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.	.81	.83
CASA ASAs have a consistent understanding of TEM principles.	.79	.79
I am satisfied with the way CASA implements new aviation safety regulations/initiatives.	.70	.66
The number of CASA staff with appropriate expertise in TEM is adequate.	.59	.58
CASA maintains an effective ongoing dialogue with the general aviation industry.	.78	.79
CASA explains new safety regulations/initiatives and how they affect the general aviation industry in a clear and succinct manner.	.77	.74

The questions were subjected to CFA to check whether the model was a good fit to the data. The initial CFA analysis showed that two fit measures in Table 6.16 were slightly outside the acceptable level of model fit statistics (Table 6.10), suggesting that the model was not a good fit. Therefore, modification of the model was required.

Table 6.16 *Achieved Fit Indices for Third Theme (Poor Implementation).*

	CMIN/DF	RMSEA	IFI	TLI	CFI
Initial	2.067	.107	.964	.939	.964
Final	1.685	.086	.980	.961	.979

Upon inspection of the MI values, it was identified that e3 and e6 was the only covariance listed with the MI value of 4.412. Therefore, an ad hoc attempt was made to retest the model after pairing the errors (Figure 6.39), which improved all the fit measures to an acceptable level of model fit statistics (Table 6.16). The composite reliability for the four questions was .875, which was again well above the acceptable level of .7.

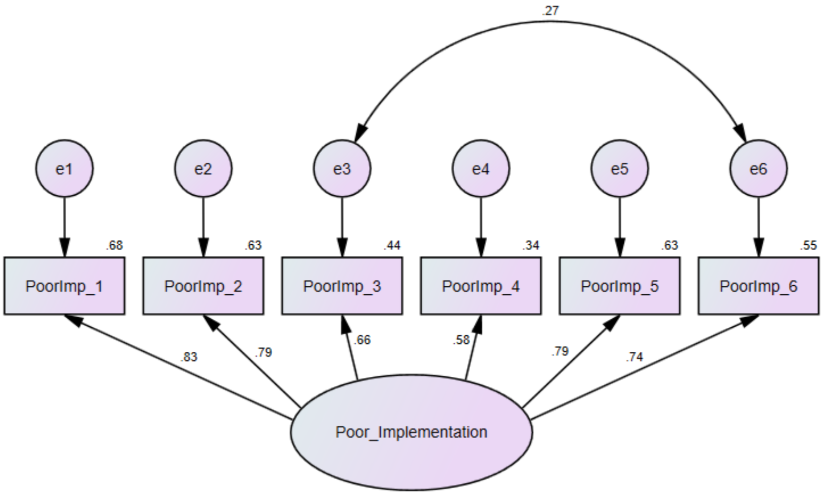

Figure 6.39. Six-item single model on poor implementation.

The fourth theme—TEM in practice—was measured using four questions (Table 6.17). The initial standardised loadings suggested that all questions were highly correlated with each other (all above .50).

Table 6.17 *Summary of Initial Findings (CFA) on Fourth Theme (TEM in Practice).*

Survey questions	Initial loadings	Final loadings
I use TEM principles when I am undertaking flying activities.	.92	.96
I have a clear understanding of what I need to do when applying TEM principles.	.85	.85
Consideration of TEM is an important part of my flight preparation.	.84	.79
I can see value in TEM when I am undertaking flying activities.	.84	.79

The questions were subjected to CFA to check whether the model was a good fit to the data. The initial CFA analysis showed that three fit measures in Table 6.18 were outside the acceptable level of model fit statistics (Table 6.10), suggesting that the model was not a good fit. Therefore, modification of the model was required.

Table 6.18 *Achieved Fit Indices for Fourth Theme (TEM in Practice).*

	CMIN/DF	RMSEA	IFI	TLI	CFI
Initial	6.578	.245	.961	.882	.961
Final	.362	.000	1.002	1.013	1.000

Upon inspection of the MI values, it was identified that e3 and e4 was the covariance with the highest MI value of 9.261. Therefore, an ad hoc attempt was made to retest the model after pairing the errors (Figure 6.40), which improved all the fit measures to an acceptable level of model fit statistics (Table 6.18). The composite reliability for the four questions was .912, which was again well above the acceptable level of .7.

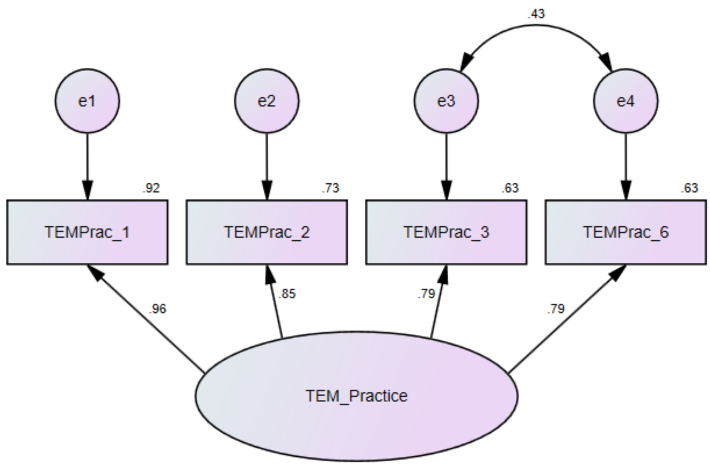

Figure 6.40. Revised four-item single model on TEM in practice.

A series of CFAs was performed to examine whether each individual measurement model had a good fit to the data, and the final fit indices for each theme confirmed that the models were a good fit. Therefore, it was appropriate to move onto the next step to test the structural relationship among the themes and their responses, as well as testing the hypotheses (H_5 to H_{10}).

Prior to testing the structural relationship among the themes, it was necessary to determine which dependence relationships existed in these themes. The central focus of this thesis was exploring the effect of TEM in Australian general aviation. Based on the findings from Studies 1 and 2, it was hypothesised that impracticality and lack of guidance and support from CASA were correlated, and each negatively affected the way TEM was viewed, taught and practised in general aviation. Therefore, impracticality and lack of guidance and support from the CASA were considered exogenous variables (independent variables) and TEM in practice was considered an endogenous variable (dependent variable). In addition, it was hypothesised that both impracticality and a lack of guidance and support from CASA affected the survey participants' views on poor TEM implementation, and this mediated the effect of impracticality and lack of guidance and support on the way TEM was viewed and practised. Based on these notions, six hypotheses were proposed, which formed hypothesised structural relationships (Figure 6.41). For instance, a single-headed arrow from impracticality to TEM in practice (H_5) shows that TEM in practice is dependent on impracticality—that is, the higher the impracticality, the less successfully TEM is practised. The hypotheses were as follows:

H_5: Impracticality negatively affected the way TEM was viewed and practised.

H_6: Impracticality resulted in survey participants' view of poor TEM implementation

H_7: Lack of guidance and support from CASA negatively affected the way TEM was viewed and practised.

H_8: Lack of guidance and support from CASA resulted in participants' view of poor TEM implementation

H_9: Poor TEM implementation mediated the effect of impracticality on the way TEM was viewed and practised.

H_{10}: Poor TEM implementation mediated the effect of lack of guidance and support from CASA on the way TEM was viewed and practised.

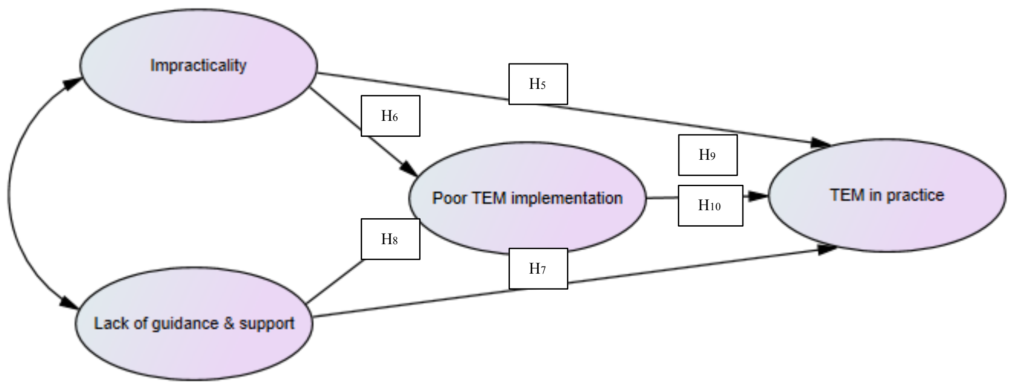

Figure 6.41. Simplified path diagram showing the hypothesised structural relationships.

To estimate and assess how well the SEM model fit, a complete model with themes and survey questions for each theme was required, as presented in Figure 6.42.

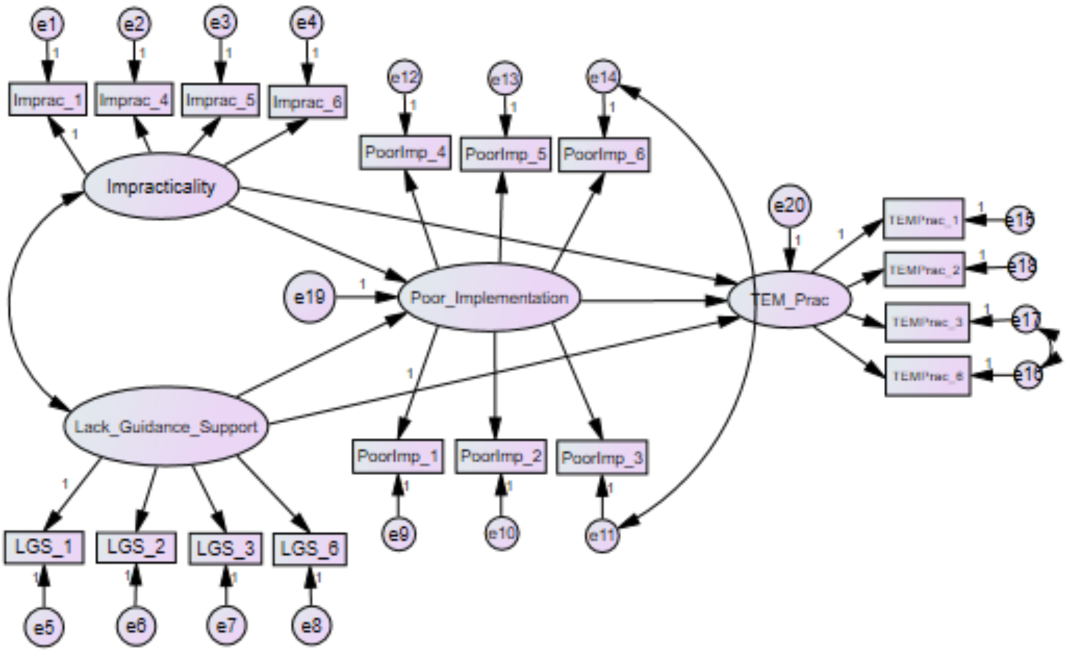

Figure 6.42. Complete path diagram showing specified hypothesised structural relationships.

The complete path diagram was subjected to CFA to check whether the model was a good fit to the data. The initial CFA analysis showed that one of the fit measures (i.e., RMSEA) in Table 6.19 needed to be improved in accordance with the SEM fit indices (Table 6.10). Therefore, modification of the model was required.

Table 6.19 *Achieved Fit Indices for Complete Path Diagram Showing Specified Hypothesised Structural Relationships.*

	CMIN/DF	RMSEA	IFI	TLI	CFI
Initial	1.657	.083	.922	.903	.920
Final	1.554	.076	.934	.918	.933

Upon inspection of the MI values, it was identified that the pair (e7 and e14) was the covariance with the highest MI value of 13.034. Therefore, an ad hoc attempt was made to retest the model after pairing the errors, which improved the fit measures to an acceptable level (Table 6.19). Figure 6.43 displays a revised complete path diagram showing the specified hypothesised structural relationships. The revised model with the standardised regression weights is presented in Appendix 11.

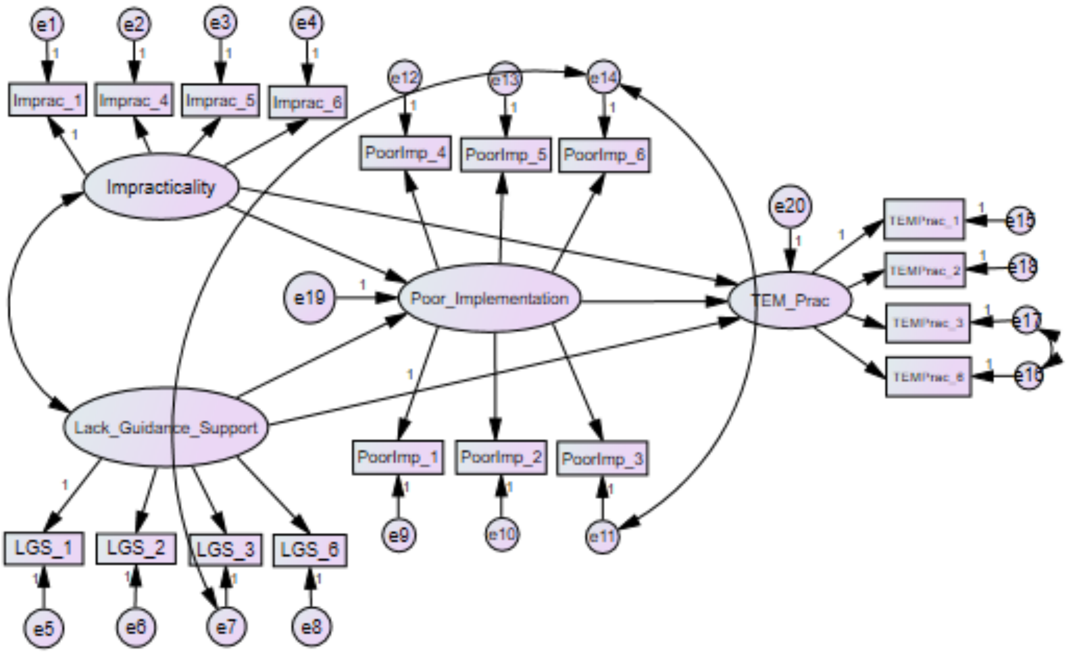

Figure 6.43. Revised path diagram showing specified hypothesised structural relationships.

Having determined the model as a good fit, path analysis was performed to determine the correlation between the themes (impracticality and lack of guidance and support), as presented in Table 6.20.

Table 6.20 *Correlation Coefficient between a Pair of Themes.*

Correlation path	Estimate
Impracticality ↔ Lack of guidance and support	.836

The pair was found to be highly and positively correlated, thereby suggesting that any changes in one theme significantly and positively affected the other. For instance, if more survey participants agreed that TEM lacked practicality, then more survey participants would agree that there was a lack of guidance and support from CASA. The final step was to test the hypotheses to determine each theme's effect on the other themes. The results of the path analysis are presented in Table 6.21.

Table 6.21 *Hypotheses Testing Results.*

Hypothesis	Hypothesised path	Estimates	*p*-value	Is hypothesis supported?
H_5	Impracticality → TEM in practice	.002	.998	NO
H_6	Impracticality → Poor TEM implementation	.541	< .001	YES
H_7	Lack of guidance and support → TEM in practice	1.219	.046	YES
H_8	Lack of guidance and support → Poor TEM implementation	.355	.019	YES

To test Hypotheses 9 and 10, the figures in Table 6.22 were used to initially calculate the test statistic (z-score), which was then converted to respective *p*-values.

Table 6.22 *Regression Weights for Impracticality and Lack of Guidance and Support in Relation to Poor TEM Implementation and TEM in Practice.*

Construct	Path	Construct	Estimate	Standard error
Impracticality	→	Poor TEM implementation	.541	.160
Lack of guidance and support	→	Poor TEM implementation	.355	.151
Poor TEM implementation	→	TEM in practice	-.365	1.257

The calculation with poor TEM implementation as a mediating variable indicated that poor TEM implementation did not mediate the effect of impracticality ($p = .77$) on TEM in practice. The calculation also indicated that poor TEM implementation did not mediate the effect of lack of guidance and support ($p = .77$) on TEM in practice. Therefore, Hypotheses 9 and 10 were not supported. The following section provides interim discussions based on the findings from this study.

6.3.3 Interim discussion.

The main findings from this study are briefly discussed in the following subsections. More thorough discussions are provided in Chapter 7: Overall Discussion.

6.3.3.1 *Impracticality.*

The survey findings suggested that a statistically significant number of survey participants were more in agreement that TEM principles and concepts lacked practicality when they were introduced and implemented in Australian general aviation. The findings on the lack of practicality were found to not align with one of the expectations of the Minister for Infrastructure and Transport that CASA's regulatory approach would be pragmatic, practical and proportionate as it applied to different industry sectors (CASA, 2018).

The general consensus was that CASA failed to provide practical guidance and examples about how and when TEM principles and concepts should be adopted in daily operations. This view was primarily based on the notion that senior executives and decision makers within CASA did not have a good understanding of how general aviation operated. This also negatively affected the survey participants' strong dissatisfaction with the way CASA developed aviation safety regulations and initiatives, with a statistically significant number of participants in the pre-TEM group displaying greater dissatisfaction. Greater emphasis on the theoretical aspect of flying, as well as difficulty encountered by the majority of survey participants in keeping up to date with changes to aviation safety regulations and initiatives, further aggravated the view that the TEM principles and concepts lacked practicality.

Mearns and Reader (2008) found that stronger health support from an organisation predicted more safety citizenship behaviours from its employees, such as assisting colleagues with safety-related matters and reporting hazards at work. Mearns and Reader (2008) explained that employees might perceive organisational investment in their health as being indicative of the organisation's strong commitment to health and safety, which led to employees placing higher priorities on safe conduct in the workplace. Indeed, this type of work environment will more likely enhance the overall safety climate and outcomes, which is also the highest priority of CASA (2018b)—to focus on aviation safety. Thus, the findings supported the first hypothesis.

In terms of impracticality and its effect on TEM in practice, impracticality was not found to have a significant effect on how TEM was viewed and practised. This was an interesting and perhaps counterintuitive finding, as the logical assumption was that, if the survey participants did not view TEM as practical, they would be less likely to use TEM in their daily flying. There were mixed views among survey participants about whether they had a clear understanding of what was required when applying TEM principles while undertaking their flying activities. This was partly caused by CASA failing to translate TEM concepts into practical guidance for general aviation. However, a statistically significant number of survey participants—particularly ones with FI and FE ratings—agreed that they used TEM principles when undertaking flying activities, and that consideration of TEM was an important part of their flight preparation. Thus, the findings did not support the fifth hypothesis.

When examining the inter-relationship between impracticality and poor TEM implementation, it was found that the relationship was positively correlated and statistically significant, thereby suggesting that the more survey participants felt TEM was impractical, the more participants would agree that TEM was poorly implemented. This is an important aspect because general aviation is vocationally oriented. Consequently, providing practical examples relating to TEM principles and concepts and placing greater emphasis on the practical aspects of flying operations would enhance the view that TEM was implemented well. Thus, the findings supported the sixth hypothesis.

Poor TEM implementation was a mediating variable, and it was hypothesised that poor TEM implementation mediated the effect of impracticality in terms of the way TEM was viewed and practised. This hypothesis was not supported. It was established earlier that, although the inter-relationship between practicality and poor TEM implementation was statistically significant, impracticality was not found to have a significant effect on how TEM was viewed and practised.

In summary, a statistically significant number of survey participants agreed that TEM lacked practicality, and this would likely have caused the survey participants to agree that TEM was poorly implemented in Australian general aviation. However, the notion of impracticality negatively affecting the way TEM was viewed and practised was not supported.

6.3.3.2 *Lack of guidance and support.*

The findings on lack of guidance and support from CASA suggested mixed views among the survey participants. A statistically significant number of participants were more in agreement that CASA did not provide adequate training and guidance material on TEM. Similarly, the majority of participants did not agree that CASA clearly and succinctly explained TEM and the way it would affect their role and activities. These views may have contributed to the result that the majority of participants did not find TEM easy to understand. Based on these findings, the second hypothesis was supported.

In contrast, there were mixed views among the survey participants on the usefulness of CASA-organised seminars (e.g., AvSafety) and CASA's safety-related publications and resources. Similarly, there were mixed views regarding regular attendance at CASA-organised seminars (e.g., AvSafety) for the participants to keep up to date with the latest aviation safety initiatives. It was found that the non-FE group, in particular, did not regularly attend the safety seminars. Based on these findings, the second hypothesis was not supported. Overall, the second hypothesis was partially supported as the survey participants were in agreement for the first three survey questions of the second theme but not the remaining three survey questions.

Numerous previous research findings have suggested that adequate guidance and support are one of the key aspects leading to successful implementation of new initiatives. For instance, according to McFadden, Stock, and Gowen (2006), lack of top management support, lack of resources and fewer educational opportunities were found to have a significantly negative effect on the implementation of patient safety initiatives in US hospitals. Similarly, Wilcock, Ball, and Fajumo (2011) found that provision of government resources and effective communication among all personnel within the company and with industry and government officials regarding specifications and guidelines were critical factors for the successful implementation of food safety initiatives. Likewise, Alderson's (2009) study on the language proficiency requirements proposed by ICAO highlighted that, even with a detailed document (a manual on the implementation of ICAO language proficiency requirements) to assist each contracting state with developing suitable and appropriate testing methods, there were serious inadequacies in the state of preparedness of national civil aviation authorities to comply with this ICAO initiative.

Based on the survey findings for Study 3, CASA failed to provide adequate training, guidance materials or explanation about how TEM affected the survey participants' role and activities in a clear and succinct manner, which led to the general consensus that the TEM concept was not easy to understand. These factors negatively affected the way TEM was viewed and practised. In addition, this negatively resulted in the participants' views of poor TEM implementation. Based on these results, both the seventh and eighth hypotheses were supported.

Poor TEM implementation was a mediating variable, and it was hypothesised that poor TEM implementation mediated the effect of lack of guidance and support from CASA on how TEM was viewed and practised. It was already established that the lack of guidance and support from CASA negatively affected the way TEM was viewed and practised, which resulted in participants' views of the poor TEM implementation. However, the poor TEM implementation did not mediate the effect of lack of guidance and support from CASA. Thus, the tenth hypothesis was not supported.

In summary, a statistically significant number of survey participants agreed that CASA failed to provide adequate training and guidance material on TEM, which would likely have caused the suboptimal level of understanding of TEM concepts and principles. However, it was found that there were mixed views among the survey participants on the usefulness of CASA-organised seminars, such as AvSafety, and the CASA's safety-related publications and resources on TEM. Lack of guidance and support from CASA had a strong and direct effect on the way TEM was implemented, viewed and practised in Australian general aviation.

6.3.3.3 Threat and error management implementation.

The survey findings suggested that a statistically significant number of survey participants were more in agreement that TEM was poorly implemented in Australian general aviation. The majority of survey participants found that CASA did not consult with all relevant stakeholders when developing or reforming aviation safety initiatives. In addition, CASA was found not to maintain an effective ongoing dialogue with the general aviation industry.

According to a guide provided by ANAO (2014), one of the key considerations when planning to implement a policy initiative is to establish a high level of stakeholder engagement with the policy initiative. This requires a clear objective for consultation and the identification of key stakeholders, while maintaining clear and timely communications with wider communities. This was closely aligned with CASA's (2018d) regulatory philosophy, where CASA aimed to develop and implement appropriate policies consultatively and collaboratively. However, this was not supported by the findings in this study, or the findings of another online survey with 1,217 stakeholder participants, where almost half of the survey participants indicated dissatisfaction with CASA's ongoing dialogue with industry, and only 11% agreed that CASA always consulted with the most appropriate people in the industry when developing and reforming aviation safety regulations (Colmar Brunton, 2016).

The findings from the present study also highlighted the inadequate number of CASA staff with appropriate expertise in TEM, and the inconsistent understanding of TEM principles among CASA ASAs. The importance of ensuring the appropriate mix of required expertise and quantity of resources is another key factor for the successful implementation of a policy initiative. In the current study, it was acknowledged that the unavailability of personnel with the required expertise to implement the planned initiative was one of the most common implementation problems (ANAO, 2016).
These findings and the survey participants' collective view on TEM lacking practicality and the lack of guidance and support from CASA most likely led to the strong dissatisfaction among the survey participants with the way CASA implements new aviation safety initiatives, such as TEM. Thus, the third hypothesis was supported,

suggesting that the survey participants were more in agreement that TEM implementation in Australian general aviation was poor.

6.3.3.4 *Threat and error management in use.*

The majority of findings thus far have suggested that TEM was poorly implemented in Australian general aviation, and the lack of guidance and support from CASA particularly contributed to this collective view among the survey participants. It was assumed that these views would lead to further negative outcomes in terms of the ways TEM was considered and practised. For example, a previous study by Conley et al. (2011) suggested that a hospital with less effective implementation of a surgical safety checklist had not been provided with dedicated education or training, which led to poor frequency and quality of checklist use. However, in the current study, the responses to questions on the way TEM was viewed and practised were somewhat contrasting. Generally, the survey participants highly valued TEM when undertaking flying activities. As such, they considered TEM an important part of their flight preparations, and used TEM principles when undertaking flying activities. Therefore, this finding did not support the fourth hypothesis—that survey participants were more in agreement that TEM was poorly practised in Australian general aviation.

There were mixed views regarding whether the participants held a clear understanding of what needed to be done when applying TEM principles and whether TEM briefings were an unnecessary additional task. A closer examination of the responses indicated that the non-FI and non-FE groups were less clear about what needed to be completed when applying TEM principles. This may be the reason that the non-FI group considered the benefits of TEM training to be overrated, and the non-FE group more disagreed with the idea of TEM training and principles being expanded to other general aviation segments, such as the recreational aviation and RPAS sectors.

6.3.3.5 *Group-specific questions.*

Based on various comments from the interview participants in Study 2, there were three additional questions for survey participants with FE ratings, and four additional questions for those with FI ratings.

A statistically significant number of FEs disagreed that CASA provided adequate guidance to examiners before TEM became a mandatory flight test item. This may be the reason why assessment of TEM during flight tests was not standardised among FEs. Another observation was that flight test candidates were verbalising their safety briefings (e.g., emergency briefing) without considering the prevailing conditions and situations.

In terms of the responses to questions for the FIs, a statistically significant number of FIs had a clear understanding of what was required when teaching TEM principles to their students, and their organisations provided adequate mentoring opportunities for junior instructors. These were very encouraging findings. However, there were mixed views on the usefulness of the CASA's FI manual and students verbalising their safety briefings (e.g., emergency briefings) without putting conscientious thought into them.

6.4 Chapter Summary

This chapter has discussed the quantitative (second) phase of the sequential mixed methods research to verify and generalise the findings from the qualitative (first) phase of the study. The primary purpose of this study was to collect information, so that the findings from Study 2 (qualitative study) could be verified and generalised. The findings suggested that TEM implementation in Australian general aviation was problematic, and this view was due to impracticality and a lack of guidance and support from the CASA. An interesting and perhaps counterintuitive finding was that, despite poor TEM implementation and the lack of guidance and support from CASA, a statistically significant number of survey participants highly valued TEM when undertaking flying activities, which led them to view TEM as an important part of their flight preparations and use TEM principles when undertaking flying activities.

Chapter 7: Overall Discussion

The findings from the three studies in this thesis suggested that TEM was poorly implemented in Australian general aviation. The findings also indicated that the two major causes for this—impracticality and a lack of guidance and support from CASA—were the main contributors to the collective views among the study participants regarding poor TEM implementation. Therefore, it was reasonable to assume that this would lead to negative views about the use of TEM in Australian general aviation. However, the collective responses from Study 3 suggested that the survey participants highly valued and applied TEM principles when undertaking flying activities, and consequently considered TEM an important part of flight preparations. This chapter provides an overall discussion of the key findings from the three studies.

7.1 Poor Threat and Error Management Implementation

The Australian National Audit Office (ANAO, 2014) provided guidance that identified key considerations 'when implementing a policy initiative—the act of translating policy into reality—so that intended benefits are realised' (p. 3). The six key considerations are governance, managing risk, engaging stakeholders, planning, resources and monitoring, and review and evaluation (ANAO, 2014). The ANAO (2014) described governance as establishing the arrangements and practices that enable the achievement of the expected outcome. Sound governance arrangements and practices require a strong commitment from executives who have a good understanding of the industry (e.g., general aviation) so that the appropriate resources can be allocated. However, an observation from Participant 1 noted a different situation in Australian general aviation:

> *'CASA particularly, it means you will have to put some resources on the ground out there and … but if you look at the profile of a lot of senior management within CASA, ones we got last probably 10 to 15 years, has very little expertise in aviation there. So, they are driving the show and, they are not too much focusing into the operational aspect, uh, what's*

> *going on in the operational field, be it in airlines or be it in the training industry. They are not going to put their focus there because, by large, they don't understand it.' (Participant 1)*

This observation was supported by the majority of survey participants (88.7%) in Study 3, who disagreed that senior executives and decision makers within CASA had a good understanding of how general aviation operated.

A strong commitment from executives is often supplemented by a senior responsible officer, whose accountability includes successful policy implementation (ANAO, 2014). If the officer does not possess the appropriate skills to ensure successful policy implementation, she or he must access an appropriately skilled implementation team to assist her or him. However, the findings from the current research suggested that the concept of TEM and the ways to implement it were not well established within CASA from the beginning. Participant 3 recounted that:

> *'… when they were introducing it, they were sort of saying things like 'we're only still trying to find out how to do it'. Even later, I spoke to a senior manager at CASA, and he said, 'we have no idea what it's about', so we were enforcing something that we didn't understand.' (Participant 3)*

The perceived lack of relevance of new safety initiatives, such as CRM and TEM, when first introduced, adversely affects the facilitation of sufficient understanding of the topics among the intended recipients. Entwistle (1992, as cited in Rudland & Rennie, 2003) identified that the degree of learning is highly correlated with the perceived relevance of subject materials, whereby a high level of perceived relevance promotes a deep approach to learning, while a lack of perceived relevance is associated with a superficial approach. Rudland and Rennie (2003) added that integration of theory and practical relevance could be enhanced by using clinical problems so that students were able to identify and apply relevant gained knowledge to core clinical problems which in turn improves learning experience and outcomes. These findings highlight the need to ensure that new safety initiatives, such as TEM, accompany a clear statement about their relevance to the individuals and industry, supported by examples, so that greater buy-in can be obtained. However, the findings from Study 3 suggested that a statistically significant number of survey participants did not agree that CASA had clearly and succinctly explained TEM and how it would affect their role and/or activities. In addition, the majority of participants (81.4%) did not agree that CASA consulted with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM.

In terms of establishing an appropriately qualified and skilled implementation team with expertise in TEM, Participant 4 commented that:

> *'I believe at that time, um, CASA had engaged in a consultant, um, who I don't think was from an aviation background at that time, uh. A lot of what came out made no sense whatsoever, um, and it proved, at that time, quite difficult to find anything aviation specific with threat and error management because it had only been introduced, um, for its integration into a training program ... there was no technical specialist actually employed by CASA.' (Participant 4)*

This was supported by a finding from Study 3 in which only seven out of 97 (7.2%) survey participants agreed that the number of CASA staff with appropriate expertise in TEM was adequate. A similar finding was obtained from a previous study's stakeholder survey, in which only 26% of survey participants agreed that CASA provided competent and capable staff (Colman Brunton, 2016). Participant 2 in Study 2 added:

> *'CASA does not ... it is a worst government department in the world, right? It does not train its staff ... So, someone's upstairs and the standard's been written all of these beautiful stuffs, right? But they really haven't trained their staff. So, their staff can't, don't understand it, can't deliver it, and that's a history of CASA.' (Participant 2)*

These findings expand on the collective responses from Study 3, in which a statistically significant number of survey participants did not agree that CASA provided adequate training and guidance materials on TEM training for general aviation.

A similar situation arose when CRM was first introduced in the US. Helmreich et al. (1999) described that the first generation of CRM involved conducting intensive training sessions, with a heavy focus on psychological testing. The exercises used in the training sessions were unrelated to aviation to illustrate the relevant concepts. As a result, this training led to resistance from some pilots (Helmreich et al., 1999). In the current study, it appeared that the lack of adequate understanding of TEM concepts and seemingly inadequate expertise available within CASA likely led to the manifestation of poor TEM implementation in Australian general aviation.

The second key consideration for successful implementation of a new policy is risk management (ANAO, 2014). In this context, risk refers to a deviation from and failure to achieve the intended outcomes. CASA (2008) introduced TEM in response to ICAO's endorsement and recommendation to include TEM training in all pilot training, with a view to further improve flight safety in general aviation. Therefore, adequate risk management aimed to ensure that TEM was properly introduced, taught, practised and assessed in general aviation to achieve the intended outcome of further improving flight safety in general aviation. More than half of the participants (56.4%) from Study 1 agreed or strongly agreed that TEM training had improved overall safety in general aviation. In addition, almost half of participants (48.7%) did not feel that the benefit of TEM training was overrated. Further, less than half of survey participants (41.2%) from Study 3 felt that the benefit of TEM training was overrated. However, the participants in Study 2 had a different view of the effect of TEM on general aviation safety: *'My observation is that this has no effect on safety, I don't think the statistics has changed. The things that would've changed statistics were probably the technology improvement on the aeroplane' (Participant 3).* Other participants in Study 2 shared a similar view. Participant 2 stated: *'Go and have a look at statistics. I dare to say the rate in GA [General Aviation] hasn't moved in 50 years' (Participant 2).* Participant 4 stated:

'I don't think you will actually find a significant change … In fact, you will find, unfortunately, during that period, there's actually been a decrease in general aviation movements, um, and therefore a decrease in movement, you've got to have a decrease in incidents, but proportionally I actually doubt you will see much of a change.' (Participant 4)

According to ATSB (2018), the number of departures involving all aircraft in Australian general aviation has fluctuated over the last 10 years. The number of departures[13] was 1,985,000 in 2010, which decreased to 1,853,000 in 2011 and 1,760,000 in 2012, before starting to increase again from 2013 (ATSB, 2018). This trend supported Participant 4's comment on aircraft movement in general aviation. When examining the data on all general aviation occurrences between 2007 and 2016, it was noted that the number of accidents per million departures was 64 in 2010, which decreased to 61.5 in 2011, 58.5 in 2012 and 49.7 in 2013, but increased to 67.7 in 2014. Overall, there were no definitive downward trends noted in the number of accidents per million departures over the 2007 and 2015 period, while the private/business and sports sectors in Australian general aviation experienced the highest average rate of reported accidents during the same period (ATSB, 2018).

Another way to evaluate an appropriate level of risk management is to determine how TEM is assessed. The current research's findings suggested that risk was not adequately managed. For instance, the majority of participants from the trainee and trainer groups in Study 1 did not believe that TEM was appropriately assessed during flight tests. Participant 5 shared his experience with regard to encountering difficulties when assessing TEM during a flight test:

'… it is a very difficult one to assess from an examiner's point of view. It's almost a bit like trying to assess common sense. You know, no one can write a sheet to actually assess common sense, and I found threat and error management falls somewhat into that category.' (Participant 5)

[13] The ATSB's (2018) report on the aviation occurrence statistics used departure data when calculating accident and fatal accident rates because departures were a more appropriate measure than hours flown, as most accidents occur either during the approach and landing or departure phases of flight.

These difficulties were due to CASA not delivering required training for FEs. Participant 1 recalled that:

> *'PDP every two years … it didn't focus on TEM, it didn't mention it and there was no separate PDP program to cover it to say, 'this is a new thing, we, CASA, have introduced to the industry … we really want to focus on it and we now want you to focus on it and therefore we will give you separate training'. No, they didn't do that.' (Participant 1)*

The findings from Study 3 supported Participant 1's experience. The majority (83.4%) of FEs in Study 3 did not receive adequate guidance from CASA before TEM became a mandatory flight test item. This led to participants' observation in Study 2 that assessment of TEM demonstration was not standardised among FEs. A finding from Study 3 indicated that only 16.7% of FEs agreed that assessment of TEM during a flight test was standardised among FEs. These findings again highlighted the importance of pre-training and education roadshows before new safety initiatives are introduced.

The third key consideration for successful implementation of new policy is engaging stakeholders (ANAO, 2014). Hartzler et al. (2013) described stakeholder engagement as activities to involve, consult and partner with key stakeholders throughout the implementation phases. However, the findings from Studies 2 and 3 did not suggest that an appropriate level of stakeholder engagement was ensured before TEM was implemented. Participant 4 shared his experience in this regard:

> *'A regulator that is quite removed from industry, so if you look at the majority of people making the rules or publishing the guidance material, they are many, many, many, many years from removed from the cold face and things are quite often done reactively, rather than proactively, and without a huge amount of engagement.' (Participant 4)*

Participant 4 continued: *'They might call for, uh, you know, they might put out, um, a Notice of Proposed Rule Making, [and] people may respond, but their response rate is about 1% in the community'*. A closer inspection of the consultation and project history on the CASA website suggested that there was no industry-wide consultation completed on TEM before it was implemented. A finding from Study 3 supported that a statistically significant number of survey participants (73.2%) disagreed that CASA had explained TEM and how it would affect their role and/or activities in a clear and succinct manner. Another finding from Study 3 indicated that a statistically significant number of survey participants (81.4%) disagreed that CASA consulted with all relevant stakeholders when developing and/or reforming aviation safety initiatives, such as TEM. The survey participants who were actively flying before TEM was implemented showed stronger disagreement. Another finding regarding this lack of stakeholder engagement indicated that only eight survey participants (8.2%) agreed that CASA maintained an effective ongoing dialogue with the general aviation industry. However, these findings did not suggest that there was no attempt to engage stakeholders before TEM was implemented. A closer inspection of the website indicated that a Civil Aviation Advisory Publication (CAAP)—titled 'Teaching and Assessing Single Pilot Human Factors and Threat and Error Management'—was released on 21 October 2008. However, Participant 1 commented that:

> *'I mean, they did the usual thing of producing some CAAP information … but that was just a document, which is then reliant on people getting hold of that document, reading it, understanding it, taking it apart, then setting up the implementation plan, and that really doesn't fit the style of industry.' (Participant 1)*

A finding from Study 3 indicated that a statistically significant number of survey participants (67%) found it difficult to keep up to date with changes to aviation safety regulations/initiatives. A lack of guidance and support from CASA, and seemingly ineffectively maintained dialogue with stakeholders, were highlighted in Participant 5's comment:

> *'… there was no ongoing support material come from CASA and, um, whether there's been something done with threat and error management, I don't know. I have to go—and I wasn't even aware of it 'til you mentioned it that there is an updated one. So, I might have to jump online,*

> *see if I can find it, go through it to see what they've published on threat and error management.' (Participant 5)*

Conley et al. (2011) emphasised the importance of clearly communicating to relevant stakeholders *why* a certain safety initiative needs to be implemented, so that stronger enthusiasm and a high level of buy-in can be built and achieved. However, Participant 4's experience was somewhat different:

> *'… if we had actually said, 'guys, we've been actually doing this … what we actually want to do is to formalise this little bit' … Um, but it's, it was still marketed as something different and I think that immediately got people's back side because we've been flying already, best part of 100 years. Um, what are we doing today that we didn't used to do and who actually says there's a problem? … Otherwise a lot more of us would've been dead if we hadn't been applying threat and error management.' (Participant 4)*

This comment suggested that the elements of TEM were already present and practised, well before TEM was officially implemented. Participant 1 in Study 2 commented that:

> *'… if you ask me my version of what TEM is, it would probably be pretty much the same as some other people, probably experienced people. I am a great believer in it, but I am going back to some basics, which is, I think it's been around for a long time, we just call it differently these days. The industry generally would probably be sort of aware of it.' (Participant 1)*

Other participants in Study 2 shared a similar view:

> *'I would go as far as saying, whilst we didn't have a name for this in the past, before every flight, we still always did this. You still try to identify any possible hazards that—it was hazard back then—it might have been associated with flight from organic hazards, um, to, you know, machine to environment. We did this anyway, we didn't have a name for it. We didn't*

know we were actually doing it … Just didn't have a name for it.' (Participant 4)

'Threat and error management really … I didn't think it was a new thing, it's been around years, for years rather, ever since I've been involved in aviation. I think it was just that [with] TEM, someone finally defined it and put a name to it.' (Participant 5)

These comments highlight the importance of maintaining an effective ongoing dialogue with the general aviation industry, so that these views can be considered while planning and allocating the appropriate mix of required expertise and quantity of resources, which are the fourth and fifth, respectively, key considerations for successful implementation of a new policy (ANAO, 2014). Planning can be considered a map that describes certain features of an initiative, such as the reasons for the new initiative and the ways it will be resourced (ANAO, 2014). Participant 4 made the comment that:

'I think the first question should've been asked is: What is this? Is this actually something new or is this something we're already doing? … if it did already exist, what was actually wrong with the way it's been, what was it actual problem here?' (Participant 4)

If TEM had been clearly planned and communicated with the general aviation industry, it would have built stronger enthusiasm and support from the industry, which would have allowed CASA to achieve a high level of buy-in from Australian general aviation. Participant 4 further added that:

'… so the first thing we should've done would've been to have actually run a series of training sessions, practically focused training session, showing threat and error management in practice, um, teaching CFIs syllabus designs, how, and examiners, um, how, what is actually a threat, what is actually an error, what are plausible management strategies, um, help them to actually understand the system first of all, and then providing the guidance materials on how to actually roll it out into a school

> *environment. If you did it that way, you probably would've had a greater pick up and a greater success.' (Participant 4)*

The comments from the participants in Study 2 indicated that the implementation that occurred was different from what should ideally have been considered, planned and actioned. Participant 1 stated: *'TEM was a buzzword, it was pushed out there and it was required by CASA to be introduced to flight schools and nobody really had a clear understanding of what it was, its components or elements'*. Participant 4 stated: *'There's a lot of psychology involved in threat and error management and that's not something that a lot of people, um, find it easy to understanding'.* The findings from Study 3 supported these participants' comments. A statistically significant number of survey participants (63.9%) disagreed that CASA had provided adequate training and guidance material on TEM training for general aviation. In addition, a statistically significant number of survey participants (52.6%) disagreed that the concept of TEM was easy to understand. Further, a statistically significant number of survey participants (73.2%) disagreed that CASA had explained TEM and how it would affect their role and/or activities in a clear and succinct manner. This lack of guidance and support from CASA likely had a negative effect on how TEM was perceived to be implemented in Australian general aviation.

The findings above highlight the importance of establishing in-house expertise in TEM within CASA. However, as discussed earlier, participant 3 in study 2 recalled that even a senior manager in CASA did not have a clear understanding of TEM. Laszlo (1998) emphasised the need for adequate allocation of resources, with a top priority being training key personnel on the project to be implemented, followed by cascading training down, based on organisational structure. In general aviation, these key personnel include, yet are not limited to, FOIs and ASAs. Participant 4's comment did not support the idea of key personnel receiving adequate training: *'To this day, I would be fairly confident in saying that you could ask 15 inspectors at CASA and you'll have 15 different opinions on threat and error management and their understanding'*. If this is deemed to be an accurate assessment, it is very likely that the ASAs who typically conducted AvSafety seminars would not have been suitably trained. A finding from Study 3 supported this notion, with only 17 survey participants

(17.5%) agreeing that the CASA's ASAs had a consistent understanding of TEM principles.

With regard to CASA-organised seminars, such as AvSafety, Participant 3 shared his experience with the seminar on TEM:

> *'CASA had run courses for private pilots where they could talk about those things, but I think I only attended one and mostly they were dealing with questions and they don't—there was no overall philosophy, except they've got to cover and demonstrate that you're considering threats and how you manage them.' (Participant 3)*

Participant 4 explained that:

> *'They have touched on TEM in seminars … but again they were done by aviation safety advisors who are not threat and error management specialists. They are safety specialists. Um, they're from all walks from life, uh, they are regurgitating materials that wasn't even necessarily developed by the department.' (Participant 4)*

These comments highlight the importance of including experts in the delivery of safety seminars. Hattie (2003) argued that, although the amount of knowledge between experts and experienced teachers does not differ much, experts possess knowledge that is more integrated by combining and relating new knowledge with prior knowledge. This allows experts to be able to plan lessons based on students' needs by changing, combining and adding content accordingly (Hattie, 2003).

In the case of TEM implementation in Australian general aviation, it was already established that the number of appropriately qualified TEM experts in CASA before and after TEM implementation was significantly lacking, and FOIs and ASAs were not adequately trained on TEM. This would likely have resulted in no specific TEM training sessions through PDP for FEs, and the collective views among the participants in all three studies indicated a lack of guidance and support from CASA. For instance, a statistically significant number of survey participants from Study 3 responded that

CASA did not provide adequate guidance and support when TEM was implemented in Australian general aviation. In addition, only 36% of the survey participants considered currently published safety publications and resources from CASA to be useful.

The last key building block for successful implementation is the establishment of ongoing active management that comprises well-coordinated monitoring, review and evaluation processes (ANAO, 2014). Ongoing active management aims to appropriately inform and facilitate other relevant key building blocks for their decision-making and support activities to enable successful implementation of an initiative (ANAO, 2014). This is consistent with one of the commonly used evaluation approaches that consists of four levels of evaluation: (i) participants' reaction to a newly implemented program or training initiatives (reaction), (ii) changes in participants' attitude and improved skills and knowledge (learning), (iii) changes in behaviour (behaviour) and (iv) achievement of the predetermined final objectives of a program (Kirkpatrick & Kirkpatrick, 2006). Participant 1 described the fourth and final level as a:

> *'… particularly commercial [CPL] level. We could say, 'okay, schools, you have a reasonable amount of time to get students prepared for CPL and to instil the principles of TEM into the way they operate'. It would be good to do some tests at CPL level by the independent examiners to get some feedback on how TEM is working and see what the standard is like.' (Participant 1)*

Participant 1 continued:

> *'Their [CASA's] annual or 18-month roadshow would go around and saying, 'we've done some appropriate sampling around the industry and the feedback is that there were still some weak areas—this, this and this*

> *... this is how we recommend to fix them'. It is the normal close loop.' (Participant 1)*

Based on an extensive literature search, it was noted that there was very limited, if any, formal post-implementation review of TEM in Australian general aviation. Participant 1's suggestions would certainly provide better insights regarding how each flight school performs in term of the way TEM principles are taught and practised, and the findings could be disseminated to other segments of general aviation, such as private flying. However, CASA's auditing role and the way the audits were conducted were noted as a perceived barrier. The findings from a recent stakeholder survey (Colmar Brunton, 2016) suggested that participants in the private flying sector were particularly dissatisfied with CASA's overly punitive approaches, rather than working with stakeholders in a more constructive manner to resolve issues and ultimately improve safety when performing its audit and compliance functions.

The review of findings from the three studies, based on the six key considerations for successful implementation of TEM, suggested that TEM implementation in Australian general aviation is perceived by many as laissez-faire. It was evident from the findings that there was lack of guidance and support from CASA before and after TEM was implemented. This was largely because of CASA failing to establish a team of expertise in TEM within CASA, and suboptimal stakeholder engagement during the planning stage. In addition, there has been very limited formal review, assessment or evaluation to examine the situation after the introduction of TEM in Australian general aviation to determine how successfully it was implemented and which improvements were required.

It was evident that TEM was poorly implemented, which would likely have led to negative views on the use of TEM in Australian general aviation. However, the collective responses from Study 3 signalled that the survey participants highly valued and used TEM principles when undertaking flying activities, and consequently considered TEM an important part of flight preparations. The following section presents further discussion on how TEM is viewed, taught, practised and examined.

7.2 Threat and Error Management in Practice

The majority of findings indicated that TEM was poorly implemented in Australian general aviation, and the impracticality of TEM and lack of guidance and support from CASA were suggested to be the main contributing factors to these collective views. These factors likely led to negative views among the survey participants on how TEM is considered and practised in Australian general aviation. However, the responses to questions on the way TEM was viewed and practised were somewhat different. The general consensus was that the survey participants in Study 3 highly valued TEM when undertaking flying activities. Consequently, the survey participants considered TEM an important part of their flight preparation and used TEM principles when undertaking flying activities. This interesting and encouraging finding suggested that, despite poor TEM implementation, TEM was valued and practised, to varying degrees, in Australian general aviation.

The participants in Study 1 (six trainees and eight trainers) who were engaged in flying activities before the mandating of TEM training confirmed that briefings and/or training syllabi were amended to incorporate TEM training. The majority of trainees and trainers in Study 1 also indicated that TEM training was incorporated in pre-/post-flight briefings, ground schools and instructional flights. Participant 4 in Study 2 described the way TEM was incorporated in his organisation:

> *'... so over a period of years, we've had, uh, development of the threat and error management material within our long briefings, within our ground school materials for PPL and CPL, um. We've probably actually completely changed the way we approached threat and error management, um, at a sequence long briefing level as well.' (Participant 4)*

Participant 5's experience with incorporation of TEM in his organisation was similar:

> *'Yes, I was, uh, involved in that, uh, CASA gave us a training package in that and we went through it, we had staff leading, all went through*

together, then I implemented some training measures, some study materials for the candidates and, uh, implemented changes to our briefing processes, to trying, uh, implement threat and error management as best as we can understand it.' (Participant 5)

Despite organisations' endeavours to implement TEM, a statistically significant number of survey participants in Study 3 had difficulty adequately understanding the TEM concept. This might be due to the collective views among the majority of survey participants in Study 3 that CASA failed to appropriately translate TEM concepts into practical guidance, and offered inadequate practical examples relating to TEM principles and concepts for general aviation. In addition, Participant 1 in Study 2 indicated that the degree of TEM implementation varied depending on the geographical locations of flight schools:

'So, I guess implementation in flight schools sort of varied from small ones—small country ones, who probably didn't have and still haven't got a clue about TEM, probably didn't believe in it and therefore really did not implement it at all—to the major flying schools, like Flight Training Adelaide—an organisation that, yeah, would've embraced TEM. They would've had a clear understanding about what it was, and uh, would've implemented it into their procedures. So clearly it depends on which part of the industry you are looking at.' (Participant 1)

Participant 1 highlighted the importance of implementing an appropriate mix of required expertise and quantity of resources when introducing a new safety initiative, such as TEM:

'Places like Flight Training Adelaide, CAE Oxford get feedback from airline clients and/or their parent company. CAE Oxford down at Moorabbin, they used to be one of the QANTAS providers, previous named General Flying Services … They would've drawn a lot of their expertise on TEM down here from their parent company in UK and Europe because EASA brought TEM in a number of years before we did, so, you are right. They had resources and they have contracts in the big end of town to bring

> *the stuff in and probably did it quite well. The lower end, they don't … that's going to take a bit of time to filter down.' (Participant 1)*

Participant 4 concurred with Participant 1's observation and added that:

> *'There're massive differences what you see there, but there's very little difference between what I have experienced in flight with applicants from both ends of spectrum. I think everybody has always done a fairly good job of in-flight training students to identify threats and errors.' (Participant 4)*

A statistically significant number of survey participants indicated that TEM principles were valued and practised when undertaking flying activities. This may have been contributed to by the inclusion of TEM considerations in pre-flight briefings, which usually comprise 15- to 20-minute one-on-one discussions on the aim of the flight and lesson sequences. Participant 4 commented that:

> *'Threat and error management better lives in a pre-flight briefing, where we are able to actually consider environmental factors of the day and the two organic beings on board the aircraft and the actual condition of the aircraft in which we are flying on that day—none of which is known at a long briefing stage. So, I prefer that it will create a discussion, threat and error management in a pre-flight briefing stage, despite the fact that, um, you know, in our long briefing format, there is a requirement for TEM.' (Participant 4)*

Participant 5 added that:

> *'… we modified all our briefings and, um, we just put TEM as an addition to the briefings, okay, so the actual pre-flight briefing didn't change, but we incorporated TEM at the end of the briefings … now, the last thing before you walk out the door, let's see what can turn to worms.' (Participant 5)*

The majority of participants in Study 2 commented that certain elements and principles of TEM already existed in Australian general aviation well before TEM was formally introduced. Participant 1 commented that: *'The elements of TEM, threat and error management, okay, it has been around for a long, long time, certainly back at World War II or probably before that'.* Participant 3 shared the similar view that:

> *'… when it was taught, you know, we talk more generally as airmanship, I see good instructors taught the students what to be aware of and how to think when it wasn't formally threat and error management as a subject.' (Participant 3)*

Similarly, Participant 4 commented that:

> *'I would go as far as saying, whilst we didn't have a name for this in the past, before every flight, we still always did this, you still try to identify any possible hazards that—it was hazard back then—it might have been associated with flight from organic hazards, um, to, you know, machine to environment. We did this anyway, we didn't have a name for it. We didn't know we were actually doing it.' (Participant 4)*

Participant 5 also added:

> *'Threat and error management really … I didn't think it was a new thing, it's been around years, for years rather, ever since I've been involved in aviation. I think it was just that [with] TEM, someone finally defined it and put a name to it.' (Participant 5)*

These collective observations highlighted the importance of careful planning through close and inclusive stakeholder engagement. However, as discussed in the previous section, this did not appear to have been well established during the planning stage of TEM implementation. Participant 4 commented that:

'… the first question should've been asked is: What is that? Is this actually something new or is this something we're already doing? … I think it already exists. If it did already exist, what was actually wrong with the way it's been, what was it actually problem here?' (Participant 4)

This lack of engagement may have resulted in some resistance from FIs after TEM was mandated. Participant 4 recalled that:

'We probably had a fair bit of resistance at that time because we had a lot of older generation instructors, uh, who I guess, had an opinion that, well, I am a pilot, I look like a pilot, I've been training pilots since way this long, I haven't had any problems yet, um … Otherwise, a lot more of us would've been dead if we hadn't been applying threat and error management prior to practical standards.' (Participant 4)

However, the findings from Study 3 suggested that the FI group—compared with the non-FI group—were using TEM principles more when undertaking flying activities, and had a clear understanding of what was required when applying and teaching TEM principles. Similarly, the FI group considered TEM an important part of their flight preparation.

Based on the findings from the series of studies in this thesis, it was identified that there was lack of guidance and support for TEM from CASA, and this might have contributed to the collective view of poor TEM implementation in Australian general aviation. However, an encouraging finding was that organisations endeavoured to incorporate TEM principles in their flying operations, and TEM principles were valued and viewed positively within general aviation.

7.3 Two Additional Questions

Two additional questions were included in Study 3. The first additional question asked whether the benefits of TEM training were overrated, and was included because, when this question was asked in Study 1, the results showed mixed views on the perceived benefits of TEM training, and varied according to groups (Lee et al., 2016). The responses from the survey participants in Study 3 also indicated mixed views on the benefits of TEM training, with the non-FI group indicating more agreement that the benefits of TEM were overrated. This may be because the non-FI group, compared with the FI group, did not have a clear understanding of what was required when applying TEM principles, which led them to place less importance on TEM in their flight preparation.

The second additional question considered the feasibility of expanding TEM training and principles to other sectors, such as the recreational aviation and RPAS sectors. The responses among survey participants in Study 3 were somewhat mixed. Participant 1's comment sheds some light on the views of the survey participants who agreed that TEM training and principles should be expanded to other sectors:

> *'Well, my immediate reaction to that would be: why they would be any different? I mean, they are operating in the same environment, uh, essentially general aviation, the threats are by and large the same, errors that people are going to make and management of those … same principles are applied.' (Participant 1)*

In contrast, Participant 3's comment supports the survey participants who did not agree or were neutral about expanding TEM training and principles to other sectors:

> *'I would … my reaction would be to say, yes, it would be helpful everywhere, but I would say that the caution to that is, if it's done the way it's been done under my observation in aviation, general aviation, it would just be a distraction.' (Participant 3)*

Participant 3's comment highlights the importance of careful and thorough planning, alongside other key considerations, such as engaging stakeholders and resources, to ensure successful implementation.

7.4 Potential Limitations of the Present Study

The main limitation of study 3 is a sample size for Structural Equation Modelling (SEM). Although the total number of participants in study 3 exceeded the recommended minimum sample size of 91 it was achieved by increasing anticipated effect size to 0.5 while keeping the statistical significance (*p* value) of 0.05 (Figure 6.35). The free online a-priori sample size calculator for SEM (https://www.danielsoper.com/statcalc/calculator.aspx?id=89) offers three options: 0.1 for a small effect; 0.3 for a medium effect and 0.5 for a large effect.

Sullivan and Feinn (2012) described that with a large effect (i.e., 0.5) the likelihood of committing a Type II error would increase. A Type II error occurs when research fails to reject a null hypothesis (Field, 2009). Another word, A Type II error leads researchers to falsely believe that there is no effect even though there is an effect. This will also likely lead to lowering the statistical power, a chance of detecting an effect if there is one exists (Field, 2009; Sullivan & Feinn, 2012). Therefore, with the large effect selected in study 3 there is a probability the researcher may have failed to detect statistically significant effects when interpreting results after completing SEM.

7.5 Where to from here

There is a sufficient number of research (e.g., Bazargan & Guzhva, 2011; Hunter, Martinussen, Wiggins, & O'Hare, 2011; Salter & Alexander, 2000) and statistics (e.g., ATSB, 2018; NTSB, 2014) that have identified suboptimal safety performance of general aviation. These are mainly due to less rigorously trained pilots with a lack of appropriate, formal human factors training offered within general aviation. It was in this notion that TEM training was introduced in Australia to improve general aviation safety. However, the three studies in this thesis identified that TEM was poorly implemented. This was mainly due to the perceived impracticality of TEM and a lack of guidance and support from CASA. Therefore, it is considered to be an important and logical step to conduct a nationwide survey with a greater number of participants to identify areas that require immediate attention. These areas may include, but not limited to, how the requirements of TEM are taught and practised in different organisations such as flight schools, aeroclubs and small charter operators and the need for practical guidance material on TEM. Comprehensively developed training sessions specifically targeting flight examiners, flight instructors, Flying Operations Inspectors (FOI) and Aviation Safety Advisors (ASA) on how to teach TEM principles and their practical applications will be valuable for the benefits of TEM training to be cascaded to general aviation pilots. A specifically tasked group may need to be established within CASA to be responsible for this purpose. The survey data and findings can also be supplemented if participants' viewpoints are also collected through interview, focus group or other similar means so that CASA is better informed as to what remedial actions should be undertaken to improve underperforming areas and how these actions are framed to best fit into appropriate organisations.

Findings from this thesis also suggested that it would be beneficial for TEM to be implemented in other sectors within Australian general aviation such as recreational aviation and Remotely Piloted Aircraft System (RPAS) sectors. Careful planning such as a nationwide consultation with key stakeholders and other important considerations listed in the guidance by the Australian National Audit Office ([ANAO], 2014) need to be followed. Also, findings from this thesis, from the nationwide survey suggested above and from other research (e.g., Conley et al., 2011) should be reflected before, during and after TEM implementation so that intended outcomes are achieved.

7.6 Chapter Summary

The findings from the three studies in this thesis suggested that TEM was poorly implemented in Australian general aviation. The findings also suggested that two major barriers—impracticality and lack of guidance and support from CASA—were the main contributors for the collective views among the participants on poor TEM implementation. It was reasonable to assume that these views would likely lead to negative views on the use of TEM in Australian general aviation. However, the collective responses from Study 3 suggested that the survey participants highly valued and used TEM principles when undertaking flying activities, and consequently considered TEM an important part of flight preparations. This is an encouraging finding because considering TEM principles and concepts prior to and during flying will further improve safety in Australian general aviation.

Chapter 8: Conclusion

Civil aviation is considered an ultra-safe industry and its safety record far exceeds other high-reliability industries, such as the healthcare industry. This high level of safety is attributed to a constant drive to develop and improve safety initiatives and training. One of the recent safety initiatives was the introduction of Threat and Error Management (TEM). Since the emergence of TEM, it has been well regarded as an effective method to improve aviation safety. Consequently, ICAO endorsed and recommended that TEM training be an integral part of pilot training and a requirement for all pilot licences. In response to ICAO's acknowledgement and recommendation of the need for TEM training, CASA mandated TEM as an additional assessment item for various levels of flight tests and ground examinations, effective from July 2009.

The findings from the three studies in this research suggested that two major barriers—impracticality and lack of guidance and support from CASA—were the main contributors to the collective views among the study participants regarding poor TEM implementation. General aviation in Australia is primarily a vocationally based industry; thus, it is important to consider practical aspects when new safety initiatives are proposed and introduced in general aviation, such as TEM implementation. CASA's failure to provide practical guidance and examples of how, why and when TEM principles and concepts should be adopted and practised in daily operations was due to CASA's perceived lack of understanding of how general aviation operates. Although CASA offered a range of seminars and workshops, such as AvSafety, to assist general aviation pilots to stay up to date and to promote aviation safety, there was a lack of systematic and well-planned educational roadshows to promote TEM. This necessary educational opportunity was not offered to FEs, despite TEM becoming an additional testing item. This resulted in general consensus that TEM was not assessed in a standardised manner.

The Australian National Audit Office ([ANAO], 2014) has highlighted six key considerations when implementing a new policy initiative, and the findings from the current research suggested that these key considerations were not adequately followed during TEM implementation. For instance, sound governance arrangements and practices were not found to be well established, and an appropriate level of stakeholder engagement was not maintained throughout the implementation process. These failures resulted in a perceived lack of relevance of TEM among the intended recipients, and a subsequent lack of buy-in, which likely led to the manifestation of poor TEM implementation in Australian general aviation.

It is evident that TEM was poorly implemented in Australian general aviation, and this would likely have led to negative views on the use of TEM principles and concepts. However, the findings suggested that the participants highly valued TEM when undertaking flying activities. Consequently, the participants considered TEM an important part of their flight preparation and used TEM principles when undertaking flying activities. This interesting, yet perhaps counterintuitive, finding suggested that, despite poor TEM implementation, TEM was valued and practised, to varying degrees, in Australian general aviation. In addition, it is understood that concepts similar to TEM have existed in general aviation for a long time; thus, although these principles and concepts were not labelled TEM, they have nevertheless been taught and exercised.

In conclusion, there are clear indications that the effects of TEM implementation in Australian general aviation have not been experienced and translated as intended. The perceived impracticality of TEM and a lack of guidance and support were found to be the two major obstacles that adversely affected the successful implementation of TEM. However, despite the poor TEM implementation, it was found that TEM principles were highly valued among the study participants, and were considered an important part of flight preparation and subsequently exercised when undertaking flying activities. This is an encouraging finding given that the ultimate aim of TEM implementation is to further improve safety in Australian general aviation.

References:

Adams, D. (2006). A layman's introduction to human factors in aircraft accident and incident investigation, B2006/0094. Canberra, Australia: Australian Transport Safety Bureau.

Amalberti, R., Auroy, Y., Berwick, D., & Barach, P. (2005). Five system barriers to achieving ultrasafe health care. Annals of Internal Medicine, 142(9), 756-764. Doi: 10.7326/0003-4819-142-9-200505030-00012

Alderson, J. C. (2009). Air safety, language assessment policy, and policy implementation: The case of aviation English, *Annual Review of Applied Linguistics, 29*, 168-187. Doi: 10.1017/S0267190509090138

Australian National Audit Office (2014). Successful implementation of policy initiatives: Better practice guide. Canberra, Australia: Commonwealth of Australia.

Australian Transport Safety Bureau (2018). Aviation occurrence statistics: 2007 – 2016, AR-2017-104. Canberra, Australia: Author.

Australian Transport Safety Bureau (2107). Aviation occurrence statistics: 2006 – 2015, AR-2016-122. Canberra, Australia: Author.

Australian Transport Safety Bureau (2103a). Collison with terrain involving Cessna 172N, VH-JGR (AO-2012-148). Canberra, Australia: Author.

Australian Transport Safety Bureau (2013b). Engine failure involving Gippsland aeronautics GA-8, VH-AJS (AO-2011-125). Canberra, Australia: Author.

Australian Transport Safety Bureau (2012). Hard landing involving Cessna P 206B, VH-EGG (AO-2012-135). Canberra, Australia: Author.

Australian Transport Safety Bureau (2011). *Aviation occurrence statistics 2001 to 2010.* Canberra, Australia: Author.

Bartsch, R. I. C. (2013). *Aviation law in Australia* (4th ed.). NSW, Australia: Thomson Reuters.

Bazagan, M., & Guzhva, V. S. (2011). Impact of gender, age and experience of pilots on general aviation accidents. *Accident Analysis and Prevention, 43*, 962-970. Doi: 10.1016/j.aap.2010.11.023.

Bazeley, P., & Jackson, K. (2013). *Qualitative data analysis with Nvivo* (2nd ed.). Thousand Oaks, CA: Sage.

Bove, T. (2002). Development and validation of a human error management taxonomy in air traffic control (Riso-R-1378[EN]), Roskilde, Denmark: Roso National Laboratory.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2). 77-101. Doi: 10.1191/1478088706qp063oa

Brinkmann, S. (2017). The interview. In N. K. Denzin & Y. S. Lincoln (eds.), *The SAGE handbook of quality research* (5th ed), (pp. 576-599). Thousand Oaks, CA: Sage.

Brinkmann, S., & Kvale, S. (2015). *Interviews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: Sage.

Broadsky, A. E. (2008). Negative case analysis. In L. M. Given (ed.), *The SAGE encyclopedia of qualitative research methods* (volume 2), (p. 552). Thousand Oaks, CA: Sage.

Bryman, A. (2006). Integrating quantitative and qualitative research: how it is done?. *Qualitative Research 6*(1), 97-113. Doi: 10.1177/1468794106058877

Byrne, B. M. (2016). Structural equation modeling with AMOS: Basic concepts, applications and programming (3rd ed.). New York, NY: Routledge.

Catchpole, K. R., Giddings, A. E. B., de Laval, M. R., Peek, G. J., Godden, P. J., Utley, M., Gallian, S., Hirst, G., & Dale, T. (2006). Identification of systems failures in successful paediatric cardiac surgery. *Ergonomics, 49*(5), 567-588. Doi: 10.1080/00140130600568865

Cheng, K., Inglis, M., & Godley, S. T. (2009). Threat and error management: Attitudes towards training and applicability of TEM to general aviation and low capacity air transport operations (AR-2006-156). Canberra, Australia: Australian Transport Safety Bureau.

Civil Aviation Safety Authority. (2018a). About us. Retrieved from https://www.casa.gov.au/landing-page/about-us

Civil Aviation Safety Authority. (2018b). Purpose. Retrieved from https://www.casa.gov.au/book-page/purpose

Civil Aviation Safety Authority. (2018c). Purpose. Retrieved from https://www.casa.gov.au/book-page/purpose.

Civil Aviation Safety Authority. (2018d). Our regulatory philosophy. Retrieved from https://www.casa.gov.au/about-us/standard-page/our-regulatory-philosophy

Civil Aviation Safety Authority. (2018e). Part 61 manual of standards instrument 2014. Retrieved from https://www.legislation.gov.au/Details/F2018C00657

Civil Aviation Safety Authority. (2018f). The Part 61 manual of standards. Retrieved from https://www.casa.gov.au/standard-page/part-61-manual-standards

Civil Aviation Safety Authority. (2017). Flight examiner handbook (D17/151743). Canberra, Australia: Author.

Civil Aviation Safety Authority. (2010). Annual report 2009-2010. Canberra, Australia: Author.

Civil Aviation Safety Authority. (2008). Teaching and assessing single-pilot human factors and threat and error management (CAAP 5.59-1). Canberra, Australia: Author.

Colmar Brunton. (2016). Civil Aviation Safety Authority: Stakeholder relationship health. QLD, Australia: Colmar Brunton Social Research.

Conley, D. M., Singer, S. J., Edmondson, L., Berry, W. R., & Gawande, A. A. (2011). Effective surgical safety checklist implementation, Journal of the American College of Surgeons, 212(5), 873-879. DOI: 10.1016/j.jamcollsurg.2011.01.052

Creswell, J. W. (2014). Research design: Qualitative, quantitative and mixed methods approaches (4th ed). Thousand Oaks, CA: Sage.

Creswell, J. W. (2013). Qualitative inquiry and research design: Choosing among five approaches. Thousand Oaks, CA: Sage.

Creswell, J. W. & Plano Clark, V. L. (2011). Designing and conducting mixed methods research (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry, *Theory into Practice, 39*(3), 124-130. DOI: 10.1207/s15430421tip3903_2

Crotty, M. (1998). The foundation of social research: Meaning and perspective in the research process. NSW, Australia: Allen & Unwin.

Davidson, J., & di Gregorio, S. (2011). Qualitative research and technology: In the midst of a revolution. In N. K. Denzin & Y. S. Lincoln (eds.), *The SAGE handbook of qualitative research* (4th ed), (pp. 627-643). Thousand Oaks, CA: Sage.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research, 50*(1), 61-77. Doi: 10.1177/147078530805000106

Dekker, S. (2006). *The field guide to understanding human error*. Surrey, England: Ashgate.

Eames-Brown, C., & Collis, G. (2007). Regional Airline Line Operations Safety Audit. ATSB transport safety report (B2004/0237). Canberra, Australia: Australian Transport Safety Bureau.

Earl, L., Bates, P. R., Murray, P. S., Glendon, I., & Creed, P. A. (2012). Developing a single-pilot line operations safety audit. Aviation Psychology and Applied Human Factors, 2, 49-61. Doi: 10.1027/2192-0923/a000027.

Evers, J. C., Silver, C., Mruck, K., & Peeters, B. (2011). Introduction to the KWALON experiment: Discussions on qualitative data analysis software by developers and users. *Forum: Qualitative Social Research 12*(1), Art. 40.

Federal Aviation Administration (1991). *Advanced qualification program* (Advisory Circular 120-54). Washington, DC: Author.

Feggetter, A. J. (1985). Human factors: The key to survival. In R. Hurst & L. R. Hurst (Eds.), *Safety in General Aviation* (pp. 1-66). London, UK: Collins.

Ferroff, C. V., Mavin, T. J., Bates, P. R., & Murray, P. S. (2012). A case for social constructionism in aviation safety and human performance research. *Aeronautica, 3*, 1-12.

Field, A. (2013). Discovering statistics using IBM SPSS statistics (4th ed.), Thousand Oaks, CA: Sage.

Fisher, J., Phillips, E., & Mather, J. (2000). Does crew resource management training work? *Air Medical Journal, 19*(4), 137-139. Doi: 10.1016/S1067-991X(00)90006-3

Flin, R., O'Connor, P., & Crichton, M. (2008). Safety at the sharp end: A guide to non-technical skills. Surrey, UK: Ashgate.

Flin, R., & Martin, L. (2001). Behavioral markers for crew resource management: A review of current practice. *International Journal of Aviation Psychology, 11*(1), 95-118. Doi: 10.1207/S15327108IJAP1101_6

Ford, J., Henderson, R., & O'Hare, D. (2014). The effects of Crew Resource Management (CRM) training on flight attendants' safety attitudes. *Journal of Safety Research, 48*, 49-56.

Foster, K., Curtis, K., Mitchell, R., Van, C., & Young, A. (2016). The experiences, unmet needs and outcomes of parents of severely injured children: A longitudinal mixed methods study protocol. *BMC Pediatrics, 16*(1), 1-7. Doi: 10.1186/s12887-016-0693-8

Friday, M. A. (2011). A mixed methods study of designated pilot examiners experiences with advanced display technology and technology advanced aircraft (Unpublished doctoral thesis). Oklahoma State University, Stillwater, Oklahoma.

Gilbert, L. S., Jackson, K., & di Gregorio, S. (2014). Tools for analysing qualitative data: The history and relevance of qualitative data analysis software. In M. Spector, M. D. Merrill, J. Elen., & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 221-236). NY: Springer.

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Market Science, 23*(4), 519-529. Doi: 10.1287/mksc.1040.0070.

Gunther, D. (2002). Threat and error management training counters complacency, fosters operational excellence. *International Civil Aviation Organisation Journal, 57*(4), 12-13.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate data analysis (6^{th} ed.). Upper Saddle River, NJ: Pearson

Hartzler, A., McCarty, C. A., Rasmussen, L. V., Williams, M. S., Brilliant, M., Bowton, E. A., … Trinidad, S. B. (2013). Stakeholder engagement: a key component of integrating genomic information into electronic health records. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, *15*(10), 792–801. http://doi.org/10.1038/gim.2013.127.

Hatch, J. A. (2002). Doing qualitative research in education settings. Albany, NY: State University of New York.

Hattie, J. A. C. (2003). Teachers make a difference: What is the research evidence? Paper presented at the Building Teacher Quality: What does the research tell us ACER Research Conference, Melbourne, Australia.

Helmreich, R. L., & Foushee, H. C. (2010). Why CRM? Empirical and theoretical bases of human factors training. In B. Kanki, R.L. Helmreich & J. Anca (Eds.), *Crew Resource Management* (2^{nd} ed., pp. 3-57). San Diego, CA: Elsevier.

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *International Journal of Aviation Psychology, 9*(1), 19-32. Doi: 10.1207/s15327108ijap0901_2

Hobbs, A. (2004). Human factors: The last frontier of aviation safety? *The International Journal of Aviation Psychology*, *14*(4), 335-341. Doi: 10.1207/s15327108ijap1404_1

Holloway, I., & Todres, L. (2003). The status of method: Flexibility, consistency and coherence. *Qualitative Research, 3*(3), 345-357. Doi: 10.1177/1468794103033004

Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review, 11*, 354-373.

Hunter, D. R., Martinussen, M., Wiggins, M., & O'Hare, D. (2011). Situational and personal characteristics associated with adverse weather encountered by pilots. Accident Analysis and Prevention, 43, 176-186. Doi: 10.1016/j.aap.2010. 08.007.

Implementation. (2018). In Oxford English online dictionary. Retrieved from https://en.oxforddictionaries.com/definition/implementation

International Civil Aviation Organisation. (2013a). Annex 19 – Safety Management, Montreal, Canada: Author.

International Civil Aviation Organisation. (2013b). Safety Management Manual (SMM), Montreal, Canada: Author.

International Civil Aviation Organisation. (2006). Personnel Licensing (10th ed.), Montreal, Canada: Author.

International Civil Aviation Organisation. (2005a). *ICAO Accident Prevention Programme*, Montreal, Canada: Author.

International Civil Aviation Organisation. (2005b). Threat and Error Management (TEM) in air traffic control, Montreal, Canada: Author.

International Civil Aviation Organisation. (2004). *Manual on the implementation of ICAO language proficiency requirements*, Montreal, Canada: Author.

International Civil Aviation Organisation. (2002). Line operations safety audit (LOSA). (ICAO document 9803, AN/761). Montreal: Author.

Janghorban, R., Roudsari, R. L., & Taghipour, A. (2014). Skype interviewing: The new generation of online synchronous interview in qualitative research. *International Journal of Qualitative Studies on Health and Well-being, 9*(1). Doi: 10.3402/qhw.v9.24152

Johnny, S. (2009). The coding manual for qualitative researchers. Thousand Oaks, CA: SAGE.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7). 14-26. Doi: 10.3102/0013189X033007014.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133. Doi: 10.1177/1558689806298224

Kharchenko, V., Chynchenko, Y., & Raychev, S. (2007). Threat and error management in air traffic control. *Bichuk HAY, 3*(4), 24-29. Doi: 10.18372/2306-1472.33.1491

Kim, S. K., & Byun, S. N. (2011). Effects of crew resource management training on the team performance of operators in an advanced nuclear power plant. *Journal of Nuclear Science and Technology, 48*(9), 1256-1264. Doi: 10.1080/18811248.2011.9711814

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). Evaluating training programs: The four levels (3rd ed.), San Francisco, CA: Berrett-Koehler.

Kirkpatrick, D. L. (1994). Evaluating training programs: The four levels (1st ed.), San Francisco, CA: Berrett-Koehler.

Klinect, J. R. (2016). Line Operations Safety Audit (LOSA). Paper presented at BASS 2016, Austin, Texas.

Klinect, J. R. (2005). Line Operations Safety Audit: A cockpit observation methodology for monitoring commercial airline safety performance (unpublished doctoral thesis). The University of Texas, Austin.

Klinect, J. R., Murray, P., Merritt, A., & Helmreich, R. (2003). Line operations safety audit (LOSA): Defining and operating characteristics. In Proceedings of the 12th International Symposium on Aviation Psychology (pp. 663-668). Daytona, OH: The Ohio State University.

Koeppen, N. A. (2012). The influence of automation on aviation accident and fatality rates: 2000-2010, retrived from http://commons.erau.edu/publications/95.

Koglbauer, I., Kallus, K. W., Braunstingl, R., & Boucsein, W. (2011). Recovery training in simulator improves performance and psychophysiological state of pilots during simulated and real visual flight rules flight. *The International Journal of Aviation Psychology, 21*(4), 307-324. Doi: 10.1080/10508414.2011.606741

Koonce, J. M. (2002). *Human factors in the training of pilots*. London, UK: Taylor & Francis.

Koonce, J. M. (1999). A historical overview of human factors in aviation. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), Handbook of aviation human factors (pp. 3-13). Mahwah, NJ: Lawrence Erlbaum.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy, 45*(3), 214-222. Doi: 10.5014/ajot.45.3.214

Kvalheim, S. A., & Haugen, S. (2014). Lessons from aviation: Applying threat and error management framework to cases from the offshore drilling industry. *Proceedings of the European Safety and Reliability Conference* (pp. 383-392). Amsterdam, the Netherlands: CRC

Lee, C-J. G., (2012). Reconsidering constructivism in qualitative research. *Educational Philosophy and Theory, 44*(2), 403-412. Doi: 10.1111/j.1469-5812.2010.00720.x.

Lee, R. M. (2008). Emory Bogardus and the new social research. *Current Sociology, 56*(2), 307-321. Doi: 10.1177/0011392107085037.

Lee, S. Y., Bates, P. R., Murray, P. S., & Martin, W. L. (2017). Training flight accidents: An explorative analysis of influencing factors and accident severity. *Aviation Psychology and Applied Human Factors, 7*(2), 107-113.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE.

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions and emerging confluences revisited. In N. K. Denzin & Y. S. Lincoln. *The SAGE handbook of qualitative research (4th ed.,)*. Thousand Oaks, CA: SAGE.

Maccoby, E. E., & Maccoby, N. (1954). The interview: A tool of social science. In G. Lindzey (ed.), *Handbook of social psychology* (pp. 449-487). Cambridge, MA: Addison-Wesley.

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. All Ireland Journal of Teaching and Learning in Higher Education. 9(3), 335(1)-335(14).

Martin, W. L. (2013). Pathological behaviours in pilots during unexpected critical events: The effects of startle, freeze and denial on situation outcome (Unpublished doctoral thesis). Griffith University, Brisbane, Australia.

Maurino, D. E. (2005a). Threat and Error Management (TEM). Paper presented at Canadian Aviation Safety Seminar (CASS), Vancouver, BC, 18-20 April.

Maurino, D. E. (2005b). TEM & LOSA: The state of affairs. Paper presented at 3rd IATA/ICAO LOSA/TEM Conference, Kuala Lumpur, Malaysia, 13-14 September.

Maurino, D. E., & Murray, P. S. (2009). Crew resource management. In J. A. Wise, D. Hopkins, & D. J. Garland (Eds.), Handbook of aviation human factors (2nd ed.). Boca Raton, FL: CRC.

McFadden, K. L., Stock, G. N., & Gowen, C. R. (2006). Implementation of patient safety initiatives in US hospitals. *International Journal of Operations and Production Management, 26*(3), 326-347.

Mearns, K. J., & Reader, T. (2008). Organisational support and safety outcomes: An un-investigated relationship?. *Safety Science, 46*(3), 388-397. Doi: 10.1016/j.ssci.2007.05.002.

Merritt, A., & Klinect, J. (2006). Defensive flying for pilots: An introduction to Threat and Error Management. *The University of Texas Human Factors Research Project: The LOSA Collaborative.* Austin, Texas.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). Thousand Oaks, CA: Sage.

Morris, A. (2015). *A practical introduction to in-depth interviewing*. London, UK: Sage.

Murray, S. R. (1997). Deliberate decision making by aircraft pilots: A simple reminder to avoid decision making under panic. *The International Journal of Aviation Psychology, 7*(1), 83-100. Doi: 10.1207/s15327108ijap0701_5

National Transportation Safety Board (2014). *Review of U.S. civil aviation accidents: Calendar year 2011 (NTSB/ARA-14/01)*. Washington, DC: Author.

Neuman, W. L. (2011). Social research methods: Qualitative and quantitative approaches (7th ed.). Boston: Pearson.

O'Connor, P., O'Dea, A., & Keogh, I. (2013). Is crew resource management training effective in healthcare settings? *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care. 2*(1), 105-107.

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. A. (2008). Crew resource management training effectiveness: A meta-analysis and some critical needs. *The International Journal of Aviation Psychology, 18*(4), 353-368. Doi: 10.1080/10508410802347044

Oxford University Press (2017). English: Oxford living dictionaries. Retrieved from https://en.oxforddictionaries.com/definition/ethics

Parker, I. (2005). Qualitative psychology: Introducing radical research. Berkshire, UK: Open University Press.

Patton, M. Q. (1987). How to use qualitative methods in evaluation. Newbury Park, CA: SAGE.

Pink, B (2012). 2012 Year book Australia (No. 1301.0). Canberra, Australia: Australian Bureau of Statistics.

Reason, J. T. (2000). Human error: Models and management. *British Medical Journal, 320*. 768-770.

Revilla, M. A., Saris, W. E., & Krosnick, J. A. (2013). Choosing the number of categories in agree-disagree scales. *Sociological Methods and Research, 43*(1). 73-91. Doi: 10.1177/0049124113509605.

Rudland, J. R., & Rennie, S. C. The determination of the relevance of basic sciences learning objectives to clinical practice using a questionnaire survey. *Medical Education, 37*. 962-965. Doi: 10.1046/j.1365-2923.2003.01671.x.

Ruskin, K. J., Stiegler, M. P., Park, K., Guffey, P., Kurup, V., & Chidester, T. (2013). Threat and error management for anesthesiologiests: A predictive risk taxonomy. *Current Opinion in Anesthesiology, 26*(6). 707-713. Doi: 10.1097/ACO.0000000000000014

Salas, E., Wilson, K. A., Burke, C. S., & Wightman, D. C. (2006). Does crew resource management training work? An update, an extension, and some critical needs. Human Factors, 48(2). 392-412. Doi: 10.1518/001872006777724444

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does Crew Resource Management (CRM) training work? *Human Factors, 43*(4). 641-674. Doi: 10.1518/001872001775870386

Sarter, N. B., & Alexander, H. M. (2000). Error types and related error detection mechanisms in aviation domain: An analysis of aviation safety reporting system incident reports. *The International Journal of Aviation Psychology, 10*(2). 189-206. Doi: 10.1207/S15327108IJAP1002_5

Seale, C. (2008). Using computers to analyse qualitative data. In D. Silverman & A. Marvasti. Doing qualitative research: A comprehensive quide. Thousand Oaks, CA: SAGE.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2). 63-75. Doi: 10.3233/EFI-2004-22201

St John, W., & Johnson, P. (2000). The pros and cons of data analysis software for qualitative research. *Journal of Nursing Scholarship, 32*(4), 393-397. Doi: 10.1111/j.1547-5069.2000.00393.x

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or Why the P value is not enough. Journal of Graduate Medical Education, 4(3). 279-282. Doi: 10.4300/JGME-D-12-00156.1.

Thomas, M. J. W. (2004). Error management training: Defining best practice (No. 2004/0500 – ATSB Aviation Safety Research Grant Scheme Project). Adelaide, SA: University of South Australia.

van der Schaaf, T. W. (1995). Human recovery of errors in man-machine systems. In: Proceedings of the 6th IFAC/IFIP/IFORS/IEA Symposium on the Analysis, Design and Evaluation of Man-Machine system, June 1995, Cambridge, MA.

Wakeman, D., & Langham, M. R., Jr. (2018). Creating a safer operating room: Groups, team dynamics and crew resource management principles. *Seminars in Pediatric Surgery, 27*(2). 107-113. Doi: 10.1053/j.sempedsurg.2018.02.008

Wiegmann, D., & Shappell, S. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Aldershot, UK: Ashgate.

Wiersma, W., & Jurs, S. G. (2009). *Research methods in education: An introduction (9th ed.)*. Boston, MA: Allyn and Bacon.

Wiggins, M. (2005). The interpretation and use of weather radar displays in aviation. MARCS Auditory Laboratories: University of Western Sydney.

Wilcock, A., Ball, B., & Fajumo, A. (2011). Effective implementation of food safety initiatives: Managers', food safety coordinators' and production workers' perspectives. *Food Control, 22*(1), 27-33. Doi: 10.1016/j.foodcont.2010.06.005.

Wittink, M. N., Barg, F. K., & Gallo, J. J. (2006). Unwritten rules of talking to doctors about depression: Integrating qualitative and quantitative methods. *Annals of Family Medicine, 4*(4), 302-309. Doi: 10.1370/afm.558

Zhang, W. (2009). Issues of practical teaching in vocational-technical schools in China and their countermeasures. *International Educaiton Studies, 2*(4), 75-78.

Appendix 1: Two types of threats with examples

Environmental threats	**Examples**
Adverse weather	Thunderstorms, turbulence, poor visibility, wind shear, icing conditions, IMC
Airport	Poor signage, faint markings, runway/taxiway closures, INOP navigational aids, poor braking action, contaminated runways/taxiways
ATC	Tough-to-meet clearances/restrictions, re-routes, language difficulties, controller errors
Environmental Operational Pressure	Terrain, traffic, TCAS, TA/RA, radio congestion
Airline threats	**Examples**
Aircraft	Systems, engines, flight controls, or automation anomalies or malfunctions; MEL items with operational implications; other aircraft threats requiring flight crew attention
Airline Operational Pressure	On-time performance pressure, delays, late arriving aircraft or flight crew
Cabin	Cabin events, flight attendant errors, distractions, interruptions
Dispatch/Paperwork	Load sheet errors, crew scheduling events, late paperwork, changes or errors
Ground/Ramp	Aircraft loading events, fuelling errors, agent interruptions, improper ground support, de-icing
Ground Maintenance	Aircraft repairs on ground, maintenance log problems, maintenance errors
Manuals/Charts	Missing information or documentation errors

Source: adapted from Merritt and Klinect (2006, p. 5)

Appendix 2: Three types of errors with examples

Aircraft Handling Errors	Examples
Automation	Incorrect altitude, speed, heading, auto-throttle setting, mode executed, or entries
Flight Control	Incorrect flaps, speed brake, auto-brake, thrust reverser or power settings
Ground Navigation	Attempting to turn down wrong taxiway/runway, Missed taxiway/runway/gate
Manual Flying	Hand flying vertical, lateral, or speed deviations, Missed runway/taxiway failure to hold short, or taxi above speed limit
Systems/Radios/Instruments	Incorrect pack, altimeter, fuel switch or radio frequency settings
Procedural Errors	**Examples**
Briefing	Missed items in the brief, omitted departure, take-off, approach, or handover briefing
Callouts	Omitted take-off, descent, or approach callouts
Cabin	Cabin events, flight attendant errors, distractions, interruptions
Checklist	Performed checklist from memory or omitted checklist, Missed items, wrong challenge and response, performed late or at wrong time
Documentation	Wrong weight and balance, fuel information, ATIS, or clearance recorded, Misinterpreted items on paperwork
Pilot Flying (PF) / Pilot Not Flying (PNF) Duty	PF makes own automation changes, PNF doing PF duties, PF doing PNF duties
SOP Cross-verification	Intentional and unintentional failure to cross-verify automation inputs
Other Procedural	Other deviations from government regulations, flight manual requirements or Standard Operating Procedures (SOP)
Communication Errors	**Examples**
Crew to External	Missed calls, misinterpretation of instructions, or incorrect read-backs to ATC, Wrong clearance, taxiway, gate or runway communicated
Pilot to Pilot	Within-crew miscommunication or misinterpretation

Source: adapted from Merritt and Klinect (2006, p. 9)

Appendix 3: Three types of UAS with examples

UAS Types	Examples
Aircraft Handling	Vertical, lateral or speed deviations; Unnecessary weather penetration; Unstable approach; Long, floated, firm or off-centreline landings
Ground Navigation	Runway/taxiway incursions; Wrong taxiway, ramp, gate, or hold spot; Taxi above speed limit
Incorrect Aircraft Configuration	Automation, engine, flight control, systems, or weight/balance events

Source: adapted from Merritt and Klinect (2006, p. 13)

Appendix 4: Ethics approval (Study 1)

GRIFFITH UNIVERSITY HUMAN RESEARCH ETHICS COMMITTEE

12-Mar-2013

Dear Mr Lee

I write further to your application for ethical clearance for your project "NR: What is the current address of Threat and Error Management training and how its effectiveness is perceived?" (GU Ref No: BPS/04/13/HREC). This project has been considered by Human expedited review 1.

The Chair resolved to grant this project ethical clearance.

This decision was made on 12-Feb-13. The ethical clearance for this protocol runs from 12-Feb-13 to 31-Jul-13.

Please refer to the attached sheet for the standard conditions of ethical clearance at Griffith University, as well as responses to questions commonly posed by researchers.

Regards

Ms Kristie Westerlaken
Policy Officer
Office for Research
Bray Centre, Nathan Campus
Griffith University
ph: +61 (0)7 373 58043
fax: +61 (07) 373 57994
email: k.westerlaken@griffith.edu.au
web:

Cc:

Researchers are reminded that the Griffith University Code for the Responsible Conduct of Research provides guidance to researchers in areas such as conflict of interest, authorship, storage of data, & the training of research students.
You can find further information, resources and a link to the University's Code by visiting
http://policies.griffith.edu.au/pdf/Code%20for%20the%20Responsible%20Conduct%20of%20Research.pdf

Appendix 5: Participant information sheet (Study 2)

University of Southern Queensland

Participant Information for USQ Research Project Interview

Project Details

Title of Project:	Exploring the effect of introduction of Threat and Error Management in Australian general aviation
Human Research Ethics Approval Number:	H16REA238

Research Team Contact Details

Principal investigator Details

Mr Paul Lee
Email: paul.lee@usq.edu.au
Telephone: (07) 3470 4559
Mobile: 0450 567 747

Principal supervisor Details

Professor Paul Bates
Email: paul.bates@usq.edu.au
Telephone: (07) 3470 4549

Associate supervisor Details

Professor Patrick Murray
Email: patrick.murray@usq.edu.au
Telephone: (07) 3470 4556

Associate supervisor Details

Dr. Wayne Martin
Email: wayne.martin@usq.edu.au
Telephone: (07) 3470 4558

Description

Following the ICAO introduction of TEM into the various Annexes to the ICAO convention more than five years ago, CASA mandated TEM training in Australia with the aim of improving aviation safety. However, to date, there has been very limited formal post-implementation review, assessment or evaluation (e.g., ATSB Transport Safety Report) to look at the 'after-state' in terms of the effectiveness of the introduction of TEM. Also, an exploratory study, comprising approximately 50 general aviation pilots, conducted by the same research team, indicated a variable uptake of TEM principles and differing opinions as to its effectiveness. Therefore it is considered timely to examine how the requirement for TEM training is currently being addressed. Areas of importance include, but not limited to, how TEM training is regarded among Australian general aviation pilots, with a strong focus in the private pilot and the flight training sectors, in terms of its use and effectiveness.

This survey is the second stage of PhD study by the principal investigator and is intended to gather information on how TEM training is conducted and whether it is perceived to be beneficial for pilot training and an effective method of improving flight safety. The findings from this will provide the basis for developing an appropriate instrument and questionnaires to generalise the findings.

Participation

Your participation will involve participation in an interview that will take approximately 1 hour of your time. The interview will take place at a time and venue that is convenient to you. The interview will be audio recorded and below are some of the questions that will be asked during the interview.

- Are you aware of any difficulties or challenges General Aviation (GA) organisations (e.g., aeroclubs or flight schools) faced when they were implementing TEM in their organisations?
- How would you rate CASA's role in introducing and providing on-going supports for TEM training?
- As an examiner, how do you examine the application/demonstration of TEM principles during a flight test?

Your participation in this project is entirely voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage. You may also request that any data collected about you be destroyed. If you do wish to withdraw from this project or withdraw data collected about you, please contact the Research Team (contact details at the top of this form).

Your decision whether you take part, do not take part, or to take part and then withdraw, will in no way impact your current or future relationship with the University of Southern Queensland.

Expected Benefits

It is expected that this project will not directly benefit you. However, the data collected through the interview are expected to provide in-depth information on the current state of TEM training in general aviation and to provide guidance for further research.

Risks

There are no anticipated risks beyond normal day-to-day living associated with your participation in this project.

Privacy and Confidentiality

The conduct of this research involves the collection, access and/or use of your non-identified personal information. The information collected is confidential and will not be disclosed to third parties without your consent, except to meet government, legal or other regulatory authority requirements. Audio recordings and notes taken during the interviews will be stored securely as per University of Southern Queensland's Research Data Management policy and will securely remain for at least six years after the completion of the study. Audio recording of the interview will be transcribed by the principal investigator and a copy of the transcript of the interview will be provided to you for your perusal and endorsement prior to inclusion in the project.

Consent to Participate

We would like to ask you to sign a written consent form (enclosed) to confirm your agreement to participate in this project. Please return your signed consent form to a member of the Research Team prior to participating in your interview.

Questions or Further Information about the Project

Please refer to the Research Team Contact Details at the top of the form to have any questions answered or to request further information about this project.

Concerns or Complaints Regarding the Conduct of the Project

If you have any concerns or complaints about the ethical conduct of the project you may contact the University of Southern Queensland Ethics Coordinator on (07) 4631 2690 or email ethics@usq.edu.au. The Ethics Coordinator is not connected with the research project and can facilitate a resolution to your concern in an unbiased manner.

Thank you for taking the time to help with this research project. Please keep this sheet for your information.

Appendix 6: Consent form (Study 2)

University of Southern Queensland

Consent Form for USQ Research Project Interview

Project Details

Title of Project:	Exploring the effect of introduction of Threat and Error Management in Australian general aviation
Human Research Ethics Approval Number:	H16REA238

Research Team Contact Details

Principal investigator Details
Mr Paul Lee
Email: paul.lee@usq.edu.au
Telephone: (07) 3470 4559
Mobile: 0450 567 747

Principal supervisor Details
Professor Paul Bates
Email: paul.bates@usq.edu.au
Telephone: (07) 3470 4549

Associate supervisor Details
Professor Patrick Murray
Email: patrick.murray@usq.edu.au
Telephone: (07) 3470 4556

Associate supervisor Details
Dr. Wayne Martin
Email: wayne.martin@usq.edu.au
Telephone: (07) 3470 4558

Statement of Consent

By signing below, you are indicating that you:

- Have read and understood the information document regarding this project.
- Have had any questions answered to your satisfaction.
- Understand that if you have any additional questions you can contact the research team.
- Understand that the interview will be audio recorded.
- Understand that I will be provided with a copy of the transcript of the interview for my perusal and endorsement prior to inclusion of this data in the project.
- Understand that you are free to withdraw at any time, without comment or penalty.
- Understand that you can contact the University of Southern Queensland Ethics Coordinator on (07) 4631 2690 or email ethics@usq.edu.au if you do have any concern or complaint about the ethical conduct of this project.
- Are over 18 years of age.
- Agree to participate in the project.

Participant Name	
Participant Signature	
Date	

Please return this sheet to a Research Team member prior to undertaking the interview.

Appendix 7: Sample extract from P3 interview (Study 2)

Facilitator:	Yep.
Interviewee:	In 2008, 2005 and 2008 when they were introducing it. They were sort of saying things like we're only still trying to find out how to do it, even later I spoke to a senior manager at CASA and he said, we have no idea what it's about so we were enforcing something that we didn't understand.
Facilitator:	Yep.
Interviewee:	Um, which I pointed out in my work.
Facilitator:	Mm.
Interviewee:	Um, so, what, the way they did that, because ATOs and FOIs came out who didn't really understand the system themselves but they had to show formally that they were training it, so we would get what I called "ritual chant to please the gods".
Facilitator:	[laughing], yep.
Interviewee:	I don't know whether you read that because I think I put that in my work. The ritual chant to please the gods. That is…
Facilitator:	I can't recall it exactly…
Interviewee:	To demonstrate it, for example, to demonstrate it that they were using Threat and Error Management, you're coming in from a training area, uh, at Archerfield, and at some stage they've got to show that they are aware of threats and how they need to be managed, so I would get this ritual chant, it was all learnt off by heart and that was, that was practically the same, everybody from a school would say virtually exactly the same words, that is, I'm coming into Park Ridge… are you aware of these places in the training area, Paul? One of the rejoin points for Archerfield circuit is a place called, Park Ridge from southern training area and there is 'Target' from eastern training area. Coming to those points and say, we are coming into this approach point, it's going to be traffic likely from the area converging here so there is likely traffic converging so I'll have to keep my eyes on that traffic and be aware of what they're doing and so they're letting me know that they're aware and thinking about it.
Facilitator:	Yep.
Interviewee:	So, they're going through the same, so I have to be aware of traffic coming in and so I'm looking out for those and I'll have to fit in and they might call first and I'll have to call second and so… [indiscernible]… what he's doing, what I need to hear it but nevertheless they were told to do it because instructors at a school said you have to make sure that you're aware of doing this.
Facilitator:	Yes.
Interviewee:	So they told me about that. Once we passed here, I will make a call, I will be given rejoin instruction, I'm expecting to land on Runway 10 or 28 and so I will go to this point and then I will come down and then I'll come around to land on the left or right runway and then I'll turn off at a taxi-way, whatever, because… [indiscernible]… taxi via, taxi to this and that and this comes back to my whole point. So they've given me whole this spiel and I'm sitting there and I'm watching this aircraft, converging on us from my right.
Facilitator:	Mmm.

Interviewee:	So we were going, he's going… so this guy's focusing on this spiel, giving this spiel, he finally finishes it and then he says, 'Ahhh', there is an aeroplane.
Facilitator:	[laugher]
Interviewee:	You haven't seen that before?
Facilitator:	Mmm.
Interviewee:	But he was so busy with saying his spiels, you know, he wasn't actually doing Threat and Error Management by keeping his eyes open.
Facilitator:	Yes.
Interviewee:	And caught him by surprise and this was common and I think because everybody said they were so worried about Competency Based Training and covering all, ticking all the boxes…
Facilitator:	Mm-hm.
Interviewee:	I would've… as a testing officer and as an instructor, I can see whether he's noticing an aircraft because he would say, there's an aircraft coming in to the right, 2 o'clock high, ½ miles.
Facilitator:	Yep.
Interviewee:	Well, that means he's watching and so then when he does something… well, he's managing threat and error, they were told to give this spiel and this happened in Archerfield at different schools, uh, Townsville, Rockhamton, at Toowoomba and the similar process for whole time because they felt that they had to produce by telling me all about it.
Facilitator:	Mm.
Interviewee:	And they are telling things that were going to happen after they landed but two seconds later, they call up and the tower will say they are landing on the other runway.
Facilitator:	Oaky.
Interviewee:	So the whole briefing meant nothing, which means they were wasting their time, trying to impress me with their Threat and Error Management and the whole thing was pointless because concentrating on that, they didn't see the aeroplane, and telling me what they were going to do after they land and the runway was changed and everything is out the window so it became such a... and CASA emphasised this is really important and so they emphasise their response to it. And that's why I said it was ritual chant to please the gods. The other ritual chant to please the gods was the before take-off and emergency briefs. You aware that they talk about what happens in case an engine failure and that kind of thing. So we're sitting, at a runup point, and they say, okay I'm taking off from runway 10, left or right, and I'm going to, my, abort speed is whatever it was and then once I'm, if I'm short of that then I'm going to stop on the runway and if it's passed that point, and again, everybody uses exactly the same practised spiels and said, after that I'm going to lower the nose, look 30 degrees either side of the nose for a suitable place to land and I'll land on that suitable point. And so, at the end of that… I would say, I don't know how many times I said this, so, okay, we're taking off on 10 Archerfield, off the end of the runway, there is a whole industrial area, full of factories.
Facilitator:	Mm.
Interviewee:	So we land… "Um, ah, there's a football field there just off the end of the runway" and I'll land on that and I said, that football field is, what, about 100m long, where it crosses is probably 50m wide. When was the last time you cleared sheds, land cuts, touch down and before the next lot of sheds in 50 m. Well, I've never done at… So do you think that's a realistic option?

Facilitator:	Mm.
Interviewee:	They've told to say that and I even said to some of the instructors… so on, well if you had an engine failure off Archerfield, how much height do you need, you think to turn back and land on the airfield somewhere, not on the runway so you're not at conflicting traffic. "Ahh, probably need a 1000ft". And I would say to instructors, how much height do you lose in a Cessna 152 or 172 in a 360° turn and they say, uh, probably 1000ft at least so I take them out to the aeroplane and say show me how much height we are losing. We are gliding, throttle's back, I want you to do a 360° turn, hold, concentrate on the gliding speed, I want you to use 30 degrees of that. Well, a 180° turn, they lose 250ft.
Facilitator:	Mm, yes.
Interviewee:	So I said, have you ever considered if you are above 300ft, you can turn back and land safely, you wouldn't crash into all those factories that would kill you.
Facilitator:	Yes, yes.
Interviewee:	So, come back to the question, how did they implement Threat and Error Management, they did it very consciously, very formally and completely impractically and quite… would be a complete failure.
Facilitator:	Yep.
Interviewee:	They actually, it actually taught people not to think of safe ways of landing.
Facilitator:	Mm.
Interviewee:	You know, I even say, you said 30° either side of heading, why that? And part of the problem is that start of the point is where long time ago people would say, some people have just gone solo and they are not that confident, I don't want you to spend time looking further than that, do the best you can from there and then at least you might crash wings level.
Facilitator:	Mm.
Interviewee:	Because what's happened is that if I think I can turn back they'll likely pull the aircraft around, stall, spin, crash and burn. But by the time they've done that little bit… the bit is now it will come automatic now, lower the nose, maintain the speed, now we can turn safely. And… that… that never got passed that first briefing and I've seen people at a commercial licence flight test. They're still saying exactly the same thing. So nobody has taken them past the point where they automatically lower the nose and we'll turn safely. They… They basically said, you can't turn back and land on an airfield.
Facilitator:	Mm.
Interviewee:	And yet, around the world, some of the most successful engine failures after take-off in single engine aircraft have been a turn back.
Facilitator:	Yes, yes, that's right.
Interviewee:	Often people criticise it, so how they've implemented it? Poorly. Consciously, very strong part of the process but to note it, it was a completely illogical way to deal with it.
Facilitator:	And the message I get, uh, from you've just said is, uh, there hasn't been much of resources, uh, made available from CASA for a start and even if there was, uh, it was quite weak, uh, in terms of, being used, in terms of, uh, worse, implemented.
Interviewee:	That's clear because I spoke to other, all the schools doing this ritual chant to please the gods as I said because CASA was… all the testing

	officers and examiners were, going and saying, that's good, that's good.
Facilitator:	Mm, yep.
Interviewee:	But nobody had thought about it. Somebody came up with a formula and they said, they're really doing it, everybody is doing the same, that's right, we've succeeded it and I was thinking 90% of those people would die if they had an engine failure after take-off, so it hasn't been implemented well at all.
Facilitator:	Sure.
Interviewee:	Completely impractical.
Facilitator:	Mm. And I guess, uh, that you as an examiner, have you got any special training from CASA in term of how you need to assess, uh, TEM, uh, when students are demonstrating it in the flight test?
Interviewee:	No, we… all ATOs had to, had to attend, about the time, 2008 onwards every two years, had to go to the professional development program. And in those things I raised some of those issues but generally everybody else was doing what I was observing.
Facilitator:	Mm.
Interviewee:	Most people were doing that, um, and so… and CASA was sort of reinforcing it in a way they were saying, you know, you've got to really do this but, and they say, you know even though in some of their, one of their civil aviation advisory publications, they say some of those things you've got to observe over a period of time but nobody ever criticised that, um, the basic approach of TEM.
Facilitator:	Mm.
Interviewee:	Nobody criticised it. In fact people used to… I mean I passed people on the course because I figured that they are producing what they were taught and what CASA reinforced but all I do was to talk to instructors and say, um, you need to get back to... and I had some results from that but over time it was almost a lone voice in, uh, in criticising the ritual chant which everybody was using.
Facilitator:	Mm.
Interviewee:	So, spread of that sort of approach was quite poor, uh, in that sense but I didn't fail anybody on that alone, I just said, look you guys, you need to think about this.
Facilitator:	Yes.
Interviewee:	Because everybody else was going to pass them so I had to be within the standard…
Facilitator:	Mm-hm.
Interviewee:	And, uh, I would've been standing as a lone voice, um, probably, um, an old… cranky old man… It will be a century…
Facilitator:	What about instructor training? Because, my… from my experience, uh, as an instructor if you really believe in strong and good airmanship, uh, and also TEM and so on, then it's more likely that you can, uh, you as an instructor actually pass onto your students but if you are a bit vague about the concept then it's unlikely, uh, the quality of whatever instructor is passing on is not going be a high quality. So, have you seen any of, uh, sort of, uh, special sessions or special training procedures, you know in a flight school, uh, when, uh, they are training their, uh, grade 3s or upgrading to grade 2 et cetera?
Interviewee:	I have to say apart from the formal way that people, schools were teaching, which I've just described, uh, and they wouldn't say this is what you've got to teach to the students but generally speaking, no. CASA didn't run anything particular, they've just said you have to

	make people aware of it, that they have admitted it, it's part of good airmanship. They said but we've spelled it out because it's so important about managing threats and errors to prevent a problem occurring or reducing the likelihood of serious problems occurring. All of that good philosophy but say that what they've produced was poor in regard and say you have to, they didn't say tick the boxes but that's the result of it. What I found was that before that we had airmanship and it wasn't, say, prescribed as it became broken down to microscopic little bits which you'd have to tick the boxes for, um, airmanship then people would talk about the ways they survive. I think the introduction of it and it's very hard to separate it from the introduction of Competency Based Training which I know you are not dealing with, um, but it was a part of that process. They added in under Human Factors and so on and so… and in the flight test and the Day VFR syllabus, where they had elements listed and the conditions under which they were to be performed and all that sorts of things. When they produced it, not everybody understood it so they had to make it something separate from what we were used to do with airmanship. Now airmanship wasn't a subject in a civil training school. It was in fact a subject in the Air Force . We had a subject called, airmanship, with an exam and that… the definition of airmanship was, uh, in effect anything to do with the safe and efficient conduct of air operations, uh, was a broad definition. So all sorts of things would come under that, from that kind of… what you discuss before you take off about how you handle emergencies, um, talking about a normal departure without emergencies before you take off so it's clear in your head about, um, lookout was always in it, but not just lookout but, um, what are you looking for so you, you go in this direction on a crosscountry but you're going to end up down there, why don't you look after take-off to see what the weather looks like down there, so I had one where we got airborne and we are going to Gold Coast and so we crossed the little range, went to Gold Coast, all the way to there, rain/showers all the way along that little range.
Facilitator:	Mm-hm.
Interviewee:	I am not sure what we can do, we did a little circuit at Coolangatta. I've seen the direction you want to go after departing Coolangatta, looked and said, we can get through there, we can get through quite a long way and it looks fine past that.
Facilitator:	Mm.
Interviewee:	So, you take off and go but if he didn't think about that until… to me, this is Threat and Error Management. We went and did the circuit, didn't look at any of that, he got airborne and I looked out, all he could see was rain/shower there and he didn't look past, say, around here is fine or at least I can go, below 4000ft then if I'm not sure what happens I can go there, if not I will turn back and come back here. He just looked up there and said, okay, we can't go, we are going back to Archerfield, okay?
Facilitator:	Mm, yep.
Interviewee:	Threat and Error Management wasn't there because the philosophy wasn't there because it became so formal and CASA never emphasised, uh, a system of dealing with that, um, I did it in my school when I was a CFI, I used to get people to look passed that, we talk about it and I wouldn't, I wouldn't want ritual chant. What I wanted, every take-off from Archerfield, different runways, I wanted a different

	briefing about where you would go, like taking off on 10, I wouldn't talk about 30 degrees either side of the heading, I would say 90 degrees to the right, there's an area which is going to be rough but you'll survive if you have to land on it and if you are higher than that it will be 180 degrees and you will land back again. Taking off from the other way there is a golf course which is, if you are high enough you can reach, if you are not high enough there was an area which you would probably damage the aircraft and probably crawl out but, no, but that's the best option from that direction.
Facilitator:	Mm.
Interviewee:	And I will run those things and I will tell FOIs that's what I was doing and none of my students ever had a problem passing the Threat and Error Management part of it.
Facilitator:	Mm.
Interviewee:	They were giving different briefings in each case. They wouldn't give spiels coming back into Archerfield. They would just say, okay, coming to Park Ridge, there's likely to be traffic and they're looking around and saying there's an aircraft there. They always got ticked on Threat and Error Management, people accepted what I taught but they also accepted ritual chant.
Facilitator:	Mm-hm.
Interviewee:	Problem solved without criticising. So, so was there any special training on it? No, CASA had run courses for private pilots where they could talk about those things but I think I only attended one and mostly they were dealing with questions and they don't, there was no overall philosophy except they've got to cover and demonstrate that you're considering threats and how you manage them.
Facilitator:	And it sounds like, uh, there was a lack of practicality, uh, in terms of…
Interviewee:	Yes.
Facilitator:	What you use of TEM.
Interviewee:	Yes, I think what happened was when Competency Based Training came in, non of the schools changed their practice at all from previous training.
Facilitator:	Okay.

Appendix 8: Ethics approval (Study 2)

OFFICE OF RESEARCH
Human Research Ethics Committee
PHONE +61 7 4687 5703| FAX +61 7 4631 5555
EMAIL ethics@usq.edu.au

UNIVERSITY OF SOUTHERN QUEENSLAND AUSTRALIA

10 October 2016

Mr Seung-Yong Lee

Dear Paul

The USQ Human Research Ethics Committee has recently reviewed your responses to the conditions placed upon the ethical approval for the project outlined below. Your proposal is now deemed to meet the requirements of the *National Statement on Ethical Conduct in Human Research (2007)* and full ethical approval has been granted.

Approval No.	**H16REA238**
Project Title	Exploring the effect of introduction of threat and error management in Australian general aviation
Approval date	10 October 2016
Expiry date	10 October 2019
HREC Decision	**Approved**

The standard conditions of this approval are:

(a) conduct the project strictly in accordance with the proposal submitted and granted ethics approval, including any amendments made to the proposal required by the HREC
(b) advise (email: ethics@usq.edu.au) immediately of any complaints or other issues in relation to the project which may warrant review of the ethical approval of the project
(c) make submission for approval of amendments to the approved project before implementing such changes
(d) provide a 'progress report' for every year of approval
(e) provide a 'final report' when the project is complete
(f) advise in writing if the project has been discontinued, using a 'final report'

For (c) to (f) forms are available on the USQ ethics website:
http://www.usq.edu.au/research/support-development/research-services/research-integrity-ethics/human/forms

Appendix 9: On-line survey for study 3

Examining the effect of post implmentation of Threat and Error Management in Australian General Aviation

This survey is the last stage of a PhD study by a student researcher and it examines the effect of post implementation of Threat and Error Management (TEM) in Australian General Aviation. The survey places a particular focus on obtaining opinions from GA pilots who primarily fly VH-registered aircraft.

The survey should not take more than 15 minutes of your time.

Thank you for your participation.

Participant information and Informed consent

Human Research Ethics Approval Number: HXXREAXXX

Research team contact details

Principal investigator	Principal supervisor	Associate supervisor
Mr Paul Lee	Professor Paul Bates	Dr Tarryn Kille
Email: paul.lee@usq.edu.au	Email: paul.bates@usq.edu.au	Email: tarryn.kille@usq.edu.au
Telephone: (07) 3470 4559	Telephone: (07) 3470 4549	

Purpose of this survey

Threat and Error Management (TEM) training was implemented more than eight years ago, but there have not been definitive data available to suggest whether TEM training has been well received and a positive effect of such training is witnessed, experienced or translated within the general aviation sector. Findings from this survey as well as previous studies in this topic will be used to examine the effect of post implementation of TEM training in Australian general aviation.

This project is being undertaken as part of a PhD project.

Expected Benefits

Findings from this survey and previous studies are expected to add to the body of knowledge that would better inform national aviation safety regulators including the Civil Aviation Safety Authority (CASA) with more comprehensive and accurately targeted education and training materials in the area of TEM and human factors.

Risk

In participating in this survey, there are no anticipated risks beyond normal day-to-day living

Privacy and Confidentiality

All responses in this survey will be treated confidentially unless required by law. The names of individual persons and their organisations are not required in any of the responses. A non-identified copy of this data may be used for other research purposes. Any data collected as a part of this project will be stored securely as per University of Southern Queensland's Research Data Management Policy (https://policy.usq.edu.au/documents/151987PL).

Consent to Participate

Clicking on the 'Submit' button at the conclusion of the survey is accepted as an indication of your consent to participate in this survey and to confirm you are over 18 years of age.

Questions or Further information about the Project

Please refer to the Research Team Contact Details above to have any questions answered or to request further information about this project.

Concerns or Complaints Regarding the Conduct of the Project

If you have any concerns or complaints about the ethical conduct of the project, you may contact the University of Southern Queensland Manager of Research Integrity and Ethics on (07) 4631 2214 or email researchintegrity@usq.edu.au. The Manager of Research Integrity and Ethics is not connected to the research project and can facilitate a resolution to your concern in an unbiased manner.

There are 9 questions in this survey

About yourself

[]Do you primarily fly VH-registered aircraft? *

Choose one of the following answers

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

[]Which of the following best describes your involvement in general aviation? *

Choose one of the following answers

Please choose **only one** of the following:

- ◯ Private operation
- ◯ Flight training
- ◯ Agricultural
- ◯ Small charter operation
- ◯ Other ____

[]Were you actively flying (commercially or privately) before Threat and Error Management (TEM) training became mandated in July 2009? *

Choose one of the following answers

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

[]On average, how many hours do you fly per month? *

Choose one of the following answers

Please choose **only one** of the following:

- ◯ Less than 10 hours

- ○ 10 ~ 50 hours
- ○ 51 ~ 100 hours
- ○ More than 100 hours

[]What is/was the highest CASA issued licence you hold/have held? *

Choose one of the following answers

Please choose **only one** of the following:

- ○ No Licence (Student pilot)
- ○ Private Pilot Licence (PPL)
- ○ Commercial Pilot Licence (CPL)
- ○ Air Transport Pilot Licence (ATPL)

[]Do you hold/have you held any of the following rating(s)?

Check all that apply

Please choose **all** that apply:

- ☐ Flight Instructor Rating
- ☐ Flight Examiner Rating

Please choose the appropriate response for each item that best states your view on TEM

[] *

Please choose the appropriate response for each item:

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
CASA has appropriately translated Threat and Error Management (TEM) concepts into practical guidance for general aviation	○	○	○	○	○
There are adequate practical examples relating to Threat and Error Management (TEM) principles and concepts for general aviation	○	○	○	○	○
When it comes to flying operations within general aviation CASA places greater emphasis on theoretical aspects of flying rather than practical aspects	○	○	○	○	○
Senior executives and decision makers within CASA have a good understanding on how general aviation operates	○	○	○	○	○
I find it easy to keep up-to-date with changes to aviation safety regulations/initiatives	○	○	○	○	○
I am satisfied with the way CASA develops aviation safety regulations/initiatives	○	○	○	○	○

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
CASA provides adequate training and guidance material on Threat and Error Management (TEM) training for general aviation	○	○	○	○	○
The concept of Threat and Error Management (TEM) is easy to understand	○	○	○	○	○
CASA explains Threat and Error Management (TEM) and how it affects my role and/or activities in a clear and succinct manner	○	○	○	○	○
I regularly attend CASA organised safety seminars (e.g., AvSafety) to keep up-to-date with the latest aviation safety initiatives	○	○	○	○	○
I find it useful to attend seminars such as AvSafety and/or others organised by CASA	○	○	○	○	○
I find the number of safety related publications and resources from CASA useful	○	○	○	○	○
CASA consults with all relevant stakeholders when developing and/or reforming aviation safety initiatives such as Threat and Error Management	○	○	○	○	○

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
CASA aviation safety advisors have a consistent understanding of Threat and Error management (TEM) principles	○	○	○	○	○
I am satisfied with the way CASA implements new aviation safety regulations/initiatives	○	○	○	○	○
The number of CASA staff with appropriate expertise in Threat and Error Management (TEM) is adequate	○	○	○	○	○
CASA maintains an effective on-going dialogue with the general aviation industry	○	○	○	○	○
CASA explains new safety regulations/initiatives and how they affect the general aviation industry in a clear and succinct manner	○	○	○	○	○
I use Threat and Error Management (TEM) principles when I am undertaking flying activities	○	○	○	○	○
I have a clear understanding of what I need to do when applying Threat and Error Management (TEM) principles	○	○	○	○	○

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
Consideration of Threat and Error Management (TEM) is an important part of my flight preparation	○	○	○	○	○
I always perform a series of briefings (e.g., emergency and take off briefings) taking into account the prevailing conditions and situations	○	○	○	○	○
I find Threat and Error management (TEM) briefing is an unnecessary additional task	○	○	○	○	○
I can see value in Threat and Error Management (TEM) when I am undertaking flying activities	○	○	○	○	○
The benefits of Threat and Error Management (TEM) training are over-rated	○	○	○	○	○
The use of Threat and Error Management (TEM) training/principles should be expanded to other sectors such as recreational aviation and Remotely Piloted Aircraft Systems (RPAS) sectors	○	○	○	○	○

[] *

Only answer this question if the following conditions are met:
Answer was at question '6 [Demog4]' (Do you hold/have you held any of the following rating(s)?)

Please choose the appropriate response for each item:

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
CASA provided adequate guidance to examiners before Threat and Error Management (TEM) became a mandatory flight test item	○	○	○	○	○
During flight tests assessment of Threat and Error Management (TEM) during a flight test is standardised among flight examiners	○	○	○	○	○
During flight tests candidates verbalise their briefings (e.g., emergency briefing) without putting conscientious thought into them	○	○	○	○	○

[] *

Only answer this question if the following conditions are met:
Answer was at question '8 [Demog4]' (Do you hold/have you held any of the following rating(s)?)

Please choose the appropriate response for each item:

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
During instructional flights my students verbalise their briefings (e.g., emergency briefing) without putting conscientious thought into them	○	○	○	○	○
I find CASA's flight instructor manual useful	○	○	○	○	○

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I have a clear understanding of what I need to do when teaching Threat and Error Management (TEM) principles to my students	○	○	○	○	○
My organisation provides adequate mentoring opportunities for junior flight instructors	○	○	○	○	○

Thank you very much for taking the time to complete this survey. Your respones and time are greatly appreciated.

Submit your survey.
Thank you for completing this survey.

Appendix 10: Ethics approval (Study 3)

OFFICE OF RESEARCH
Human Research Ethics Committee
PHONE +61 7 4631 2690| FAX +61 7 4631 5555
EMAIL human.ethics@usq.edu.au

UNIVERSITY OF SOUTHERN QUEENSLAND

11 June 2018

Mr Paul Lee

Dear Paul

The USQ Human Research Ethics Committee has recently reviewed your responses to the conditions placed upon the ethical approval for the project outlined below. Your proposal is now deemed to meet the requirements of the *National Statement on Ethical Conduct in Human Research (2007)* and full ethical approval has been granted.

Approval No.	H18REA135
Project Title	Examining the effect of post implementation of Threat and Error Management (TEM) in Australian general Aviation.
Approval date	11 June 2018
Expiry date	11 June 2021
Status	**Approved with standard conditions**

The standard conditions of this approval are:

(a) responsibly conduct the project strictly in accordance with the proposal submitted and granted ethics approval, including any amendments made to the proposal;
(b) advise the University (email: ResearchIntegrity@usq.edu.au) immediately of any complaint pertaining to the conduct of the research or any other issues in relation to the project which may warrant review of the ethical approval of the project;
(c) promptly report any adverse events or unexpected outcomes to the University (email: ResearchIntegrity@usq.edu.au) and take prompt action to deal with any unexpected risks;
(d) make submission for any amendments to the project and obtain approval prior to implementing such changes;
(e) provide a progress 'milestone report' when requested and at least for every year of approval;
(f) provide a final 'milestone report' when the project is complete;
(g) promptly advise the University if the project has been discontinued, using a final 'milestone report'.

For (d) to (g) forms are available on the USQ ethics website:

Appendix 11: The final measurement model with the standardised regression weights.