ATTRIBUTES OF AN EFFECTIVE FLIGHT INSTRUCTOR IN A UNIVERSITY AVIATION DEGREE PROGRAMME

A thesis presented in fulfilment of the requirements for the degree of

Master of Management

at Massey University, Palmerston North

John Magnus Murray

2004

Abstract

The flight instructor is at the heart of aviation practice and safety. Problems have arisen with the practice of flight instruction and it has been identified that improvements in flight instruction are required for the maintaining of good aviation practice and safety. Possible reasons for these problems are identified including the practice of flight instruction being used by newly qualified pilots as a career route towards an airline position.

Two methods of evaluating teaching, student evaluations of teaching (SET) and self reports from teachers, were used in this study to investigate the perceived attributes of effective flight instructors as rated by both students and flight instructors in a university aviation degree programme.

Twenty two flight instructors and seventy eight student pilots completed surveys which included questions on the importance of technical attributes, personal attributes, interpersonal skills and teaching ability for effective flight instructors.

Similar to previous research, students and flight instructors perceived domain expertise (piloting skills) and being a good teacher as attributes of effective flight instructors. Traditional measures of aviation expertise, namely the number of hours accumulated by pilots – were not rated as attributes of effective flight instructors. Differences between the students and flight instructors were in the area of interpersonal skills which were rated more highly by the students. This also closely aligns with findings from the literature, especially medical studies, where the teacher – learner relationship is highly valued by medical students.

General limitations and possible areas for future research are discussed. It is concluded that the findings from this study also align with the regulatory requirements for the training of new flight instructors.

Acknowledgements

I would like to express my thanks to Anne-Marie, Fiona, Gurjeet and Michelle for their support, guidance and assistance in completing this work.

John Murray

December 2004

Table of Contents

ABSTRACT II

ACKNOWLEDGEMENTS IV

TABLE OF CONTENTS V

LIST OF TABLES VII

LIST OF FIGURES VIII

OVERVIEW 1

CHAPTER ONE INTRODUCTION 4

1.1 The Flight Instructor 4

1.2 History of Instructing 8

1.3 The Problem 14

CHAPTER TWO LITERATURE REVIEW 21

2.1 Teaching Effectiveness 21

2.2 Student Evaluation of Teaching (SET) 24

2.3 Evaluating Teacher or Course 29

2.4 Reliability of SET 30

2.5 Validity of SET 31

2.6 Factors which may bias student evaluations of teaching 34
2.6.1 Grading Leniency and Student's GPA (Grade Point Average). 34
2.6.2 Instructor popularity 35
2.6.3 Experience of instructor 35
2.6.4 Student interest in subject 35
2.6.5 Course workload and difficult 36
2.6.6 Level of Student 36
2.6.7 Gender 37
2.6.8 The chosen discipline 38
2.6.9 Instructor personality 38

2.7 Summary of SET 41

2.8 Attributes of Effective Teachers 41

2.9 Self-Reports of Effective Teaching 42

2.10 SET and Flight Instruction 45

2.11 Medicine and Aviation 46

2.12 Defining the Research Problem 51

CHAPTER THREE METHODOLOGY 55

3.1 Design 55

3.2 Participants 55

3.3 Procedure 55

3.4 Measures 56
3.4.1Flight Experience 56
3.4.2 Technical Attributes 57
3.4.3 Personal Attributes 57
3.4.4 Interpersonal Skills 58
3.4.5 Teaching Ability 58
3.4.6 Meeting Organisational Goals 59

CHAPTER 4 RESULTS 60

4.1 Sample Description 61

4.2 Results 63

4.3 Attributes of an Effective Flight Instructor 64

CHAPTER FIVE DISCUSSION 71

5.1 Attributes of an Effective Flight Instructor - students 71

5.2 Attributes of an Effective Flight Instructor – flight instructors 73

5.3 Comparison between students and flight instructors 77

5.4 Differences between experience levels 80

5.4 Limitations 81

5.5 Future Research 84

5.6 Conclusion 87

REFERENCES 89

List of Tables

Table 1: Reliability Analysis for the five effective instructor scales (Alpha) 63

Table 2: Means and standard deviations for attributes rated as most important for effective flight instruction – all students 64

Table 3 Means and standard deviations for attributes rated as least important for effective flight instruction – all students 65

Table 4 Means and standard deviations for attributes rated as most important for effective flight instruction – all instructors 66

Table 5 Means and standard deviations for attributes rated as least important for effective flight instruction – all instructors 67

Table 6 Means and standard deviations on attributes of effective flight instructors for instructors and students 68

Table 7: Means and standard deviations on attributes of effective flight instructors across flight experience levels 70

List of Figures

Figure 1: Frequency of Flight Instructor and Student respondents.....62

Overview

This study seeks to establish the attributes of effective flight instructors as identified by students and flight instructors.

Flight Instructors are the foundation of the aviation industry (Henley, 1991). From this person the student aviator learns the physical flying skills, the cognitive knowledge of the operating environment and is imbued with a safety discipline. The instructor is the pilot, the teacher and the role model laying down the foundations of the student's future flying. Despite the increasing availability of technology in flight instruction and changes in the aircraft being used for flight, there has been little change in the way that flight instructors teach students how to fly. Koonce (2002) describes how many training organisations actually pride themselves on being old fashioned in the flight training offered to students.

Due to the industry structure, most new commercial pilots become instructors not for a career option but as a stepping stone (Henley, 1995) on the pathway to future employment by an airline. Instructor training has become a continuation of the training provided for the pilot as he/she progresses to their goal of employment with an airline (Woodhouse, 2001). As most new Commercial Pilot Licence holders need to become instructors to advance their employment credentials there is little or no attempt to select or train the type of person who may have the attributes to be an effective flight instructor. Complaints about the decline in instructing standards have come from all sectors of the aviation industry, (O'Hare & Roscoe, 1990). The New Zealand Civil Aviation Authority's "Towards 2005 – The Aviation Safety Plan Forum" (Civil Aviation Authority of New Zealand, 2001) identified 24 specific areas of concern regarding

aviation safety and nearly one third of these areas of concern were to do with flight training and the standards of flight instruction offered by current flight instructors. Discussions arising from this Safety Forum about improving the standard of flight instruction focus on the instructor (instructor-centred, Telfer, 1993a), as in broader educational settings it has been noted that improving the instructor can improve the effectiveness of the instruction provided (McKeachie, 1997a).

There has been little or no examination of the students' expectations of flight instruction and flight instructors. Yet within many academic disciplines students are used to evaluate the effectiveness of the teaching they have received in order to improve the teaching (Marsh, 1987). Not only are learning institutions using student evaluations but national government bodies are also drawing upon student evaluations as an indication of effective teaching (Ballantyne, Borthwick & Packer, 2000).

The present study uses the ratings of students within a university aviation programme to evaluate their perceptions of the attributes required of flight instructors that are effective or ineffective in helping them learn to fly. This input by the students into shaping the learning process is considered beneficial for instructors (Henley, 2003).

This study will also evaluate the perceptions of the flight instructors themselves regarding the attributes of effective flight instructors and note any lack of congruence or alignment between students and flight instructors perceptions, (Moore, Lehrer & Telfer, 2001).

Chapter One will describe the role of the flight instructor. It will describe the current flight training practices arising from the flight instruction practices developed in the first and second world wars. The problems created for flight instructing by the

necessity of newly qualified Commercial Licence Holders using flight instructing as a means to advance their own professional aviation careers will also be discussed (Henley, 1995; Woodhouse, 2001).

Chapter Two reviews the literature surrounding the evaluation of effective teaching and how teaching and teachers can be evaluated. The use of students to evaluate teaching (SET – Student Evaluation of Teaching) has been found to be a valid and reliable method of evaluating teaching (Marsh, 1987). Although there is a paucity of material regarding flight instruction, the field of medicine has been regarded as similar to aviation as both are complex environments where the future actions of the student may cause serious injury or death and where teamwork is required (Helmereich, 2000b). Therefore the literature related to the attributes of effective medical teachers, both classroom and clinical, as evaluated by medical students is reviewed. The aims of this study are presented at the conclusion of Chapter Two.

Chapter Three outlines the methodology, design and procedures of the present study. A survey instrument formulated through the review of material surrounding the evaluation of effective teaching by students and teachers was used for data collection. The data was collected from (a) students in the flight training major of a university aviation degree programme, and (b) their flight instructors through a cross-sectional survey.

Chapter Four describes the participants and presents the results of the present study. Chapter Five discusses the findings of the present study in the context of previous research. Limitations of the study and implications for future research are also discussed.

Chapter One

INTRODUCTION

This chapter will describe the role of the flight instructor. It will also describe how the current flight training practices have arisen from the instructional practices developed in the first and second world wars. It will further describe problems created for flight instructing by the necessity of newly qualified commercial pilot licence holders becoming a flight instructor as a means to advancing their own professional aviation careers, (Henley, 1995; Woodhouse, 2001).

1.1 The Flight Instructor

The flying instructor is the pilot and teacher who introduce the *ab initio* or beginning student to the physical motor skills and cognitive demands of flying. Koonce (2002) identifies the performance of the flight instructor as having a bearing on the future safe flying of the student pilot. Skills, knowledge and attitudes learnt in the formation period of mastering the initial tasks of flying will provide a platform for later flying, including possible professional piloting careers (Hunter, 1999). Ramsey and Ramsey (1996, p. 1) describe the role of the flight instructor as having "a formative influence on the attitudes and behaviours of new pilots".

The skills involved in successfully flying the aircraft have been identified by de Montalk (2000) as including physical motor skills, aviation knowledge and attitudes. While traditionally the first two tasks have been the focus of flight instructor attention, more recently the focus of pilot performance has become less on skill levels and deficiencies and more on attitudinal problems and knowledge deficiencies (Besco, 1992;

Thomas, 2001). This is due to the recognition that pilot error has caused a disproportional number of aviation accidents. An examination of New Zealand air accidents during the 1980's indicates that decisional factors account for a majority of accidents (O'Hare, Wiggins, Batt & Morrison, 1994). This is not confined to New Zealand. Li, Baker, Grabowski & Rebok (2001) examined data files over a similar period from the National Transportation Safety Board and found that 85% of General Aviation crashes in the U.S.A. were caused by pilot error. It has thus been recognised that decision-making is an important piloting skill that must be learnt by trainee pilots (Jenson 1995). The attitudes driving the decision making of pilots are formed as a result of the training the student receives, including how the flight instructor makes aeronautical decisions (Jenson, 1995). Therefore the teaching of safe attitudes and decision making has become an important part of the role of the flight instructor (Koonce, 2002).

The task of flight instruction requires the flight instructor to be more than pilot. As well as mastering and utilising the skills involved in flying the aircraft the flight instructor must also be able to transfer these skills to the student, a role for which it has been suggested the flight instructor has been poorly prepared (Henley, 1995). The role and methods of the flight instructor and the curriculum that the instructor teaches to, follows the patterns of military flight instruction (O'Hare & Roscoe, 1990). One military pilot describes the training offered to pilots in initial stages of training as involving handling competence in a single-engined aircraft (Lewis, 2000), in short; the physical motor skills of "stick and rudder".

It is these different skill-sets which must be demonstrated and taught together for the flight instructor to become effective at flight instruction. Competence as described by

Hager and Gonci (1996) is an integrated concept, not merely the listing of a series of task descriptions, in the example of flight instruction the technical flying skills and descriptions of each flying exercise. It also involves broader aspects of attributes and capabilities such as planning, problem solving and communication, which are context dependent. Task performance is observable – it is easy to describe limits for a medium turn with angle of bank being 30° and holding of altitude to plus or minus fifty feet – whilst attributes such as cognitive skills for decision making or interpersonal skills for multi-crew operations are inferred and less easy to describe and measure. Aviation has concentrated on task performance and ignored or poorly defined the abilities and capabilities, other than task performance, that are required to be displayed by pilots and flight instructors (Damos, 1996). Defining competence exclusively in terms of either tasks or attributes without the other is to lower the competence required of the flight instructor (Hagar & Gonci, 1996). The competent flight instructor must be able to perform the observable tasks of flight instruction such as flying but must also possess and display the other non-observable attributes and capabilities required for flight instruction.

Arnold (1991) notes that flight instructors see themselves as pilots firstly and not as teachers, receiving little if any training in person to person communications. Henley's (1995) research also identifies the poor to non-existent training flight instructors' receive to prepare them for the role of teacher. The training instead focuses on observable physical flying skills. In wider educational settings McKeachie (1997a) identifies this idea of subject knowledge being sufficient for good teaching as being a barrier to good teaching. In the case of flight instruction the subject knowledge is seen as flying skills. Teaching and inter-relational skills are components of flight instruction that must also be

demonstrated for competence in flight instruction but receive little attention in the training of flight instructors (Henley, 1995).

As the flight instructor requires an integrated competency so too do the students taught by the flight instructor. There is growing awareness that students emerging from education institutions need not only be skilled in the immediate tasks of the educational programme but also need to have developed the skills required for future employability, life long learning and taking a leadership role in society (Wilson & Lizzio 1997). Not only are the flying and training tasks that are required of the flight instructor an important part of the role of a flight instructor, there is also a realisation that how the instructor goes about achieving the tasks of flight instruction is just as important (Termoehlen, 1987).

The regulatory requirements for becoming a flight instructor describe the experience levels required to be a flight instructor and the training to be undertaken to become a flight instructor but there is no description of the personal attributes required of the flight instructor. The personal attributes that the flight instructor brings to the role of flight instruction are just as important as the tasks expected to be performed. However the personal attributes and instructing methods of flight instructors can for students form barriers to their learning. There is evidence that the flight instructor-student relationship can hamper the development of a safe and effective pilot through the raising of student anxiety levels during the flight instruction process (de Montalk, 2000).Other research has found within student pilots a heightened fear of the flight instructor relative to other fear (Grandchamp, 1971). In addition Henley (2003) states that for close to 40 per cent of

student pilots the cause of psychological stress most likely to impair their learning is the flight instructor.

The role of the flight instructor has been variously described as the foundation of the flight training system in the development of competent and safe pilots (Henley, 1991), a fundamental part of aviation training (Woodhouse, 2001), and at the heart of the aviation culture (Ramsey & Ramsey, 1996).

1.2 History of Instructing

The first people involved in heavier-than-air flying were self-invited. Having the inclination was enough. Of necessity these first aviators were self-taught, there was no instructor to show the way.

An example of this attitude was early aviation pioneer Claude Grahame-White (Gwynn-Jones, 1996). Inspired by Channel conqueror Bleriot, Grahame-White sought to purchase an aircraft from Bleriot. As the aircraft was built towards the end of 1909 Grahame-White was not only allowed to help with the construction but also began to acquire tips on how to fly. Once the aircraft was completed Grahame-White and a friend taxied back and forth across the military parade ground that was at the time the aerodrome for Paris until they felt confident enough to lift off into ground effect and "fly" for a few feet in a hopping manner across the parade ground. Moving to a larger airfield Grahame-White continued to learn to fly by teaching himself until he was competent enough to be awarded the French licence, a *brevet de pilotaviateur*. Grahame-White went on to be an accomplished pilot in both United Kingdom and America. He opened an aviation school to instruct new comers to aviation and took part in early air races (Gwynn-Jones, 1996).

In New Zealand one of the earliest Flying Schools was set-up in Auckland by the Walsh brothers who had to first build their own aircraft and then teach themselves to fly it (Bentley 1969). From these self-instructed beginnings the Walsh brothers were able to instruct other pilots who in turn became instructors. This then was the beginnings of flight instruction; there was no flight instructor. The earliest pilots were required to teach themselves to fly (Clouston, 1954). As the first pilots began to instruct other pilots in the art of aviating differing methods of instruction were used.

Henley (1995) identifies three different methods of flight training during the beginnings of flight training. The Wright method - which became the accepted practice - had the instructor and the student each occupying their own seat with their own controls. This method has continued through to current practice. Other methods involved the instructor sitting on the wing of an aircraft that could not aerodynamically get out of ground affect. The instructor shouted instructions to the student and as the aircraft picked up speed and lifted off to a level of a few feet above the ground the instructor would jump off and the student would "fly" the aircraft to the far end of the field. Another method required the student and instructor to squash into one seat with the instructor reaching around the student to manipulate the controls. There is little doubt as to why the Wright method has become the standard method.

The commencement of the First World War with its increasing demands for trained pilots (500 pilots a month on the western front (Mackersey, 1998)) saw the beginnings of more formalised flight instruction. However, those chosen to be instructors were often those who had little interest or competence in being flight instructors, (Orange 2001) or students found themselves being instructed by those men who had been at the

front but were thought to have lost their nerve (Lewis, 1936; Mackersey, 1998). This shifting into flight instruction roles those men who were no longer considered fit to be at the front in active service was to continue in the second world war (Orange, 2001).

Not that there was, at the beginning of the war, much instruction given to students. Training consisted of taxi practice during which a high enough speed was reached so that the aircraft lifted off and was placed on the ground again, followed by one dual flight and then the student was sent solo which included having to loop the loop (Cobby, 1981). Lewis (1936) found himself going solo after one and a half hours dual and Keith Park (later Air Chief Marshal Sir Keith Park) went solo in two and half hours (Orange, 2001). Park, after further training, was sent to be a flight instructor with a mere 35 hours experience. In his biography of Air Chief Marshal Sir Keith Park, AOC of 11 Group during the Battle of Britain, Orange (2001, p.16) notes "that a man of his brief experience could officially be considered capable of instructing others is a revealing comment on the dire standard of flying training".

Towards the latter half of the war there were efforts to improve the standard of instructing through the development of schools dedicated to the training of flight instructors and the standardisation of the training. Major Smith-Barry at the Central Flying School in the United Kingdom developed standardised flying lessons and expected the lesson to follow a set pattern with the instructor demonstrating the manoeuvre whilst talking about what he the instructor was doing (Elshaw, 1993). This lead to the development of the instructors patter, known as the "Gosford patter" after the station at which it was developed (Cobby, 1981). Learning the instructor's patter continues to be part of becoming a flight instructor (Civil Aviation Authority of New

Zealand, 1999). Cobby (1981) recalled his instructing of students required him when airborne to fly along side the student who then demonstrated the exercises that had been discussed on a blackboard on the ground prior to the flight. The patterns of flight instruction developed in the first two decades of flying – the Wright method of instruction and the instructor's patter – form the basis of current flight instructional praxis.

During the intervening war years many of the instructors were returned servicemen. There were few openings for pilots in fledging airlines and to continue to fly they had to instruct. The standards required to be reached by the recipients of flight instruction during this period were not great. In the early part of the 1920's there was no pilot's licence in the United Kingdom. Instead, a certificate was issued by the Royal Aero Club. The test for this certificate as flown by author and aircraft designer Nevil Shute entailed "an ascent to six thousand feet, flying several figure-of-eights at about a thousand feet and a landing without damage" (Shute, 1966, p. 52). By 1939 there were 39 pilots in New Zealand with an instructors rating (Rendel, 1975).

At the beginning of World War 2 the instruction given to student pilots was not any different to that in peace-time (Lewis, 2000). However the Royal Air Force (RAF) at the commencement of hostilities foresaw the need to train many more aircrew than was possible in the limited spaces of Great Britain (Marshall, 2000). As it was estimated that 20,000 to 30,000 trained aircrew would be required to replace those aircrew lost, the British government turned to its commonwealth dominions and established the British Commonwealth Training Plan. This involved initially Canada, Australia and New Zealand with later Rhodesia and South Africa becoming part of the plan. Initial training

was to be provided in the country of origin and then advanced training was to take place in Canada. These cadets often flew an aeroplane before they learnt to drive a car (Marshall, 2000). As New Zealand had to provide initial training of pilots who were to proceed to war service, the newly established Royal New Zealand Airforce established a flying instructor's school at Mangere, Auckland (Bentley, 1969). Civil instructors were sent to this school to be refreshed and updated in instructional methods before becoming part of the Commonwealth scheme. The influence of the military on flight instruction in Commonwealth countries in the years following the cessation of hostilities has been very strong (Henley, 1995) with New Zealand being no exception.

In the United States aviation psychologists became involved during World War 2 in the training of airmen (Koonce, 1984). In addition, the Committee of Selection and Training of Civilian Pilots was established. One of the results arising from this committee in 1943 was the development of ground instruction as part of flight instruction, especially pre and post briefings and the development of planned lessons (Henley, 1995).

The planned lessons, along with the instructors' patter are still the main focus of how the instructor operates (Elshaw, 1993) and how the instructor is trained (Henley, 1995). The patterns of flight instructor training laid down during the First World War, that became entrenched in the 1930's and the Second World War through the British Commonwealth Training Plan, have remained as the training pattern of flight instructors down to the present time (O'Hare & Roscoe, 1990; Arnold, 1991). The establishment of university aviation programes, firstly in the U.S.A. after World War 2 (Koonce, 1984), and more recently in Australia and New Zealand, has not greatly changed the operation of the flight instructor.

This is perhaps not surprising as it is believed that efficiency and success in flight instruction requires standardisation of teaching methods and skills amongst instructors, (Roscoe, 1980). Yet as Elshaw (1993) notes, whilst there have been large developments in the training hardware (e.g. the aircraft, the simulators including low fidelity simulators, the increasing use of personal computers), there has been little or no development in the practice and techniques of flight instruction. Despite the increasing development of multimedia and other instructional tools such as Web CT there is an acknowledgment that well qualified and competent humans (i.e. flight instructors) are unable to be replaced by technologies (Laurillard, 1997). The human component of flight instruction remains the most important component but suffers from the least attention. This has meant that instructional practices have changed little over half a century. Since the second world war hand-me down ideas and methods are passed from one generation of instructors to another as the top layer of experienced instructors proceed to airline employment taking all they have learnt of the art and science of instructing with them (Henley 1995). This has meant that instructors follow "folkways" (Telfer, 1993b), unwritten patterns of behaviour that are considered to be correct. These folkways are patterns of routine and habit and are difficult to change (Merton, 1957). The result of learning about teaching through observing their own teachers may lead the new instructor to have fragmented and /or erroneous ideas about what is required for effective teaching (Hativa, Barak & Simhi, 2001). Hativa *et al* (2001) note that exemplary teachers (as compared to non-expert teachers) at tertiary level thought more about teaching and linked their increased knowledge of effective teaching strategies with their actual teaching practices. Within flight instructing there is very little reflection on what is effective

instruction as the traditional patterns and practices are not only observed by the intending instructor but are also taught as the correct manner of instructing (Henley, 1991). Woodhouse (2001) notes that inexperienced instructors teach the student instructors who in turn, whilst still inexperienced, teach the next group of instructors who in turn teach the new instructors and so on. These inexperienced instructors are reverting back to learned patterns. These habitual patterns of performing the role of flight instructor become ingrained and establish a culture that influences the understanding of the role of the instructor. Pomson (2002, p. 23) notes that "research has consistently shown that culturally embedded archetypes, inherited images and traditions of teaching powerfully shape the ways in which teachers conceive of their work and lives, their roles and responsibility".

There is a growing sense of the need to improve the standard of flight instruction. Calls for the improvement of instructor knowledge, skills and abilities range from academics (Hunt, 2000), to airline pilots in the form of their union ALPA (Fiorino, 2001), to the instructors themselves (Henley 1991 & 1995), as well as regulatory authorities (O'Hare & Roscoe, 1990) including Civil Aviation Authority of New Zealand (2001).

1.3 The Problem

The supposedly strongest link in the aviation chain, the instructor, is actually the weakest link (Henley, 1995). This criticism appears to be widespread as O'Hare and Roscoe, (1990, p. 61) note " most national and international aviation organisations are increasingly accepting the position that flight training, particularly in general aviation, is a problem".

As in other similar professional fields to aviation where there is an interaction between humans and safety critical environments, such as medicine, (Sexton, Thomas & Helmreich, 2000) it would be expected that the pre-eminent and most experienced practitioners of aviation were the ones entrusted with the responsibility of providing the new student pilots with the best possible training as they enter into the aviation industry. Unfortunately this is not so. The experienced pilots fly for the airlines and have little or no participation in general aviation (O'Hare & Chalmers, 1999), including the training of students. Inexperienced pilots become instructors who train the next generation of pilots who, despite being even more inexperienced than their instructors, become instructors and so the cycle continues (Woodhouse, 2001). The entry requirements for training and becoming an instructor are very low, allowing the inexperienced pilot to become an instructor. Further cause for concern are the ever changing tasks required to be performed by the modern pilot, tasks that the few remaining experienced instructors were not trained for (Roscoe, 1980).

It is not surprising that newly qualified pilots with a commercial pilot licence become instructors although it is normally not seen as or expected to be a career choice by the new pilots (Henley, 1995). Nearly 75% of Californian Flight Instructors surveyed as part of a study into personality characteristics of successful flight instructors responded that in five years time they wanted to be flying as an airline or corporate pilot (Arnold, 1991). Woodhouse (2001, p. 2) notes "far too many instructors are there to gain the hours they need to go on to the airlines ... it's not their fault, it's an industry structure problem".

The industry structure requires newly qualified commercial licence holding pilots to gather further experience in terms of hours flown, to be seen as ready to enter the job market, (Frederick-Recascino & Hall, 2003) and have the right credentials for airline employment. Hunter (1999) describes this attitude as where "experience is equated with expertise" (p. 27).

An example of this experience and expertise link is illustrated by one flight training school owned by a regional airline in North America that offers students gaining a CPL through the school an opportunity to interview for a first officer's position with the airline. However, before getting this chance, the students after gaining a CPL, need to first work for a year as a flight instructor with the training school. In this year working as a flight instructor the pilot gains another 800 hours flight experience which makes them eligible for employment as a first officer with the airline. The airline employs as first officers 97% of the flight instructors working at the flight training school (Fowler, 2000).

The method employed to achieve this "hour building", that is to gain the experience required to be seen to be an expert, is to become an instructor. There are very few, if any, other avenues apart from instructing open to these newly qualified pilots to achieve the large number of hours listed by the airlines as necessary for employment with the airline. Instructing is not seen as a career option for most pilots who become instructors (Henley, 1995). Koonce (2002) notes that the genuine desire of flight instructors to fly and teach is often absent as they use instruction to merely gain flying hours needed to get employment with an airline. It is therefore not surprising that the commitment and competence of most instructors can be called into question. Instructing

is viewed as a transitory stage or a stepping stone (Henley, 1995) of a pilot's career which has the potential to affect the standards of flight instruction offered to students.

For the aviation industry credential inflation has developed in regards to the CPL and the 'hour building' exercise that has seen pilots completing the flight instructors rating. Credential inflation occurs when qualifications that were required by job seekers are no long enough and higher qualifications must be held by the next generation of job seekers to get the same type of work. It is perhaps not surprising as Nash (1985) notes that all educational systems have credential inflation as one of their characteristics.

The credential value of the CPL has decreased in terms of being able to gain entry to airline employment. A new credential for employment has arisen based on a minimum number of 1,000 flying hours. This is based on the idea that greater experience as measured in number of hours equates to expertise (Hunter, 1999).

A tightening labour market also saw the increase in the number of flight training organizations. Historically in New Zealand, ab initio training was carried out by Aero Clubs but over the last 15 years there has been a transition to Flying Schools and Universities (de Montalk, 2001). As McKenzie, (1987) notes, the retention in educational institutions is linked to the need to remove students from a tight labour market and avail them of the increased availability of educational opportunities. The same has applied for pilots as the labour market shrunk and there were increased educational opportunities. The instructors rating became the next qualification to obtain after gaining the CPL. Therefore there grew within New Zealand aviation a culture of pilots becoming instructors who did not intend to make instructing their career but were merely passing through, as they sought to gain the credential of 1,000 hours. This has meant that the

instructor rating that was once a qualification for employment in itself is becoming used to gain the next credential. Instructing becomes a credentialing exercise where it is not seen as a specialty career option but as a step on the pathway to future employment by an airline (Arnold, 1991; Henley, 1995). This has been described as a necessary evil, (Henley, 1995). For the training organizations and the commercial license holding pilot, the instructor training is merely a continuation of the training provided for the pilot (Henley, 1995). As more pilots use the instructors rating to gain a further credential before moving to other employment opportunities, the standard of instructing decreases because as Woodhouse, (2001) notes, the transitory nature of the flight instructor has lead to inexperienced instructors teaching inexperienced pilots to become instructors who in turn teach the next group of trainee instructors.

Whilst airlines will employ different forms of selection testing from the use of psychometric measures to simulator rides, there are no selection criteria for those wanting to become flight instructors. There is no description of the type of person and attributes of the person who will be an effective flight instructor. . There is no selection policy as to who should be an instructor. Yet research from Grandchamp (1971), de Montalk (2000) and Henley (2003) indicate that the flight instructor can be the cause of stress that blocks or hinders student learning.

The standard of training and the resultant standard of flying skills demonstrated by pilots are often criticised as not meeting the needs and expectations of the aviation industry. The Civil Aviation Authority of New Zealand's Safety Forum (2001) in formulating its safety plan identified nine areas of aviation activity where the prevention of accidents could come about through teaching and instilling in students appropriate

skills and attitudes. However it was noted that accident prevention was hindered by the quality of the flight instructors providing instruction to students. Identified instructional deficiencies included the lack of adequate initial training provided to *ab initio* students, the lack of experience amongst instructors with inexperienced instructors training student instructors. Further, once pilots have qualified as instructors there is poor or non-existent supervision of new instructors and a lack of ongoing development and refresher training. (Civil Aviation Authority of New Zealand, 2001). The accident rates evidenced in the accident statistics stays stubbornly high, especially - but not exclusively - amongst GA activity. Hunter's (2001) study of accident risk indicators showed that those involved in single engine piston-powered flight were most likely to be involved in an accident. This is the area that *ab initio* flight instructors' work.

The flight instructor has been variously described as the foundation (Henley, 1991), the heart (Ramsey and Ramsey, 1996), and fundamental (Woodhouse, 2001) to aviation. Yet the supposed strongest link has become the weakest link (Henley, 1995). Reasons for this have been identified as:

- the inexperience of instructors (Woodhouse, 2001)
- the low entry threshold for those wishing to become instructors (Woodhouse, 2001)
- the low status of the instructor within the aviation industry and accompanying poor employment conditions (Arnold, 1991)
- the use of instructing as a stepping stone by instructors to other more desired positions in aviation (Henley, 1995)
- the poor training of instructors and the instructors of instructors (Henley, 1995),
- the emphasis of flying skills as against teaching skills and abilities (Henley, 1991),

- the lack of ongoing training and development (Civil Aviation Authority of New Zealand, 2001).

Improvement in the performance of flight instructors is a major aim of the Civil Aviation Authority of New Zealand as it strives to increase safety in aviation. Current practice has qualified holders of a Commercial Pilots Licence proceed towards gaining airline employment by becoming flight instructors (Henley, 1995; Woodhouse, 2001). There has been little attempt to define who should be a flight instructor and the attributes that the flight instructor should possess to provide effective flight instruction.

Chapter Two

LITERATURE REVIEW

This chapter reviews the literature surrounding the effectiveness of teaching offered at tertiary level education. The use of students to evaluate the effectiveness of teaching they receive is reviewed. As there is a paucity of material regarding the effectiveness of flight instruction, the effectiveness of the teaching of medical studies (being considered comparable to aviation) is examined to ascertain which attributes are valued by students involved in clinical and classroom medical studies.

2.1 Teaching Effectiveness

There is growing awareness that the quality of teaching in tertiary institutions needs to improve (Ramsden, 1992). There has been a greater desire for accountability within higher education from governments and the consumers – the students – themselves (Wilson & Lizzio, 1997).

Teachers and their behaviours have an impact on student learning and attitudes to learning (McKeachie, 1997a; Murray, 1997). For Murray (1997) the teacher is a more important factor for effective teaching than are other factors such as individual and social differences of students. Patrick and Smart (1998) identify the teacher as being crucial to the teaching process. Therefore this crucial component of the teaching process, the teacher and their teaching, should be examined and evaluated. Centra (1993) lists three methods for evaluating effective teaching in the tertiary sector;

- evaluations by students,
- reviews by peers or colleagues or chairs, faculty development specialists,

- self-reports.

Marsh and Roche (1997) also add a further method of evaluating teaching effectiveness by using former students.

Teaching effectiveness is difficult to define (Cashin & Downey, 1992; Martin, 1998; Wachtel, 1998). One of the difficulties in defining effective teaching is its subjective nature. What works well for one student may not for another student within the same class (McKeachie, 1997a). There is though general agreement amongst researchers that student learning is one of the most important critera of teaching effectiveness (Cohen, 1981; Marsh & Roche, 1997; Stringer & Irwing, 1998). However, whilst student learning is an important indicator, it is not the only one (Marsh, 1987). Marsh and Roche (1997) caution against interpreting student learning in a narrow and constraining manner such as from results in multi-choice tests. In the field of flight instruction some national aviation authorities such as Canada's regulatory body, Transport Canada, judge the teaching effectiveness of flight instructors by the results of students sitting flight tests for private or commercial flight tests (Henley, 1991). A study of Californian flight instructors used as its definition of success in flight instruction the percentage of students who completed primary flight training (Arnold, 1991).

d'Apollonia and Abrami (1997) see definitions of effective teaching falling into two categories: the process of instructing and the products or results that arise from the instruction. Within some North American Universities students are not asked to directly rate the teacher but to rate their achievement of educational objectives. Helping students achieve educational goals is seen as being part of teaching effectiveness (McKeachie, 1997a). Abrami, d'Apollonia and Rosenfield (1997) see effective teaching as being

multi-dimensional and bringing about positive changes in students in cognitive, affective and psychomotor domains. Therefore for these researchers there must be an inclusion of other measures apart from student learning. They also identify the process definition of effective teaching where the actions of the instructors before and during teaching contribute to effective teaching. Stringer and Irwing (1998) identify teaching effectiveness with the improvement in student performance in the domains of declarative and procedure knowledges and motivation. For these researchers there is the notion that effective teaching allows students to learn more which will be evidenced in higher grades. Shevlin, Banyard, Davies and Griffiths (2000) indirectly define teaching effectiveness by asking how to value teachers and answer by seeing change and development in students or academic work achieved by the students as possible ways of valuing teachers. Shevlin *et al.* (2000) and Patrick and Smart (1998) see a common definition between that of effective teaching and leader charisma. Both these groups of researchers conclude however that the agreement on what is effective teaching is low.

Despite different disciplines utilising different modes of teaching (Neumann, 2001) - humanities predominantly using lectures, tutorials and seminars and sciences using laboratory and field work - Murray (1997) finds no support for arguing that the elements of teaching effectiveness show much discrepancy across different disciplines. This similarity between disciplines was also noted in a cross-university study in Spain where the attributes of the ideal teacher were similar in different degree courses (Pozo-Munoz, Rebolloso-Pacheco & Fernandez-Ramirez, 2000).

The two groups involved in the teaching process, students and teachers, can view the concept of effective teaching differently. Within definitions of effective teaching

there can be a tension between the role of the teacher or instructor and the role of the student. This can be seen in a review of the evaluation of teaching at Queensland University of Technology (Ballantyne, Borthwick & Packer, 2000). The instructors thought the effectiveness of the teaching would be improved by students becoming more involved in the learning process whilst the students wanted improvements in feedback, assessment and interaction between staff and students. Each group placed an emphasis on effective teaching on areas that were under the control of the "other side" of the teaching process. Ballantyne *et al.,* (2000, p.228) conclude, "students and staff hold quite different understandings of what good teaching is".

There is therefore a low agreement in precisely defining effective teaching but there is a common thread of having the students involved to differing degrees in the evaluation of effective teaching.

2.2 Student Evaluation of Teaching (SET)

Interest in student evaluations of teaching (SET) began in the 1920's at different North American Universities (Marsh, 1987). Both Marsh (1987) and Greenwald (1997) recognise that student evaluation of teaching has been mostly studied in the decades of the 1970's and 1980's with the peak of the research coming in the early 1980's. From that time there has been a decline in research activity into the validity of student evaluations which has lead Greenwald (1997) to infer that the major issues associated with student evaluations of teaching had been resolved by the late 1990's. Student Evaluations of Teaching are now in widespread use across North American universities where most students complete some form of evaluation and rating of the course and teaching they have experienced over the past semester (Kolitch & Dean, 1999). Saroyan

and Amundsen (2001) identify 94% of Canadian universities as using students' ratings, either alone or in conjunction with other methods, to evaluate teaching. This figure was also identified by Arnold (1991) when reviewing the literature of effective teaching. Although the bulk of the studies have come from North American Universities and the tools used by students to evaluate teaching effectiveness have been developed in North America there is evidence that SET can cross national boundaries (e.g. United Kingdom, Shevlin, *et al.*, 2000). Watkins, Marsh and Young (1987) used a tool developed in North America in a New Zealand tertiary setting and discovered that the results were very similar between the two national student groups. Further support for the cross cultural use of SET is seen in the confirmation of an evaluation form developed by Marsh for the U.S.A. tertiary sector being used in United Kingdom tertiary settings with similar results (Coffey & Gibbs, 2001).

Many opinions have been expressed about student evaluations of teaching (SET):

- their usefulness ranges from useless to useful (Wachtel, 1998),
- they are influenced by factors not necessarily aligned with teaching effectiveness (Clayton, 1998),
- they express more about the students doing the evaluation rather than the instructor or the courses being taught and evaluated (Hinton, 1993),
- the students themselves may not be reliable and consistent evaluators of teaching effectiveness (Obenchain, Abernathy & Wiest, 2001),
- they are a useful and good measure of teaching effectiveness (Marsh, 1987).

Wachtel 's (1998) review of SET noted that despite considerable research about SET, much of it confirming the reliability, validity and usefulness of SET, there is still criticism and scepticism about the use of students to evaluate teaching. However in some countries it is becoming (or is already) government policy that students provide feedback on the course and the teaching (Johnson, 2000).

For some educators whilst there is a recognition that students have a role to play in the evaluation of their teaching, the educators have a distrust of the instruments used by the students (Obenchain, *et al.*, 2001). Wachtel (1997) notes that this mistrust was due to the perceived self-interest of those marketing and selling these instruments to institutions. Those same people were the researchers who had developed the instruments and who were also declaring them to be a reliable and valid manner by which to measure teaching effectiveness.

Students are however recognised as stakeholders (Watson, 2003) and customers (Ballantyne *et al.*, 2000) and student evaluations are seen as part of improving the service to and satisfaction of the customers. But for other educators student ratings distort teaching in favour of the lowest common denominator (Clayton, 1998). For some critics of student evaluations of teaching, students do not have sufficient experience or subject expertise to be able to adequately judge effective teaching (Martin, 1998). Kerridge and Mathews (1998) suggest that students could be considered similar to a hospital patient asked to rate the standard of care provided by the doctors. The patient knows nothing about medicine and must therefore base the ratings on peripheral services such as meals, communication skills, décor of the room, rather than core medicine services. Other researchers recognize that students in most subjects experience many different teachers in

their many hours in the classroom thus having the required knowledge and expertise to make evaluations of teaching (Perry, 1997). The students themselves are not reluctant to provide evaluations of their instructors (Spencer & Schmelkin, 2002).

Watson (2003) provides three advantages in having student feedback:

- it is the view of the person participating in the learning process
- it is direct
- information can be provided for prospective students.

Overall Abrami *et al* (1997) describe the use of students to evaluate teaching as being reasonable and suggest that instructors and students have a high degree of agreement as to what are considered important characteristics of teaching.

One corner of the debate surrounding the use of SET is whether effective teaching can be marked with one global score (Cashin & Downey, 1992) or a recognition that as teaching is multi-dimensional there needs to be the measurement of many different dimensions of teaching (Marsh, 1987; Marsh & Roche, 1997). For those such as Cashin and Downey (1992) who see a global score as being attainable, do not reject the multi-dimensional nature of teaching but compare a single global score for teaching with that single grade the instructors give to students of their course which evaluates the students' learning, a multi-dimensional task. Cashin and Downey (1992, p.569) hypothesize that "a better criterion measure of the instructors effectiveness will correlate highly with an item such as 'Overall. I rate this INSTRUCTOR an excellent teacher'". Marriott and Litzelman (1998) found that a single global score reliably rated the medical instructors being rated by students. They also noted there were administrative advantages in collecting a global score of teacher effectiveness.

Marsh (1987; 2001) notes that as the nature of teaching is multi-dimensional, there is no one criterion of effective teaching, therefore it requires multiple ratings where an instructor can rate highly in one area and not so highly in another. Attending to only the scores from global questions is seen by Brightman, Elliott, and Bhada (1993) as one of the weaknesses of student evaluations.

Part of the discussion of single global rating versus multiple ratings is the purpose for which the rating will be used. For formative purposes to improve teaching Cashin and Downey (1992) note that many ratings covering different areas would be useful so improvements in the instructor's work can be made in the areas that need further work. But these researchers contend, for summative purposes, a global score is adequate. Overall, there appears to be general acceptance that student evaluations can be helpful in improving teaching standards. McKeachie (1997a, p. 1219), argues for student evaluations being "the single most valid source of data on teaching effectiveness." Marsh and Roche (1997, p.1190) go further and suggest there is a "lack of support for the validity of any other indicators of effective teaching." So not only is it appropriate to use student evaluations to assess effectiveness of the instruction performed by teachers, there is perhaps nothing else to challenge or replace it.

The greatest discussions about SET come when it is discussed as to how the evaluations should be used (McKeachie, 1997b). Kember, Leung and Kwan (2002) list three purposes for which student evaluations may be made:

- formative – for the improvement of teaching
- summative – for making decisions about employment conditions of the instructor
- requirement – the institution requires a student evaluation to be made

Using student evaluations for summative purposes – deciding promotions and tenure and industrial purposes – generates the most discussion about the applicability and relevance as against when the evaluations are used for teaching improvement purposes, i.e. formative purposes (Wachtel, 1998; Obenchain *et al.*, 2001). One study at an Israeli University found staff somewhat ambivalent about student evaluations whilst a further study of staff found that the formative purpose of SET was paid little heed to as no more than 10% of the lecturing staff made large changes to their teaching as a result of feedback from SET (Nasser & Fresko, 2002). Students involved in evaluations thought formative purposes were the most important task of SET, however the students are uncertain as to whether their evaluations are taken heed of by the instructors (Spencer & Schmelkin, 2002). Amongst the five different uses of student evaluations described by Marsh (1987), the most common purpose is feedback for the improvement of teaching (Marsh & Dunkin, 1997).

2.3 Evaluating Teacher or Course

Abrami *et al.* (1997) view SET as measuring student satisfaction with teaching rather than what has been learned by a class of students. Other researchers found that the student ratings were a measure of the teacher rather than the course itself (Marsh, 1987; Marsh & Dunkin, 1997; Weimer,1997). Further, Marsh (1987, p. 278) believes that there is no support for the idea of using students to rate the course as student ratings of courses are "primarily a function of the instructor who taught the course." Evidence of the distinction between teacher and course is seen in a study of the evaluation form used at the University of Queensland (Timpson & Desley, 1997) where students found it confusing evaluating the course and instructor in one form. At the Queensland University

of Technology students are asked to evaluate the effectiveness of teachers using one form and provide an evaluation of the actual course itself - as distinct from the inputs from an individual – with a different form (Ballantyne. *et al.*, 2000). Marsh (1987) suggests this distinction between instructor and course is due to the autonomy university lecturers enjoy – (as most of the research appraised by Marsh arose from tertiary institutions) - but he could not generalise this across to courses where the teacher has little or no autonomy, the situation with flight instruction. Although student evaluations may and sometimes do get used for evaluation of courses Marsh and Dunkin (1997) found that the evaluation still measures the effectiveness of the instructor who teaches the course rather than the course itself.

2.4 Reliability of SET

Watkins, *et al.* (1987) in reviewing the literature on the value of student evaluations note that early studies supported the reliability of SET. Abrami *et al.* (1997) acknowledge that student ratings of teaching effectiveness are stable and consistent. When students have been asked to evaluate effectiveness of the teaching after a period of time after the completion of the course the ratings are stable (Marsh & Hocevar, 1991). These researchers examined results from evaluations performed consistently on the same instructors over a period of thirteen years and found the means of the ratings to remain stable. There was little or no change in that period from the perspective of the students. The period of time over which ratings have remained stable have included at the completion of the course, at graduation and after a period of employment (Marsh & Dunkin, 1997). Longitudinally, the means of student ratings of instructors have been found to be stable.

The internal consistency in SET among items measuring the same area of effective teaching was found by Marsh (1987) to be consistently high. Obenchain *et al.* (2001) whilst acknowledging the reliability of SET and the reliability of instruments used as described in earlier studies (especially Marsh, 1987), examined the reliability of the individual student as an evaluator. These researchers used a group of student teachers and asked the students on one instrument to give one global rating for the lecturer of the course and with another instrument provide rankings for the same lecturer on an attribute scale. Although the assumption was that there would be a close match between the global score and the mean of the attribute rankings, this was only true for less than one third of the students who provided evaluations of their lecturers. They found that students were not reliable evaluators of effective teaching. Despite this, there appears to be common agreement on the reliability of SET. Hobson & Talbot (2001, p.29) conclude "the research on SETEs has provided strong support for their reliability, and there has been little dispute about it

2.5 Validity of SET

For any measure to be valid it must measure what it purports to measure. To validate student evaluations they must be related to a variety of indicators of effective teaching, in other words, establish construct validity (Marsh, 1987). Greenwald (1997) identified three kinds of studies that have provided support for the construct validity of SET: multi-section validity studies, path-analytic studies and multitrait – multimethod studies.

Marsh and Roche (1997) describe teaching as multi-dimensional, therefore the SET should relate to a wide range of indicators of effective teaching. As there are

different measures of teaching effectiveness SET are more effective when multi-dimensional scores rather than one single score averaged across a class of students (Marsh, 1987). There is however little understanding of how the dimensions are related.

d'Apollonia and Abrami (1997) whilst acknowledging that teaching is multi-dimensional, believe that different items on the ratings form are not sufficiently distinct and thus student ratings come down to measuring one global score of general instructional skill. Therefore they looked at using many sections – multi-sectional - of a course to validate SET. This design they conclude "provides most generalizable evidence for the validity of student ratings" (d'Apollonia & Abrami, 1997, p 1201).

As discussed above there is no agreement on what effective teaching is and therefore no agreement on criteria or markers of effective teaching that can be used to validate SET (Marsh, 1994). Bosshardt and Watts (2001) identified possible indicators of effective teaching as objective tests, instructors and administrators evaluations. Stringer and Irwing (1998) look to variables involving student learning or overall course ratings to validate SET. Marsh (1987) notes that there is good agreement between student evaluations and instructor self-reports but poor agreement between peer's reviews and any other indicator of effective teaching. The factors to be measured should provide an understanding of teaching effectiveness. The research carried out by Marsh (1987) discovered that across different subjects and course levels (graduate and undergraduate level) there were nine factors to be used to measure teaching effectiveness.

Martin (1998) questions the validity of SET because they are the opinions of students who do not have the expertise or qualifications to rate many of the conditions of effective teaching. Martin (1998) believes Deming's (1986) theory of management

perspective should be applied to the tertiary teaching setting where – as with any group of workers – evaluating or ranking is misconceived as "94% of the variation in any system is attributable to the system, not to the people working in the system" (Martin, 1998. p. 1082).

Hinton (1993) in examining SET from a discipline other than education claims that current student evaluations are more akin to a public opinion survey with the resulting different approach to the dimensions of reliability and validity. Hinton (1993) instead argues for SET to be considered as a survey where the information provided by the students gives an insight into the students rather than the instructors "If we want knowledge about the characteristics and quality of faculty teaching, we should study a random sample of faculty" (Hinton, 1993, p. 567). The survey doesn't presume that students are objective observers and raters. Surveys have different reliability and validity criteria. Therefore, argues Hinton (1993), SET are a survey to inform us about those who undertake the survey – students - rather than the supposed subjects of the ratings – instructors.

Marsh (1987) argues against the use of criterion-related validity especially with student learning being the criterion of effective teaching. Marsh (1987) does not believe that student learning and teaching effectiveness necessarily need be synonymous although agreeing that student learning is one indicator – even an important indicator – of effective teaching. Further he argues that student learning itself is a construct. Overall, Watchell's (1998) and Marsh's (2001) reviews of SET find them to be reliable and stable and a relatively valid measure of teaching effectiveness as against other indicators of effective teaching.

2.6 Factors which may bias student evaluations of teaching

Despite SET generally being considered to be reliable and valid there are still many instructors who believe there are factors not associated with the instructors behaviour or effective teaching which bias the evaluations made by students. One report described these biasing factors as "pander pollution" (Clayton, 1998). Instructors interviewed by Clayton (1998) believed that to gain better ratings they are required to dilute their own standards and expectations, to mark leniently and put on a bit of a show by being funny and entertaining. However, Marsh (1987, p. 309) describes this search for biasing factors as not only a witch hunt but has been "extensive, confused, contradictory, misinterpreted and methodologically flawed". For a factor to be a biasing influence on SET Marsh and Roche (1997) contend that there must be a validation, i.e. it has a similar effect on other indicators of teaching effectiveness and other forms of assessment. A number of supposedly biasing factors have been identified by instructors (Marsh, 1987):

2.6.1 Grading Leniency and Student's GPA (Grade Point Average).

This bias is viewed by many instructors as the instructor achieving good ratings by marking leniently and giving good grades (McKeachie, 1997b). It is considered by Watchtel (1998) to be one of the more controversial factors considered to bias SET. Greenwald (1997) discovered a positive relationship between grades students received and ratings of the instructors through reviewing experiments performed mostly in the 1970's. In the experiments grades were manipulated up or down and evaluations of instructors were observed to follow the same path. Greenwald (1997, p. 1183) concludes; "grading leniency-strictness affects ratings ...(this) has been supported with some clarity

in virtually all published experimental tests". Marsh's (1987) studies indicate there is a positive relationship between expected grades and student ratings. However, as d'Apollonia and Abrami (1997) state, if grading leniency makes the student work harder, increasing their learning, then this factor isn't a biasing factor. Marsh and Roche (1997) believe that if there is any biasing of SET by grading leniency the effect is at best unsubstantial.

2.6.2 Instructor popularity

Shevlin *et al.* (2000) examined the biasing effect of charisma, explained as the power of leadership offered by the instructor. The ratings were demonstrated to have been affected by the students' view of the lecturer where variation was explained by the charisma of the instructor rather than teaching effectiveness. This, claim these researchers, is so strong that they conclude "students rate specific features of teaching on the basis of a global evaluation. That global factor is lecturer charisma" (Shevlin *et al.*, 2000, p. 403).

2.6.3 Experience of instructor

Whilst it has been noted that teaching assistants are rated lowlier than existing lecturing staff, once the teaching assistants are removed the rank or experience of an instructor do not influence student ratings of an instructor (Marsh, 1987).

2.6.4 Student interest in subject

The interest the student has in the course prior to the commencement of the course has been thought to bias student evaluations. Marsh (1987, p. 135) argues that prior

interest is not so much a bias but "a variable that influences some aspects of effective teaching". The reason some may have for considering prior interest as a biasing factor is that it is something that the instructor has no control over and thus is unfair if the ratings are used for human resource decision making.

2.6.5 Course workload and difficult

Proponents of biasing factors claim that for instructors to get better ratings they should make the course workload easier and lighter (Clayton, 1998). However those courses that are perceived to be more difficult tend to get higher ratings (Marsh 1987). One study found the workload – SET relation to be positive (Marsh & Roche, 2000).

2.6.6 Level of Student

There is evidence that the level of the course being undertaken will influence ratings. Graduate level courses tend to have higher ratings from both students and staff (Marsh, 1987). Santhanam and Hicks (2002) found that class or year level did in fact influence ratings. The further through the course or "older" the student the higher the ratings of teachers apart from a slump at year two. Thus as the student progresses through the course there are changes in the perceived effectiveness of the teaching of the course. The student would appear to become socialised and molded into the accepted model and course delivery formed by the existing structures such as the curriculum, the instructors themselves and their mode of instructing, and the assessment regime. Within aviation there is evidence of this socialisation and fitting into expected patterns of behaviour and learning. For example, Turney, Bishop, Karp, Niemczyk, Sitler and Green (2002) examined the problems women, who are very much the minority in all branches of

aviation including training, faced when they entered a tertiary aviation programme. Part of the study compared the responses of male and female to the same questions. The women who were student or private pilots gave responses that were not only different to their male counterparts but were also different to women who were further through the programme. The researchers describe this as an "adaptive process" whereby the women pilots who have gained a degree of experience within aviation (and had progressed further through their training) have attitudes more in keeping with the prevailing culture dominated by males. The women are molded and socialised to respond more in keeping with the culture than when they first joined the training programme. The biasing factor of the level of the student could be evidenced within aviation as more experienced women students evaluate the teaching effectiveness of their flight instructor differently as they become immersed and influenced by the prevailing attitudes contained within the aviation culture. Spencer and Schmelkin (2002) would not call the differences in student level a biasing factor but the differences in responses due to the influence on a student who has been in the education system for a longer period of time. For these researchers, class level was one of the variables expressed by different groupings of students when exploring student attitudes to SET. In contrast to Santhanam and Hicks (2002) Spencer and Schmelkin (2002) attributed this difference due to class level to scepticism and becoming less idealistic as the students progressed through the course.

2.6.7 Gender

The gender of either students or instructor has been found to have little influence on ratings (Marsh, 1987; Wachtel, 1998). A Spanish study by Pozo-Munoz (2000) found no gender differences in the descriptions of an ideal teacher and Santhanam and Hicks

(2002) reviewing gender bias studies found no or very little gender bias affecting SET. The gender role may play a role in affecting student evaluations whereby students rated more highly those instructors who displayed both masculine and feminine traits (Freeman, 1994).

2.6.8 The chosen discipline

The discipline within which the evaluating student is studying may have an effect on ratings. Humanities disciplines are rated higher than social and physical sciences by students (Marsh, 1987; Neumann, 2001). However, Santhanam and Hicks (2002) noted that whilst humanities students rated their course higher they rated the teaching effectiveness of their teachers lower than their counterparts in sciences.

2.6.9 Instructor personality

Instructor personality is believed by some to be a biasing factor although there has been little study into the area. The Dr Fox experiment provides some insight into the biasing effects of instructor personality. The Dr Fox experiment (Naftulin, Ware, & Donnelly, 1973) saw an actor schooled in a nonsensical topic who then gave a one hour lecture to a group of health workers who were undertaking a course in training other health workers. The actor – Dr Fox – was introduced to the group as a bona fide academic with appropriate credentials including a list of publications. The object of the lecture was to entertain, be humourous without delivering much if any worthwhile content on the topic of "Mathematical Game Theory as Applied to Physical Education" (Naftulin *et al.,* 1973, p. 631). At the end of the presentation, which included a question and answer session, the recipients of the lecture were given a questionnaire to measure

their satisfaction with the lecture. All responses were positive. The "lecture" was repeated with two other groups of professional people with similar results, with one respondent to the questionnaire replying that they had read some of Dr Fox's non-existent publications. Naftulin *et al.* (1973, p. 633) concluded that these students – who were professional people, not undergraduate students – were "seduced into an illusion of having learned (by) a style of authority and wit" and they "failed as competent crap detectors". Marsh (1987, p. 331) defines the Dr Fox effect as "the overriding influence of instructor expressiveness on students' evaluations of college/university teaching ..... mean(ing) that an enthusiastic lecturer can entice or seduce favourable evaluations, even though the lecture may be devoid of meaningful content".

Despite many researchers – including the authors of the original Dr Fox study - identifying many problems with the Dr Fox study, critics of student evaluations have picked up on entertainment, enthusiasm and expressiveness as being biasing factors upon student evaluations (Marsh, 1987). Further researchers developed extended Dr Fox studies where an actor was video-taped with different levels of content and different levels of expressiveness. Ware and Williams (1979) (cited in Marsh 1987) concluded that the differences in expressiveness had more effect on evaluation variance than did content. Further analysis showed that if students were told they were to be examined on the content of the lectures then the Dr Fox effect diminished. If there was no testing then Dr Fox and expressiveness become relevant to student evaluations (Marsh 1987). Martin (1998) sees a danger that the Dr Fox effect is centred around a teacher who is deemed to know all and spends the time telling all which is inconsistent with more recent

approaches to teaching which centres around the student and has the student being actively involved in their learning.

Abrami *et al.* (1982) (cited in Marsh, 1987) conducted a meta-analysis of Dr Fox studies and concluded:

> "Expressiveness manipulations had a substantial impact on student ratings and a small effect on achievement while content manipulations had a substantial effect on achievement and a small effect on ratings..... the rating factors that were most logically related to the expressiveness manipulation were most affected by it ..... none of these interactions accounted for more than 5 percent of the variance in student ratings" (p. 333).

One of the issues with the Dr Fox phenomenon is the contrast effect between different instructors (Marsh, 1987). Instructors begin to feel that that the SET becomes a ranking device to list those instructors who are better than others. The suspicion is that the instructor who can provide better entertainment value will be rated higher than others. Naftulin *et al.* (1973, p. 635) even suggest that actors should be trained up "to give legitimate lectures as an innovative educational approach toward student-perceived satisfaction with the learning process". This gives rise to many instructors viewing student ratings as nothing more than popularity contests (Li-Ping Tang, 1997; Obenchain *et al.*, 2001).

Marsh and Dunkin, (1997) contend that if there is to be a biasing factor there must be a construct approach to its validity. Not only must a biasing factor be correlated with SET, it must also not be correlated other indicators of effective teaching. Studies reviewed by Marsh (1987) indicated that factors thought by instructors to possibly

negatively influence SET did not actually have much if any effect. Any variances in SET were not accounted for by the perceived biasing factors as discussed above.

2.7 Summary of SET

As effective teaching is difficult to define it is equally difficult to devise a measure of effective teaching. The evaluations of teaching by students, a commonly used method of evaluation, have been found to be an appropriate way to evaluate teaching. Marsh (2001, p. 184) in reviewing the research on SET describes them as:

multi-dimensional

reliable and stable

primarily a function of the instructor who teaches the course rather than the course that is taught

relatively valid against a variety of indicators of effective teaching

relatively unaffected by a variety of variables hypothesised as potential biases

2.8 Attributes of Effective Teachers

As SET is a reliable, valid and oft used means of measuring effective teaching, a review of the results from SET can give indications of what attributes are possessed by teachers who provide effective teaching. From different national cultures and differing academic disciplines as well as class levels there are many studies reporting the results of students rating the attributes of effective teachers. (e.g. Abrami *et al.,* 1997; Amin, 2002; Bosshardt & Watts, 2001; Cashin & Downey, 1992; Das & El-Sabban, 1996; Irby, 1978; Irby, Gillmore, & Ramsey, 1987; Irby & Rakestraw, 1981; Jackson *et al.*, 1999; Jameson

Beaux, 2000; Li-Ping Tang, 1997; Litzelman, Westmoreland, Skeff & Stratos, 1999; Marsh, 1987; McLean, 2001b; Murray, 1997; Patrick & Smart, 1998; Pozo-Munoz *et al.,* 2000; Sander, Stevenson, King & Coates, 2000; Riesenberg, Biddle & Erney, 2001; Stringer & Irwing, 1998; Watkins, Marsh & Young, 1987). There are many different forms used to evaluate effective teaching measuring different dimensions of effective teaching (Abrami, *et al.,* 1997). The results are as diverse as the setting from which they arise. However, to some extent there is a commonality across cultures and disciplines regarding the attributes of an effective teacher that students regard as playing an important part in effective teaching.

Most studies assessed students' review of teaching effectiveness through the students' retrospective evaluation of the teaching they have received. Some studies have used SET to investigate the students' perception of the ideal effective teacher. Irby (1978) assessed the desirable and undesirable characteristics medical students had of their ideal clinical teacher. Sander *et al.,* (2000) used student responses to learn what their expectations of teacher attributes are with a prospective viewpoint. This cross discipline study in United Kingdom with medical, psychology and business studies students surveyed in their first week of university study found that the expectation of teacher attributes in all three disciplines was teaching skills, followed by approachability, knowledge, enthusiasm and lastly organisation.

2.9 Self-Reports of Effective Teaching

Evaluation of teaching at the tertiary level can also be made by self-reports from the instructors (Centra, 1993). The benefits of using self-reports include that they are collectable in all educational settings, they can be stimulating in getting a self-reporting

teacher to alter teaching methods and contribute a view into how a teacher understands their own teaching (Marsh & Roche, 1997).

Problems with self-reports have included inflation, self-serving and expected correct responses being provided by the lecturer (Fisher, Alder & Avasalu, 1998). It has also been reported that self-reports can be influenced by SET, where initial overestimations of the instructors' teaching effectiveness are lowered once results from SET are known (Marsh & Dunkin, 1997).

Despite these problems the beliefs about teaching held by instructors are thought to influence student outcomes (Lehrer, Moore, Telfer & Freeman, 2000). As Ballantyne, *et al*, (2000, p. 231) state, there is a "need for staff and students to share their perceptions and so gain a better understanding of the whole process and context of teaching and learning". Self-reports made by lecturing staff and SET can be compared to help the learning process, and teaching effectiveness (Das & El-Sabban. 1996). The value of comparisons between student evaluations and those of the teaching staff help to highlight unrecognised problems that could hinder effective teaching. Killen (1994) identifies lecturers being able to increase the positive factors and lesson the sway of negative factors identified in student and lecturers evaluations.

There is disagreement on how close the match is between the evaluations of students and teachers. Similar factors regarding effective teaching were identified by staff and students when using the same form (Marsh, 1987). However Hativa, *et al.* (2001) noted teachers' self-ratings of effective teaching were not matched by those evaluated by students. This was especially so for teachers who evaluated highly their own teaching. A study investigating the benefits of publishing student ratings of teacher effectiveness

found that teachers and staff differed in their views (Howell & Symbaluk, 2001). Responsibility for academic success was placed on the other group involved in the teaching and learning process (Killen, 1994). The study of teaching staff and students by Ballantyne, *et al.*, (2000) found considerable differences between the two groups. For students, improvements in the effectiveness of teaching were in areas that pertained to the staff such as assessment and feedback. For the teaching staff, improvements were required of the students in areas such as student participation and learning.

Moore *et al.* (2001) examined the differences between aviation students and flight instructors regarding learning (surface, deep or achieving) which is influenced by the flight instruction the students receive. Failure to align students and flight instructors approaches to learning was identified as leading to problems such as dislocated objectives (what is believed to be taught and what is actually taught), instruction rather than learning (teaching what the student already knows) and pedagogy rather than andragogy. The largest difference between the students and the flight instructors was that were more concerned about just passing examinations than were the flight instructors. The researchers maintain the importance of an alignment in the beliefs about learning held by students and the beliefs held by flight instructors about teaching. Specifically, beliefs held by flight instructors influence the students (Moore *et al.,* 2001). A difficulty with the reports from this study is that the flight instructors did not have instructional interaction with the students. The students were pilots from one airline whilst the instructors instructed elsewhere.

2.10 SET and Flight Instruction

There is a considerable amount of research literature on the use and effectiveness of SET and those factors students rate as important in their instructors from many different academic disciplines (Greenwald, 1997). Unfortunately flight instruction has not been one of the disciplines much studied. The research material that is specific to flight instruction is small to non-existent. If there is a list of the attributes required to succeed at flight instruction it is based on general principles of effective teaching as Trollip and Jensen (1991) admit.

For flight instruction Trollip and Jensen (1991) list professionalism and being prepared as the requirements for effective flight instruction. Professionalism for these authors is seen as setting a personal example, maintaining a current knowledge of flying domain, being punctual and informing the student if any changes should happen to booking times and being polite to the students, (e.g. not yelling or swearing at the student, not losing temper with the student). Koonce (2002) also believes professionalism to be an important part of being a flight instructor and even uses the word professional as a mnemonic to state qualities of flight instructors. Further qualities deemed important by Koonce (2002) include communication and teaching skills, technical knowledge and personal qualities. Hawkins (1993) also details desired qualities of a pilot within an airline environment who could become a flight instructor – as against the General Aviation setting where most flight instructors work. These include flying abilities and technical knowledge, personality traits as well as teaching and communication abilities.

These are qualities and attributes expected by the authors to be useful to flight instruction and not based on student ratings. They tend to be a useful wish list that could be applicable to most areas of instruction, not necessarily unique to flight instruction.

These stated attributes and qualities appear to be based on the authors' opinions as opposed to students' reviews of flight instructing or other research into what makes an effective flight instructor.

The material concerning the views of flight instructors themselves is also small. Arnold (1991) in his study of successful flight instructors noted the instructors recorded satisfaction from the teaching of students as their most enjoyable part of the instructing process. Working with people was rated second and flying was rated sixth. When asked to describe the personality characteristics of successful flight instructors, the instructors nominated patience (61%) as the characteristic most needed and teaching skills second with (19%).

2.11 Medicine and Aviation

With a paucity of material regarding the attributes required by an effective flight instructor, a review of the literature relating to the instruction of students in a discipline considered to have similar characteristics to aviation would seem appropriate. Almost all of the studies using SET are carried out in a standard university course in which the teaching modes may include only lectures or variations and sometimes lab work and as such there is little or no difference in the description of effective teaching across disciplines (Murray, 1997). The teaching context for flight instruction is very different from that investigated in these studies, with its requirements for motor skills, decision making and aviation knowledge and the need to relate to a student in an intense one-on-one setting. While flight training combines components found within the teaching of other disciplines such as ground lectures, personal study and "lab" work, in flight training a mistake can be fatal. Quite literally the student's life rests with the flight instructor. The

quality of the instruction provided to the student during dual instruction contributes to keeping the student safe during solo flight. The result of the flight training – the licenced pilot - then has the lives of the passengers relying on the piloting skills tutored by the flight instructor. There is an immediacy to life within flight instruction and its outcomes that is very different from almost all other disciplinesThis difference in evaluating teaching effectiveness has also been noted in the clinical teaching of medical studies due to the uniqueness of clinical teaching (Irby, 1978).

As stated earlier there is a growing awareness of the close similarities between medicine, and aviation; they are both safety critical industries (Sexton, *et al.*, 2000). These researchers note that the University of Texas has spent over twenty years researching teams that work within a safety critical environment, which include medical teams and cockpit teams. Comparisons and similarities between the two disciplines include the team make-up and the need for the team to function without error (Sexton, *et al., 2000*). Aviation teams have captains with support from first and second officers and surgical medicine teams have surgical and anaesthetic consultants supported by residents and nurses. Helmereich (2000b) further describes similarities between doctors and pilots as needing to not only engage in teamwork but also to operate in complex environments. "Pilots and doctors have common interpersonal problem areas and similarities in professional culture", (Helmereich, 2000b, p. 783).

As well as strong similarities existing between aviation and medicine in operating and team cultures, Helmereich (2000a) describes the similarities in individual professional cultures between doctors and pilots. Both groups are self-selected members of a perceived elite profession, much effort has gone into the individual reaching their

current position, both groups have a strong sense of personal invulnerability so much so that members believe their decision making is as good in emergency situations as in normal situations.

Reason (2000) describes two approaches to error, person approach where the sharp end operator (e.g. surgeon or pilot) is at fault or a system approach whereby latent errors combine with active errors (those errors committed by operator). The Swiss cheese model of system accidents relies on defensive layers created by being engineered, provided for administratively, or provided by people such as pilots or surgeons or anaesthetists (Reason, 2000).

Howard, Gaba, Fish, Yang and Samquist (1992) describe the complex and dynamic world of anesthesiology and compare its similarities to other domains but especially aviation. As with aviation, 65 –70% of accidents in anesthesiology are attributable to human error. Because of the similarities between the two fields of medicine and aviation, these researchers applied the training techniques of aviation in the area of crisis management to anesthesia training. Helmereich, (2000b) also believes that medicine can help improve patient safety by learning from how aviation manages error. Further similarities can be seen between medical teaching and flight training. Medical teaching has a mix of teaching in the classroom setting and in a clinical setting, (Das, El Sabban, 1996), which approximates to flying with its twin setting of ground and flight teaching environments.

With the close similarities between medicine and aviation an examination of the material describing the attributes of effective teachers within a medical setting can be expected to help with the understanding of attributes of an effective instructor in the

aviation setting. McLean, (2001b) surveyed second year medical students in a South African Medical School about their first and second year experiences of teaching staff within the medical school. They were interviewed by the researcher with a mix of open and closed questions. As part of the interview process the students were asked to identify and rank the attributes they thought were required of an effective medical instructor. The students, who were in their second year of medical studies, rated communication skills along with rapport with students to be the most important. In the middle of the student preferences was subject knowledge. The technical skills or "nuts-and-bolts" of teaching such as being on time and being organised were the lowest rated attributes (McLean, 2001b). The students who participated in this study had yet to experience patient contact in clinical teaching and work.

The Stanford Faculty Development Program (SFDP) is often used in medical studies of SET to investigate and improve clinical teaching effectiveness (Vu, Marriott, Skeff, Stratos & Litzelman, 1997). In medical settings it has been demonstrated to be reliable (Marriott & Litzelman, 1998) and have strong construct validity (Litzelman, Westmoreland, Skeff & Stratos, 1999). The attributes that the SFDP measures include stablishing a positive learning climate, controlling the teaching session, communicating goals, enhancing understanding and retention, evaluation and feedback, and promoting self-directed learning.

Vu *et al.*,'s (1997) study of interns, sub-specialists and residents and their ratings of teacher attributes used a Likert type survey on the SFDP. These students were based on wards in teaching hospitals for a four week block. The results indicate that the interns and sub-specialists rated learning climate and self directed learning significantly higher

than other attributes, while control of session and communication of goals were significantly lower. Vu *et al.*, (1997) contrasts this high rating of teaching skills to earlier findings from SET of medical teaching where the interpersonal skills displayed by the teacher were more highly rated. Irby and Rakestraw (1981) reported that interpersonal relations and enthusiasm along with clinical supervision, knowledge and demonstration of clinical skills were the factors most likely to correlate with effective clinical teaching in a study of the clinical teaching of obstetrics and gynecology. This finding of the importance of the teacher - student relationship was reinforced in a later study in which relational skills were considered important by the students undergoing clinical teaching (Irby, Gillmore & Ramsey, 1987). Involvement with the student and personal interest and concern in the personal and professional development of the student by the clinical teacher were also considered important, (Irby, *et al.*).

A study of the ideal medical teachers in a classroom setting as rated by both students and faculty saw a close alignment between the two groups (Das & El-Sabban, 1996). Both groups valued a willingness to help the student as being the most highly rated attribute of an effective teacher. There were differences between faculty and students regarding the importance of personal qualities of the effective instructor – rated more highly by the students, and the importance of teaching skills – rated more highly by the faculty (Das & El-Sabban, 1996).

Risenberg, Biddle and Erney (2001) compared the ratings of third year medical students with their tutors' ratings of what was more important in the training of students for ambulatory care; the site for the teaching versus the characteristics of the tutors. The researchers found both groups agreed that the characteristics of the tutors were more

important to the education and training of the students. Further, there was close agreement between the students and the tutors that tutor characteristics of enthusiasm, communication skills and clinical skills were important.

The similarities between the tasks and cultures of medicine and aviation may allow the research into the attributes of effective medical teachers to shape and guide the study of attributes of effective flight instructors. Reference to medical studies is required due to the lack of research conducted into flight instruction. A review of the literature regarding effective medical teachers attributes showed teaching skills, including enthusiasm and clarity of communication, and student interaction with the teacher were attributes of teachers that were consistently rated most highly by students (Risenberg *et al.*, 2001).

2.12 Defining the Research Problem

Chapter one referred to the role of the flight instructor, the difficulties in attempting to be both teacher and pilot and the need to achieve competence in both roles. Due to credential inflation the position of flight instructor is used as a stepping stone (Henley, 1995) by newly qualified commercial licence holders to gain the necessary credentials the airlines deem necessary for acceptance into airline flying. This has lead to increasing problems with flight instruction and flight instructors as the inexperienced pilots seek to gain the necessary experience required for employment (Woodhouse, 2001). The least experienced members of the piloting profession have become the instructors of the *ab initio* pilots relying on methods of instruction developed by the military in the two world wars. These methods are handed down from generation to generation of instructors thus developing "folkways" in flight instruction (Telfer, 1993b).

There is wide-spread criticism of the standard of flight instruction and the effectiveness of the flight instruction received by the students. Civil Aviation Authority of New Zealand (2001) has identified nine different areas of concern regarding flight instructional deficiencies in its drive for greater safety in New Zealand aviation. Despite this there has been little attempt to define who should be the flight instructor and the attributes that the flight instructors should possess to provide effective flight instruction. This need to improve the effectiveness of teaching and instructing is not limited to aviation within New Zealand. There is a desire to identify and improve the effectiveness of teaching across different nationalities (e.g. Li-Ping Tang, 1997), and in many different disciplines (e.g. Worthington, 2002). There has been a recognition amongst teachers at tertiary level education that systemically, teaching is unrecognised, undervalued and unrewarded by institutions (Centra, 1993; Neumann, 2001) and yet it would seem to be a core business or activity.

The literature reviewed indicated the difficulty in defining effective teaching. In attempting to define effective teaching there is increasing recognition that students and their learning is a large part of effective teaching. Centra (1993) lists three methods for evaluating effective teaching; evaluations by students, reviews by peers (or colleagues or chairs or faculty development specialists), and self-reports from the teachers themselves. Evaluation of teaching by students and the teachers themselves are therefore two methods of evaluating teaching effectiveness.

In the main students evaluations are accepted as useful for formative evaluations of teachers although there are disagreements about their use in summative evaluations including promotion and tenure decisions (e.g. d'Apollonia & Abrami, 1997).

Students' evaluations of teaching are considered to be reliable. The validity of students' evaluations is more difficult to establish not least because of the difficulty in defining effective teaching and therefore establishing other measures of effective teaching against which student evaluations can be correlated (Marsh & Roche, 1997). However SET is considered to be the most valid measure of teaching (e.g. McKeachie. 1997). Suggested biasing factors are found to have at most a small effect on student evaluations of teaching (Marsh, 1987).

Students' and their evaluations are an acceptable method of identifying effective teaching and the attributes of teachers who provide this teaching, (Marsh, 1987; Centra, 1993). While self-reports from the teachers are also considered to be a method of evaluating the effectiveness of the provided teaching, they are not considered as valid a measure of teaching effectiveness as are student evaluations (Marsh & Roche, 1997). Comparing the evaluations of the two groups involved in the teaching and learning process can highlight differences that that if attended to can help improve the effectiveness of teaching (e.g. Ballantyne *et al.*, 2000).

Although there have been many studies conducted into the establishment of effective teaching attributes in many academic disciplines and in different cultures there is a paucity of material concerning flight instruction. The lack of research in the area of effective flight instruction can be overcome by examining the findings from other disciplines. An examination of the material regarding medical students, an area of study that has been considered similar to aviation, found that personal attributes, the ability to establish rapport with the students and teaching skills were attributes considered by students to be important for effective teaching (Risenberg *e tal.*, 2001).

The curriculum of what should be taught in *ab initio* flight training is known, but how the curriculum should be delivered and who should be conducting the flight instruction is less well understood. It is the attitudes, qualities and attributes of the effective flight instructor that this study is concerned with. The distinction is made between personality traits, which are seen as being fixed and attitudes and attributes which are malleable and can be changed (Chidester, Helmreich, Gregorich & Geis, 1991).

The present study will use evaluations by student and flight instructors to investigate the perceived attributes of an effective flight instructor.
Both groups have an interest in improving the standards of flight instruction and as teachers and students in the learning process have been identified by Centra (1993) as being part of the processes of evaluating teaching effectiveness.

The aims of the study are:

- to identify the attributes of flight instructors that students believe are both effective and ineffective in helping them learn to fly
- to identify the attributes of flight instructors that flight instructors themselves believe are both effective and ineffective in helping students learn to fly
- to compare and contrast the attributes of an effective flight instructor as identified by students and instructors
- to identify any differences in perceptions of attributes of effective flight instructors between *ab initio* students and more experienced students and instructors

Chapter Three

METHODOLOGY

3.1 Design

Data was collected by cross-sectional survey method. Survey material was formulated through a review of literature surrounding the evaluation of effective teaching by students and teachers. The factors underlying effective teaching have been found to be similar across different disciplines (Marsh, 1987). Consultation was undertaken with subject matter experts including senior instructors in both the general aviation and airline sectors of the aviation industry to assess and review survey materials for content and accuracy.

3.2 Participants

Both flight instructors and students training towards pilot licences were obtained through the Flight Systems Centre of a university aviation degree programme. The entire population of both sets of participants studying and instructing within the degree programme were offered the opportunity to complete the survey. Twenty two flight instructors and 78 students (response rate 87.6%) responded providing a pool of 100 respondents. A description of the participants is provided in the results section.

3.3 Procedure

Students were approached to participate in the study by a member of the lecturing staff, independent of the researcher, at the commencement of a ground lecture. Each participant was provided with an information sheet along with the survey form. The

information sheet introduced the researcher and outlined the aims of the study as well as information regarding participants' rights for those taking part in the study. The rights and responsibilities of the students in participating as well as those expected of the researcher were also explained verbally. The survey form was handed to the entire class and time was provided for those who wished to participate in the study. The survey forms were collected as each participant completed the survey.

Flight instructors were approached by the researcher at a staff meeting. A period of three days was made available for the instructors to complete and return the survey form. Those instructors not present at the meeting were individually approached by the researcher at a later date and invited to participate in the study. The time taken to complete the survey form was approximately 15 – 20 minutes.

Ethical approval for the study was given by the Massey University Human Ethics Committee, approval 02/092.

3.4 Measures

3.4.1 Flight Experience

Data was gathered regarding the amount of flight experience held by the students by asking them to indicate the semester of study they were currently undertaking. The university programme for pilot training has four distinct semesters of study. Students must pass through each semester for which there is an allocation of flying sorties that must be achieved. A sortie can be a flight, either dual or solo, or a lesson in a ground based part task trainer. There are milestones that must be achieved in each semester in terms of licences or ratings to be passed.

The flight instructors were asked to indicate the level of the flight instructor rating they held.

3.4.2 Technical Attributes

Students and instructors were asked to rate the importance of flying skills and instructional abilities in an effective flight instructor on seven items using a 5 point Likert Scale (from 1 for not important and up to 5 for very important). For flight instructors technical competence is equivalent to subject expertise. Telfer (1993a) states that effective flight instructors are required to have subject knowledge such that their expertise cannot be doubted. The literature surrounding SET indicates that subject expertise is a key quality in student evaluation (Abrami, *et al*,, 1997). In consultation with subject matter experts questions regarding subject expertise were expressed as accurate flying skills, experiential hours flying and instructing– as in aviation the more hours a pilot possesses the more expert s/he is deemed to be (Hunter, 1999) – as well as the range of flight ratings held by the flight instructor. The final question asked about the importance of instructing being a career option for the flight instructor as most instructors used instructing as a stepping stone towards an airline career (Arnold, 1991; Henley, 1995).

3.4.3 Personal Attributes

Data was gathered regarding the personal attributes of an effective flight instructor. Students and instructors were asked to rate the importance of personal attributes in an effective flight instructor on seven items using a 5 point Likert Scale (from 1 for not important to 5 for very important). The personal characteristics of the

instructor are part of the Student Instructional Report, one of the two most widely used evaluation forms used in North America (Hobson & Talbot, 2001). Items included questions regarding, approachability and friendliness of the instructor, the willingness to help, sense of humour, enthusiasm about flying and willingness to learn more about flying, satisfaction with the role of instructing and punctuality.

3.4.4 Interpersonal Skills

Students and instructors were asked to rate the importance of interpersonal skills in an effective flight instructor on seven items using a 5 point Likert Scale (from 1 for not important to 5 for very important). These skills have been identified by Murray (1997) as being a strong predictor of instructional outcomes. The teacher – student relationship has been a consistently highly rated attribute in SET involving medical studies instructors, (e.g. McLean, 2001b; Risenberg *et al.*, 2001). Items included questions on communication skills and the rapport established with the student including interest in and sensitivity to the student and having realistic expectations of the student.

3.4.5 Teaching Ability

Students and instructors were asked to rate the importance of teaching ability in an effective flight instructor on seven items using a 5 point Likert Scale (1 for not important to 5 for very important). Items included the use of objectives, feedback and evaluation provided to the student including responding to students' questions, clarity of instruction and ability to adapt instructing to the learning style of the student (Moore, *et al.*, 2001).

3.4.6 Meeting Organisational Goals

Students and instructors were asked to rate the importance of how effective the flight instructor was at meeting organisational goals on 5 items using a 5 point Likert Scale (from 1 for not important to 5 for very important). Students, in part, enroll with a training organisation because of the stated aims of the organisation (e.g. training to professional or recreational level) and the instructor is a representative of that organisation (Arnold, 1991). Items included questions regarding the management of the student's progress, integration of ground and flying learning, instructing to airline standards (as per the stated aim of the organisation in which the participants were enrolled and employed) and whether the flight instructor conducted flying activities as expected by the organisation and as laid out in written procedures manuals. The training organisation has been found to have an effect on student learning (Moore *et al.,* 2001).

An open ended question was also included to allow all participants to provide additional comments about the attributes of an effective flight instructor.

Chapter 4

RESULTS

The SPSS (11.5 for Windows) package was used to examine the data gathered from the surveys of the participants.

The population for the study was the flight instructional staff and students in the Bachelor of Aviation (Airline Transport Pilot) programme at Massey University School of Aviation's Albany and Palmerston North campuses. The Bachelor of Aviation (Airline Transport Pilot) programme is unique within New Zealand because of an "equivalence" agreement with New Zealand Civil Aviation Authority. This means that the examinations and flight testing procedures conducted by Massey University School of Aviation are considered to be "equivalent" to the examinations and testing conducted by the New Zealand Civil Aviation Authority and its designated agent. The Airline Transport Programme is for students who want to pursue a professional pilots' career. The students will look to leave the course of study with a New Zealand Commercial Pilots Licence (CPL) and Multi-Engine Instrument Rating (MEIR) so as to have the minimum regulatory requirements for airline employment.

Within each semester of study the students' study six papers, with some having flying practicum attached to the paper. Practicum can include both flight in a fixed wing aircraft (single or multi-engined aircraft) or exercises in a part task trainer or flight simulator. The number of flight hours achieved by each student per semester is 50 – 55.

4.1 Sample Description

Students were surveyed from all four semesters of the flight programme. The students enrolled in semester 1 of the programme fly 50 – 55 hours during the semester of study. Of the total sample of students surveyed, 51% were enrolled in their first semester of study and the response rate from these students was 97.5%, (N = 40). Second semester students fly another 50 – 55 hours and sit a flight examination to gain their Private Pilots Licence. Of the total sample of students surveyed, 14% were enrolled in their second semester of study. The response rate from these students was 100%, (N = 11). Third semester students fly another 50 – 55 hours and sit a flight examination to gain their Commercial Pilots Licence. Of the total sample of students surveyed, 19% were enrolled in their third semester of study. The response rate from these students was 68%, (N = 15). Fourth semester students fly another 50 – 55 hours, gain a multi-engine type rating and sit a flight examination for an instrument rating. Of the total sample of students surveyed, 15% were enrolled in their third semester of study. The response rate for these students was 80%, (N = 12).

The flight instructors delivering flight practicum (instruction) within the Bachelor of Aviation Airline Transport Programme adhere to the regulatory requirements as established by New Zealand Civil Aviation Authority Rules Part 61 Subpart G. For category A flight instructors a minimum of 1250 flight hours are required, including 1,000 hours Pilot-In-Command and 750 hours flight instructing. Of the total sample of instructors surveyed the number of Category A instructors who responded was 2, a response rate of 66.6%. For Category B flight instructors a total of 500 flight hours are required, this includes 450 hours Pilot-In-Command and 250 hours flight instructing. Of the total sample of instructors surveyed the number of Category B instructors who

responded was 10. For Category C flight instructors a total of 200 flight hours are required, this includes 150 hours Pilot-In-Command. As the most junior instructor the Category C instructor has an additional restriction of having to instruct under the supervision of a Category B or A flight instructor. Of the total sample of instructors surveyed the number of Category C instructors who responded was 10. The response rate for category B and C flight instructors was difficult to ascertain due to the fluctuating numbers of staff employed over the duration of data collection. The frequency of respondents in each grouping is provided in Figure 1.

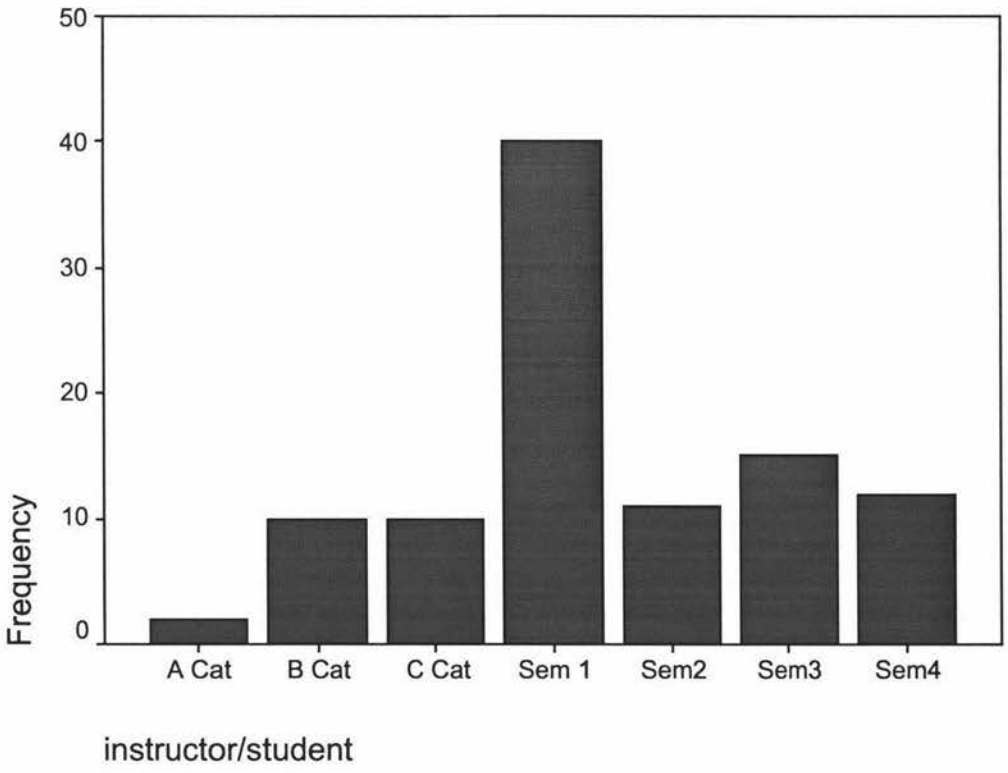

Figure 1. Frequency of Flight Instructor and Student respondents

No data was collected regarding the gender of participants due to the low number of women involved in the programme, either as flight instructors or students (less than 10%). This is similar to the research of Turney *et al.,* (2002).

4.2 Results

All data met assumptions of normality (Tabachnick & Fiedell, 2001). Cronbach alpha coefficients for each of the five measures are presented in Table 1.

Table 1

Reliability Analysis for the five effective instructor scales (Alpha).

Scales	**Alpha**
Technical attributes	.75
Personal attributes	.66
Interpersonal Skills	.73
Teaching Ability	.77
Training Organisation	.67

Ideally the Cronbach alpha should be at least .7 (Nunnally, 1978). The internal consistency of scores for the measures of personal attributes (α .66) and the importance of the training organisation (α .67) for effective flight instruction do not quite meet this criterion and findings relating to these scales need to be interpreted with caution.

4.3 Attributes of an Effective Flight Instructor

The first research question was to identify the attributes of flight instructors that students believe are both effective and ineffective in helping them learn to fly

The five top ranking attributes of an effective flight instructor as described by students are presented in Table 2. The students describe an effective flight instructor as someone who has *good flying skills* and is *able to accurately demonstrate the flight manoeuvre* with *clear instruction* by *communicating in a positive manner* and *responds to the questions the students have.*

Table 2

Means and standard deviations for attributes rated as most important for effective flight instruction – all students (N = 78)

Variable	**Mean**	**SD**
Has good flying skills	4.79	.47
Responds to students questions	4.69	.52
Communicates in a positive manner	4.67	.62
Has clarity of instruction	4.65	.55
Is able to accurately demonstrate flight manoeuvres	4.62	.61

The five bottom ranking attributes of an effective flight instructor as described by students are presented in Table 3. The attributes ranked as least important for effective flight instruction by students concern the measuring of experience by flight hours for

both total flight time (over 1,000 flight hours) and instructing time (more than 500 hours) and the breadth of flight experience. Being a career instructor and having a sense of humour are also not rated as important for effective flight instructors by students.

Table 3

Mean and standard deviations for attributes rated as least important for effective flight instruction – all students (N = 78)

Variable	**Mean**	**SD**
Has over 1,000 hours of flying experience	2.40	1.17
Instructing is a career	2.75	1.29
Has over 500 hours instructing experience	2.97	1.24
Has a wide range of flying experiences	3.03	1.17
Has a good sense of humour	3.62	1.03

The second research question was to identify the attributes of flight instructors that flight instructors themselves believe are both effective and ineffective in helping students learn to fly. The five top ranking attributes of an effective flight instructor as described by instructors are presented in Table 4. The instructors describe an effective flight instructor as someone who is *willing to learn more about flying* and is *able to describe while demonstrating the flight manoeuvre* with *clear instruction* and *remaining calm* during flight. The instructors also see *punctuality* as an attribute that is required for an effective flight instructor.

Table 4

Mean and standard deviations for attributes rated as most important for effective flight instruction – all instructors (N = 28)

Variable	**Mean**	**SD**
Is willing to learn more about flying	4.41	.59
Is punctual for flights	4.41	.67
Has clarity of instruction	4.36	.66
Is able to describe the flight manoeuvre whilst demonstrating the manoeuvre	4.32	.72
Remains calm during flight	4.27	.70

The five bottom ranking attributes of an effective flight instructor as described by instructors are presented in Table 5. The attributes that are described as least important in flight instruction by instructors concern the measuring of experience by flight hours for both total flight time (over 1,000 flight hours) and instructing time (more than 500 hours) and the breadth of flight experience. Being a career instructor and teaching to airline standard are not seen as important for effective flight instructors by instructors.

The third research question was to compare the importance ratings of attributes of an effective flight instructor as identified by the students and the flight instructors. Independent t-tests were conducted to evaluate differences in importance ratings on the five effective instruction variables between students and instructors. Results are presented in Table 6. There were significant differences between students and instructors on the

Table 5

Means and standard deviations for attributes rated as least important for flight instruction – all instructors (N = 22)

Variable	**Mean**	**SD**
Has over 1,000 hours of flying experience	2.23	1.11
Has over 500 hours instructing experience	2.64	1.22
Instructing is a career	2.75	1.09
Teaches to a standard that meets airline standards	3.14	1.17
Has a wide range of flying experiences	3.14	1.21

Importance of the Training Organisational measure ($t(98) = 2.89$, $p<.01$) and the Interpersonal Skills measure ($t(98) = 2.72$, $p<.01$). The students rated both the Importance of the Training Organisation and Interpersonal Skills measure more highly than did the instructors. There were no significant differences between students and instructors on the remaining measures of Teaching Ability ($t(98) = 1.25$, $p = .214$), Technical Attributes ($t(98) = 1.21$, $p = .231$) and Personal Attributes ($t(98) = 1.97$, $p = .051$).

The fourth research question was to identify any differences on importance ratings between the differing experience levels of both students and instructors. A one-way between-groups analysis of variance was conducted to discover whether importance

Table 6

Means and standard deviations on attributes of effective flight instructors for instructors and students

Attributes	Students (N = 78)		Instructors (N = 22)		t
	Mean	SD	Mean	SD	
Technical Attributes	69.78	4.36	68.48	4.64	ns
Personal Attributes	142.62	1.03	142.14	0.79	ns
Interpersonal Skills	30.60	3.19	28.43	3.48	**
Teaching Ability	30.62	3.22	29.62	3.41	ns
Training Organisation	20.67	2.76	18.71	2.70	**

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

ratings differed due to flying experiences and findings are presented in Table 7. Due to the small number of A Category Flight Instructors (N = 2) the responses of the A Category Flight Instructors were combined with the B Category Flight Instructors (n = 12 for combined A and B Category Flight Instructors). In actual flight instructing there is a difference between C Category Flight Instructors and B and A Category Flight

Instructors. This is most notably evident in the direct and indirect supervision of C Category Flight Instructors. For the first 100 hours of delivering flight instruction the C Category Flight Instructor must be under the direct supervision of a B or A Category Flight Instructor. There after while remaining a C Category Flight Instructor the Instructor must be under indirect supervision by a B or A Category Flight Instructor.

There were no significant differences across the different experience groups for Technical Attributes ($F(5,94) = .969$, $p = .44$) or for Personal Attributes ($F(5,94) = 1.36$, $p = .24$).

There were significant differences across experience levels for Interpersonal Skills ($F(5,94) = 3.40$, $p < .01$), Teaching Ability ($F(5,94) = 4.47$, $p < .01$) and the Importance of the Training Organisation ($F(5,94) = 4.07$, $p < .01$).

Post-hoc comparisons using Scheffes Ranges tests show that C Category Flight Instructors rated the importance of Interpersonal Skills significantly lower than Semester 4 students ($p<.05$). Regarding Teaching Ability, senior instructors (A & B Category Flight Instructors) rated the importance of this factor significantly higher than C Category Flight Instructors ($p<.05$). C Category Flight Instructors also rated this factor significantly less important than Semester 1 ($p< .05$), Semester 3 ($p< .05$) and Semester 4 ($p< .05$) students. There was no significant difference between C Category Flight Instructors and students engaged in their second semester of study. For the Importance of the Training Organisation C Category Flight Instructors again rated this variable as less important than Semester 1 ($p< .05$), Semester 3 ($p< .05$) and semester 4 ($p< .05$) students.

Table 7

Means and standard deviations on attributes of effective flight instructors across flight experience levels

Variables	Senior Instructors		C Cat Instructors		Semester 1 Students		Semester 2 Students		Semester 3 Students		Semester 4 Students		
	n = 12		n = 10		n = 40		n = 11		n = 15		n = 12		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	F
Technical Ability	69.58	5.11	66.9	3.51	69.38	4.3	70.63	4.34	70.13	4.6	70.25	4.69	ns
Personal Attributes	142.25	0.87	142.1	0.74	142.65	1.05	142.27	1.27	142.47	0.99	143.0	0.74	ns
Interpersonal Skills	29.67	3.82	26.86	2.11	30.38	3.42	30.18	2.64	30.33	3.04	32.33	2.87	**
Teaching Ability	31.42	2.81	26.5	3.37	30.78	2.72	29.45	2.66	30.93	3.92	31.67	2.9	**
Training Organisation	19.67	2.84	17.1	2.18	20.84	2.64	19.55	2.88	21.13	2.17	21.08	3.15	**

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

Chapter Five

DISCUSSION

This chapter will discuss the findings of the research in relation to the aims of the research and the reviewed literature. Limitations of the study and possible implications for flight training are also discussed. The aim of the present research was to identify the attributes of an effective flight instructor by using the evaluations of students and flight instructors within a university aviation programme.

5.1 Attributes of an Effective Flight Instructor - students

The first research question concerned the identification of attributes of flight instructors that students believed were both effective and ineffective in helping students learn to fly.

The students described an effective flight instructor as someone who has *good flying skills*, is *able to accurately demonstrate the flight manoeuvre* with *clear instruction* by *communicating in a positive manner* and who *responds to the questions the students have.*

For the flight instructor good flying skills can be analogous to subject expertise of lecturers in other disciplines. Abrami, *et al., 's* (1997) factor analysis of SET found that *knowledge of domain* was one of the least important factors of instructional effectiveness. Whilst in a safety critical industry such as aviation it should not be a surprise that the students want an instructor who has subject expertise, studies from other safety critical

domains such as medicine do not necessarily support this. McLean's (2001b) and Das and El-Sabban's (1996) studies of medical students indicated that the teacher student relationship and attributes that enhanced this were rated more highly than subject expertise. However these studies were of classroom based medical students. A closer parallel to flight instruction are the studies by Irby (1978) and Irby and Rakestraw (1981) of medical students' ratings of instructors in a clinical setting. The students from both these studies rated highly for effective teaching the attributes of instructor knowledge and clinical skills. The study of Riesenberg *et al.,* (2001) also found clinical skills were highly rated by medical students. Irby, Gilmore and Ramsey (1987) noted in their study a difference in the ratings for the attributes of knowledge – highly rated – and demonstration of clinical skills – lowly rated – in effective teaching. This is in contrast to the aviation students in the present study who ranked the attribute of demonstrating skills very highly. Telfer (1993a) describes the effective flight instructor as requiring to know their subject very well.

However the measuring of this required flying skill by *flight hours* for both total *flight time* (over 1,000 flight hours) and *instructing time* (more than 500 hours) and the *breadth of flight experience* was ranked as least important for effective flight instruction by students. The equating of expertise with experience (Hunter, 1999) has been a traditional method of measurement within aviation. Woodhouse (2001) suggests that in an effort to lift the standard of flight instruction the criteria for pilots wanting to become flight instructors be 1,000 flight hours. This is not supported in the findings of this study. Concern has also been raised about the lack of career opportunities for newly qualified professional pilots, (Woodhouse, 2001) and the ensuing use of flight instructing as a

stepping stone to other employment within aviation having a negative influence on the standard of flight instruction offered to students (Henley, 1995). Again, these concerns have not been supported in this study. Students ranked the need for an instructor to be a career instructor as the second lowest desired attribute. Having a sense of humour was rated by the students as the attribute that was least important for effective flight instruction.

The other highly ranked attributes of an effective flight instructor as described by the students were those pertaining to teaching abilities; of having clarity of instruction, being able to respond to the student's questions as well as the interpersonal skill of communicating positively. The high ratings of these attributes regarding teaching abilities match those by students from other disciplines (Abrami *et al.,* 1997). Clarity of instruction is one of the main dimensions of effective teaching, (Irby, 1978, Jackson *et al.*, 1999; Hativa *et al.,* 2001). With regards to the importance of teaching skills for effective teaching, studies of medical students showed that the students rated this attribute highly, (Das, El-Sabban, 1996; Vu *et al.,* 1997).

5.2 Attributes of an Effective Flight Instructor – flight instructors

The second research question concerned the identification of attributes of flight instructors that flight instructors believe are both effective and ineffective in helping students learn to fly. The instructors describe an effective flight instructor as someone who is *willing to learn more about flying* and is *able to describe while demonstrating the flight manoeuvre* with *clear instruction* and *remaining calm* during flight. The instructors also see *punctuality* as an attribute that is important for an effective flight instructor.

As with the students, the flight instructors describe the importance of subject expertise and teaching abilities as being the most important attributes of an effective flight instructor. Henley (1995) notes that one of the sources of stress facing a newly qualified flight instructor is their lack of confidence in their flying abilities. In continuing to be expert in their subject area, flight instructors ranked as the highest attribute of a flight instructor the need to be willing to learn more about flying.

The most common teaching method within flight instruction is the demonstration – performance method (Civil Aviation Authority of New Zealand, 1999). This method involves the instructor guiding the student through the required skills to perform a flight manouevre by talking (the instructors patter) and flying the manouevre. Part of the requirements to gain a C Category flight instructors rating is to complete a 25 hour training course in the "techniques of basic flying instruction" (New Zealand Civil Aviation Rules, Advisory Circular Part 61). Combining these two cognitive tasks of talking and motor skill can be daunting for the new flight instructor. Therefore much emphasis is placed on being able to perform these two tasks during training in preparation for actual flight instruction duties. It is not a surprise that the flight instructors in this study rated this attribute highly.

The students rated a variation of this attribute highly where the students wanted a flight instructor to be able accurately demonstrate only rather than producing the patter at the same time. The flight instructors focused on their own expected requirements for instruction. This is what Telfer (1993a) describes as instructor–centred rather than trainee–centred. The instructor-centred approach to learning means the instructor's task of teaching becomes more important than the students' learning outcomes. This is further

supported from this present study by the flight instructors' high ranking of the need to be willing to learn more about flying as an attribute of effective flight instructors. To improve the effectiveness of flight instruction the flight instructors see the need to improve their subject knowledge. This is focused on the instructor rather than the student for effective teaching. However Abrami, *et al.*, (1997) note that effective teaching can be defined from the different perspectives of students' product – the changes that happen within students because of teaching - or instructors' processes – the acts of teaching. Both perspectives contribute to effective teaching.

The teaching attribute or skill of clear instruction was also rated highly by the surveyed flight instructors. Henley's (1991; 1995) research amongst Canadian and Australian flight instructors indicates that the flight instructors were aware of their role as teachers but felt they were not being properly prepared for the role of teaching that is inherent in flight instruction. The instructors in these studies recognised the need to be trained in teaching skills but were often not given any training in how to acquire these skills, "basic teaching skills such as lesson planning, learning skills and interpersonal and communication skills were left to be acquired through experience or by trial-and-error" (Henley, 1991, p. 321). Teaching skills was the second most highly rated characteristic of a successful flight instructor in Arnold's (1990) study of personality characteristic of successful Californian flight instructors behind patience. The flight instructors in the present study recognise the need for teaching skills to be an important part of the instructing process. A third semester student emphasised this point when they wrote:

"I think teaching skills are extremely important and don't

seem to be emphasised much on a tertiary level." (Respondent 79)

Like the students participants, the flight instructors did not see the measuring of expertise by the number of flight hours – either total or instructional hours – obtained by a flight instructor as being important to be an effective flight instructor. The generally accepted notion within the aviation industry of more hours equals greater expertise (Hunter, 1999) has not appeared to have been supported by the students and flight instructors surveyed in this study. Further, both groups of respondents did not see the need for a flight instructor to be committed to making a career out of flight instruction. One Category B flight instructor sums up the attitude required of hour building instructors with the response:

> *"It is essential to take a pride in the work – even if only hour building for an airline career – otherwise you should not be instructing."* (Respondent 5)

It may be that the students' and flight instructors' low ranking of instructing as a career attribute indicates that if a perceived good job is being performed by the flight instructor, the final career of the flight instructor is not an issue for effective flight instruction.

Despite the stated aim of the aviation programme from which participants were drawn was to prepare students for entry into the airline environment, the attribute to train to airline standard was ranked lowly by the flight instructors. This may be an indication

that the flight instructors see the programme as an end in itself rather than the beginnings for the students.

5.3 Comparison between students and flight instructors

The third research question was to examine any differences in the importance ratings of the attributes of an effective flight instructor as identified by the students and the flight instructors. The students rated both Interpersonal Skills and Importance of the Training Organisation more highly than did the flight instructors. There were no differences between the two groups on the other measures. The similarities of importance ratings on these three measures between the two groups indicates that they may have a common understanding of the attributes required for flight instruction. The similarities between students and instructors in identifying common attributes that contribute to effective teaching has been noted by Marsh (1987).

However, other research has tended to report more differences than similarities in the evaluations of staff and students. For instance Hobson and Talbot (2001) note that correlations between the ratings of students and teachers range form .20 to .65. Hativa *et al.,* (2001) noted strong evidence for differences between the perceptions of effective teaching of students and staff. Two studies from Australian universities (Killen, 1994; Ballantyne, *et al.*, 2000) found significant differences between the students' evaluations and the reports from the lecturing staff. The differences between the two groups in the present study is also consistent with the findings of Moore *et al.,* (2001) who found there was a lack of alignment in instructional issues between flight instructors and students.

The students higher rating of the Interpersonal Skills attribute compared to the flight instructors is a common theme in the literature covering a broad spectrum of academic subjects. The importance of the relationship between student and instructor is

consistently rated by students as an attribute of effective teaching (d'Ápollonia & Abrami, 1997). Across academic disciplines studies note the importance of the establishment of rapport between teachers and students in effective teaching (Hativa, *et al.,* 2001). Student rapport is a first order factor identified by Jackson, *et al.,* (1999) in a factorial study of nearly 8,000 university classes at a North American University across a broad spectrum of academic disciplines. Within McLean's (2001) study of third year medical students attributes of the teacher-learner relationship such as friendly and approachable, and being sensitive to the needs of the students, were highly valued by medical students in their medical instructors. Clinical teachers of medical students who were more involved with the students were considered to be more effective teachers (Irby, *et al.,* 1987). Within aviation Henley's (2003) review of research into flight instructors indicates that where flight instructors do not establish a positive teacher-learner relationship with the student, stress is created that is not conducive to the student learning to fly. Besco (1992) also notes flight instructor attitudes towards students as being an important factor in the quality of training received by the student.

The present study confirms the establishment of positive rapport between student and instructor as being perceived by the students as important for effective flight instruction. The following responses from semester one students illustrate the importance which students' place on the importance of rapport between the flight instructor and themselves:

> *"I find the best instructors are the ones that become more of a friend than an instructor. You can learn well, but still be able to have a*

joke and be flying in a relaxed situation" (Respondent 41)

"They must have patience and a good interpersonal skills (sic)
to build a stable relationship between instructor and student" (Respondent 43)

Establishing rapport need not be a difficult task to achieve as one second semester student recorded:

"Say hello in the mornings or just say something, small things like that makes the entire day great." (Respondent 70)

The importance of communication within the development of rapport not only involves the instructor talking to the student but also as a fourth semester student plainly states:

They should listen to students. (Respondent 97).

The students rating more highly the importance of the training organisation is similar to Moore *et al.,* (2001) noting the importance of the training organisation has in determining the outcomes of learning. These researchers describe how organisational factors such as syllabus, resources, supervision will affect the interaction between the student and the flight instructor.

As to why there are differences between the students and flight instructors, Ballantyne *et al.,* (2000) and Fisher *et al.,* (1998) identify the differences between instructors and students as perhaps arising from different understandings of what

effective teaching is. Killen (1994) and Ballantyne *et al.,* (2000) also note the tendency amongst both groups involved in the teaching and learning process to identify factors for the improvement of teaching and courses that are under the control of the other group.

5.4 Differences between experience levels

The fourth research question was to identify any differences on importance ratings of flight instructor attributes between the differing experience levels of both students and instructors. Turney *et al.,* (2002) noted that attitudes amongst female flight students in American colleges and universities changed as they progressed through their flight training programme whilst for male students there was little change. These findings for the male students seem to be supported in the present study where the overwhelming majority of students were male and there were no differences on ratings of the importance of flight instructor attributes between the semester 1 students and more experienced students in semesters 2, 3, or 4.

Between the different levels of flight instructor experience C Category Flight Instructors rated the importance of Teaching Ability as an attribute of an effective flight instructor lower than did the senior Flight Instructors (A & B Category Flight Instructors). Between the differing levels of experience amongst students and flight instructors, the C Category Instructors rated the importance of Interpersonal Skills lower than Semester 4 students. The C Category Flight Instructors rated the importance of both Teaching Ability and the Importance of the Training Organisation as lower than Semester 1, 3 and 4 students. It has been noted that the educational socialisation of a teacher will influence the way they teach (Oosterheert, Vermunt & Denessen, 2002). This socialisation can be transferred from the mentor teacher. Woodhouse (2001) notes the depletion of experienced flight instructors to other positions with the aviation industry

leaves less experienced flight instructors to train the next generation of flight instructors. The newly qualified flight instructor may have learnt the attributes required for flight instruction from their observations of the flight instruction process whilst students. Due to the lack of experienced flight instructors to they newly qualified Category C flight instructors fall back on their learnt folkways (Telfer, 1993b).

5.4 Limitations

The limitations of this study are centred on the unique nature and size of the population sampled and the measures used.

The respondents to the survey were *ab initio* fixed wing students and their flight instructors operating within a flight training programme offered by a university. This may make it difficult to generalise results across to flight training programmes offered by other providers of flight training.

The university flight training programme undertaken by the participants of the study has unique features that are not replicated amongst other providers of flight training within New Zealand. It is the only tertiary education institution to offer organic flight training through an integrated flight practicum and academic studies programme. The end result for students successfully completing the course is not only New Zealand Civil Aviation issued pilot licences and ratings but also a bachelor degree in aviation (BAv ATP) conferred by the training organisation. Other flight training organisations within New Zealand offer training to gain New Zealand Civil Aviation issued pilot licences and ratings and sometimes the opportunity to earn tertiary level qualifications to the diploma level as part of the flight training. However the tertiary qualification that can be earned is conferred by education providers other than the training organisation. That is, the flight

training organisation is a contractor to another institution for the provision of flight training.

Even more uniquely, the programme from which the participants were drawn has an agreement with the Civil Aviation Authority of New Zealand that is not known to be replicated anywhere else in the world. The examinations and flight tests conducted by the university flight training programme are considered by the Civil Aviation Authority of New Zealand to be "equivalent" to Civil Aviation Authority of New Zealand conducted examinations and flight tests, (in New Zealand the task of conducting examinations and flight tests has devolved to an industry owned company, Aviation Services Limited). This may make it difficult to generalise results across to flight training programmes that follow more closely the requirements of New Zealand Civil Aviation Rules Part 61. Whilst the programme followed by the university does cover Part 61, it goes further and is broader in scope.

It may also be difficult to generalise results across to other sectors of the flight training sector such as rotary wing and agriculture that were not sampled. The desired attributes of flight instructors as identified by students within these sectors may be different to those identified by the fixed wing *ab initio* students of this study.

The small sample size may mean a bias towards the less experienced students (semester 1 students totalled more than the other three semester students combined) and flight instructors. The number of experienced Category A flight instructors sampled was very small, Category A respondents being only 10% of the Categories B and C flight instructors. The relatively small sample size from participants within the university aviation programme compared to the numbers of students in other university disciplines

may limit the comparisons with findings from student evaluations of teaching from the other disciplines.

Most SET studies use the evaluations of specific teachers and teaching situations. The present study used a more theoretical situation where the students were asked not to evaluate their existing flight instruction. The study measured the perceptions of the students and flight instructions of effective flight instruction. The use of SET has been used in this manner to measure the perceptions of medical students regarding attributes of effective clinical teachers (Irby, 1978). This however does not necessarily mean that the results will lead to improved flight instruction. SET has also been used prior to a course of academic study commencing to get a prospective view of what students expected in their instructors, (Sander *et al.,* 2000). In using SET as an evaluation tool Fisher *et al* (1998, p162) note that students may approach a survey instrument such as was used in this study with the view point of "What do I want from my flight instructor?"

Within the use of SET there is much debate about the shape of the ratings form and the items to be included, whether it be with multidimensional ratings (Marsh & Roche, 1987) or more global factors (d'Apollonia & Abrami, 1997). Abrami, *et al.,* (1997) observe that different survey instruments measure different dimensions of effective teaching. The item content and how many items are selected to assess chosen dimensions of effective teaching are problems associated with the construction of instruments to measure students' evaluations (Jackson, *et al.,* 1999).

On the construction of survey instruments, Marsh (1987, p. 263) comments:

> If a survey contains separate groups of related items that
> are derived from a logical analysis of the content of effective

> teaching and the purposes that the ratings are to serve .... then it is possible to interpret what is being measured.

The survey material used in this study was drawn up by reference to the literature and in consultation with senior flight instructors in the general aviation and airline sectors of the aviation industry. Although Marsh (1987) found that factors underlying effective teaching were similar across different disciplines Jackson *et al.,* (1999) notes that different items should be chosen for the survey instrument if it to be used for a specific domain of study as was the case in the present study. There were no specific studies examining effective flight instruction within the literature that could be used to assist the formulation of the survey instrument. This may mean that within the survey instrument, items were asked of the students and the flight instructors that may be unrelated to effective teaching in the flight instruction domain. Internal reliability scores of .66 to .67 for two of the scales suggest this may be the case.

Although there are limitations to this study, the use of students is a widely used method in tertiary education to evaluate the effectiveness of the teaching the students receive. The results from these evaluations are routinely used for formative and summative evaluations of teachers.

5.5 Future Research

Although Wilkinson (1999) states that making the statement "further research needs to be done" is a waste of space, in this instance this is what needs to be done. To increase the effectiveness of flight instructors is seen by Civil Aviation Authority of New Zealand (2001) as one way of increasing aviation safety. To achieve this a definition of effective flight instruction is required. As discussed earlier however, there is difficulty in defining effective teaching (Watchel, 1998). Like other domains of study, students have a

role to play in contributing to the rising of the standards and effectiveness of flight instructors and have already contributed to earlier research (e.g. Arnold, 1991; Henley, 2003). Possible areas of further research into SET in flight instruction could be in the sampling and size of students populations with a wider and larger group of respondents from a more representative group of *ab initio* students being sampled.

Other modes of evaluation could also be used other than a questionnaire. Johnson (2000, p.420) identifies a limitation in using a questionnaire format for SET where students and instructing staff can not "discuss, evidence, explain, justify, negotiate, or gain new insights into their own or the others' views, interests, values and assumptions during the process of the evaluative activity". Reliance solely on questionnaires for evaluation of effective teaching may be too narrow. Other forms of evaluation may be useful in helping students (and flight instructors) determine those attributes that contribute to effective flight instruction. Kolitch and Dean (1999) note that most of the forms used for the evaluation of teaching effectiveness are based on one particular paradigm of student learning, the transmission model of student learning where the instructor gives and the student takes in. While in flight instruction there may be a lot of rote learning of flight maneouvres, for the student there is also the need to develop other skills and attitudes which need to be taught but for which the transmission model of teaching is not as appropriate, such as decision making (Jenson, 1995). Whatever form of evaluation is used there should be agreement and input from the flight instructors to help get "buy-in" to the results of the evaluation (Snell, *et al.,* 2000).

Research also needs to be done on whether those attributes perceived by students to be important components of effective teaching do actually lead to more effective

instruction. Bosshardt and Watts (2001, p. 14) in writing about the teaching of economics students comment, "things perceived to be true are true in their consequences". If the students of flight instruction perceive that they are the recipients of effective flight instruction when the flight instructor displays attributes of good pilot and good teacher then the instruction offered goes a long way towards being effective. Hobson and Talbot (2001) also note that student evaluations report the perceptions of students, perceptions that may not be accurate, but never the less present the students' view of the teaching process.

Brightman, Elliot and Bhada (1993) comment there is little use in developing valid and reliable survey instruments if there is no mechanisms within the training organistion to improve and reward enhanced teaching. Hativa, *et al.*, (2001) in reviewing the literature surrounding effective teaching note that teachers' thinking and knowledge about teaching differed from their actual practice. Whilst it is necessary for teachers to gain knowledge about teaching, better teaching practice is not necessarily achieved. Using student evaluations to attempt to increase the effectiveness of flight instruction will only be useful if flight instructors, and their training organisations are prepared to follow through on the results, (Ballantyne, *et al*, 2000; Kember, *et al* , 2002). So that alignment between the perceived attributes of effective teaching held by students and flight instructors occurs future research needs to examine what Lehrer *et al.,* (2000, p. 155) describes as "professional socialistaion", the taking on of "the values and attitudes of a particular group to which one belongs".

5.6 Conclusion

The students and flight instructors participating in this study identified attributes they perceived to be important for effective flight instruction. The number of instructional and flight hours accumulated by a flight instructor – the traditional method of measuring expertise (Hunter, 1999) - was not rated as a desirable attribute of an effective flight instructor by students or flight instructors. The identification of piloting and teaching skills as attributes of effective flight instructors matches the call made by Hunt (2000, cited in Thomas, 2001) for the development of the flight instructor as an "aviation-teacher" so as to improve aviation training and safety. The findings of this study would also seem to be similar to the regulatory requirements for training to become a flight instructor as listed in Civil Aviation Rules Part 61. Candidates to become a Category C flight instructor are required to undergo an instructional techniques course of at least 12 hours and undergo a flight training programme of at least 25 hours in the practice and theory of flight instruction. It would appear therefore with these similarities between researchers, students, flight instructors and the regulators the "what" is required of an effective flight instructor is known.

One area of improvement to the training and development of flight instructors that arises from the findings of this study is the importance students placed on rapport as an attribute of an effective flight instructor as compared to the flight instructors. The importance of the teacher –learner relationship (McLean, 2001a) is a common theme in the literature surrounding SET. This is not specifically mentioned in the syllabus of the instructional techniques course required to be undertaken by the trainee category C flight instructor.

Calls to improve the standards and effectiveness of flight instructors (e.g. Civil Aviation Authority of New Zealand, 2001) can perhaps be seen as an indication that the achieving of what is required (i.e. piloting and teaching skills) of a flight instructor is not being met. This is borne out in the research of Henley (1995) where new flight instructors stated they were not fully confident of their piloting skills, especially when having to demonstrate a manoeuvre to a student (an attribute that was highly ranked by the flight instructors in this study). The new instructors were also aware of shortcomings in their training to be a teacher of aviation. There would appear to be a need to ensure pilots training to become category C flight instructors are properly educated and prepared in the areas of piloting and teaching skills.

References

Abrami, P.C., d'Apollonia, S., & Rosenfield, S. (1997). The dimensionality of student ratings of instruction: what we know and what we do not. In R. P. Perry, & J. C. Smart (Eds.), *Effective teaching in higher education: research and practice* (pp. 321-367). New York: Agathon Press

Amin, M.E. (2002). Six factors of course and teaching evaluation in a bilingual university in Central Africa. *Assessment & Evaluation in Higher Education, 27(3),* 281 – 291.

Arnold, G.H. (1991). *Personality characteristics of successful general aviation flight instructors.* Unpublished doctoral dissertation, United States International University, San Diego, USA.

Ballantyne, R., Borthwick, J., & Packer, J. (2000). Beyond student evaluation of teaching: identifying and addressing academic staff development needs. *Assessment & Evaluation in Higher Education 25*(3), 221 – 237.

Bentley, G. (1969). *RNZAF: A short history.* Wellington: Reed.

Besco, R.O. (1992). Analyzing knowledge deficiencies in pilot performance. *International Journal of Aviation Psychology 2*(1) 53 – 74.

Bosshardt, W., & Watts, M. (2001).. Comparing student and instructor evaluations of teaching. *Journal of Economic Education 32*(1) 3 – 17.

Brightman, H.J., Elliott, M. L., & Bhada, Y. (1993). Increasing the effectiveness of student evaluation of instructor data through a factor score comparative report. *Decision Sciences 24*(1) 192 – 199.

Civil Aviation Authority of New Zealand. (1999). *Flight instructor's guide.* Wellington,

NZ: Author

Civil Aviation Authority of New Zealand. (2001) Towards 2005 – the aviation safety plan. Retrieved April 1, 2003, from http://www.caa.govt.nz

Civil Aviation Authority of New Zealand (2002). Civil Aviation Rules Part 61. Retrieved October 15, 2004, from http://www.caa.govt.nz

Cashin, W.E., & Downey, R.G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology. 84*(4), 563 – 572.

Centra, J. (1993). *Reflective faculty evaluation.* San Francisco: Jossey-Bass.

Chidester, T.R., Helmreich, R.L., Gregorich S.E., & Geis, C.E. (1991). Pilot personality and crew coordination: Implications for training and selection. *The International Journal of Aviation Psychology. 1*(1), 25 – 44.

Clayton, M. (1998, March 17). Give me an 'A' professor – I'll give you one, too. *Christian Science Monitor 90(76),* B6.

Clouston, A.E. (1954). *The dangerous skies. London:* Cassell & Company Limited.

Cobby, A.H. (1981). *High adventure.* Melbourne: Kookaburra Technical Publications Pty. Ltd.

Coffey, M., & Gibbs, G. (2001). The evaluation of the student evaluation of educational quality questionnaire (SEEQ) in UK higher education. *Assessment & Evaluation in Higher Education, 26*(1), 89 – 93.

d'Apollonia, S., & Abrami, P.C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*(11), 1198 – 1208.

de Montalk, R. J. (2000). *Anxiety as a factor in student pilot performance in a university*

aviation degree programme. Unpublished master's thesis, Massey University, Palmerston North, New Zealand.

Damos, D. (1996). Pilot selection batteries: shortcomings and perspectives. *International Journal of Aviation Psychology, 6*(2), 199 – 209.

Das, M., & El-Sabban, F. (1996). Student and faculty perceptions of the characteristics of an ideal teacher in a classroom setting. *Medical Teacher, 18*(2), 141 – 146.

Deming, W.E. (1986). *Out of the crisis.* Cambridge, MA: Massachusetts Institue of Technology Center for Advanced Engineering Study.

Elshaw, C. (1993). Preparing better flight instructors. In R. A. Telfer (Ed.), *Aviation instruction and training* (pp. 253 – 270). Aldershot, UK: Ashgate.

Fiorino, F. (2001, July 16). ALPA proposes single level of training. *Aviation Week & Space Technology, 155(3),* 64.

Fisher, A.T., Alder, J.G., & Avasalu, M.W. (1998). Lecturing performance appraisal criteria: staff and student differences. *Australian Journal of Education, 42*(2) 153 – 168

Fowler, R. (2000, May). Comair Aviation Academy. *Plane & Pilot,* pp72 – 73.

Frederick-Recascino, C., & Hall, S. (2003). Pilot motivation and performance: theoretical and empirical relationships. *International Journal of Aviation Psychology, 13*(4), 401 – 414.

Freeman, H.R. (1994). Student evaluations of college instructors: effects of type of course taught, instructor gender and gender role, and student gender. *Journal of Educational Psychology, 86*(4), 627 – 630.

Grandchamp, R.J. (1971). *Student pilots' attitudes towards fear concepts in flight*

training. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Greenwald, A.G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist, 52* (11), 1182 – 1186.

Gwynn-Jones, T. (1996). Dashing Claude Grahame-White was Britain's favourite pilot during aviation's wind-in-the-wire days. *Aviation History, 6*(6), 12 – 15.

Hager P., & Gonci, A. (1996). What is competence? *Medical Teacher, 18*(1), 15 - 19.

Hativa,N., Barak, R., & Simhi, E. (2001). Exemplary university teachers: knowledge and beliefs regarding effective teaching dimensions and strategies. *Journal of Higher Education, 72*(6), 699 - 730

Hawkins, F.H. (1993). *Human factors in flight.* Brookfield, VT: Ashgate.

Helmreich, R.L. (2000a). Culture and error in space: implications from analog environments. *Aviation, Space and Environmental Medicine, 71*(9), A133 - A139.

Helmreich, R.L. (2000b). On error management: lessons from aviation. *British Medical Journal, 320*, pp. 781 - 785

Henley, I. (1991). The development and evaluation of flight instructors: A descriptive survey. *International Journal of Aviation Psychology, 1*(4), 319 – 333.

Henley, I. (1995) *The quality of the development and evaluation of flight instructors in Canada and Australia*. Unpublished doctoral dissertation, University of Newcastle, NSW. Australia.

Henley, I. (2003). Factors that affect learning. In I.M.A. Henley (Ed.), *Aviation, education andtTraining,*(pp. 137 – 174). Aldershot: Ashgate

Hinton, H. (1993). Reliability and validity of student evaluations: testing models

versus survey research models. *PS: Political Science and Politics.* September, pp. 562 – 569.

Hobson, S.M. & Talbot, D.M. (2001). Understanding student evaluations. *College Teaching, 49*(1), 26 – 32.

Howard, S.K., Gaba, D.M., Fish, K.J., Yang, G., & Sarnquist, F. H. (1992). Anesthesia crisis management training: teaching anesthesiolgists to handle critical incidents. *Aviation, Space and Environmental Medicine, 63*(9), 763 – 770.

Howell, A.J., & Symbaluk, D.G. (2001). Published student ratings of instruction: revealing and reconciling the views of students and faculty. *Journal of Educational Psychology, 93*(4), 790 – 796.

Hunt, G,J,F. (2000). System of regulating aircrew licensing in need of surveillance-based approach. *ICAO Journal, 55*(8), 13 – 15.

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

Hunter, D.R. (2001). Retrospective and prospective validity of aircraft accidents risk indicators. *Human Factors, 43*(4), 509 – 519.

Irby, D. (1978). Clinical teacher effectiveness in medicine. *Journal of Medical Education, 53*, pp. 808 – 814.

Irby, D., Gillmore, G.M., & Ramsey, P.G. (1987). Factors affecting ratings of clinical teachers by medical students and residents. *Journal of Medical Education, 62*, pp. 1 – 7.

Irby, D., & Rakestraw, P. (1981). Evaluating clinical teaching in medicine. *Journal of*

Medical Education, 56, pp. 181 – 186.

Jackson, D.L., Teal, C.R., Raines, S.J., Nansel, T.R., Force, R.C., & Burdsal, C.A. (1999). The dimensions of students' perceptions of teaching effectiveness. *Educational & Psychological Measurement, 59*(4), 580 – 597.

Jameson Beaux, L.F. (2000). Attributes of effective economics instructors: an analysis of student evaluations. *The Journal of Economic Education, 31*(3), 211 – 228.

Jenson, R.S. (1995). *Pilot judgement and crew resource management.* Brookfield, VT. Ashgate.

Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education, 5*(4), 419 – 435.

Kember, D., Leung, D.Y.P., & Kwan, K.P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education, 27*(5), 411 – 425.

Kerridge, J.R., & Mathews, B.P. (1998). Student ratings of courses in HE: further challenges and opportunities. *Assessment & Evaluation in Higher Education, 23*(1), 71 – 82.

Killen, R. (1994). Differences between students' and lecturers' perceptions of factors influencing students' academic success at university. *Higher Education Research and Development, 13*(2), 199 – 211.

Kolitch, E., & Dean, A.V. (1999). Student ratings of instruction in the USA: hidden assumptions and missing conceptions about 'good" teaching. *Studies in Higher Education, 24*(1), 27 – 43.

Koonce, J.M. (1984). A brief history of aviation psychology. *Human Factors, 26*(5),

499 – 508

Koonce, J. M. (2002). *Human factors in the training of pilots.* London: Taylor and Francis.

Laurillard, D. (1997). Learning formal representations through multimedia. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning.* Edinburgh, UK: Scottish Academic Press.

Lehrer, H.R., Moore, P.J., Telfer, R.A., & Freeman, A. (2000). Future airline training: What has been learned from pilots and instructors? In A.R. Lowe & B.J. Hayward (Eds.), *Aviation resource management (pp. 149 – 157).* Aldershot: Ashgate.

Lewis, B. (2000). *Aircrew.* London: Cassell Military Paperbacks.

Lewis, C. (1936). *Sagittarius rising.* London: Warner.

Li, G., Baker, S.P., Grabowski, J.G., & RebokG.W. (2001). Factors associated with pilot error in aviation crashes. *Aviation, Space and Environmental Medicine, 72,* pp. 52 – 58.

Li-Ping Tang, T. (1997). Teaching evaluation at a public institution of higher education: factors related to the overall teaching effectiveness. *Public Personnel Management, 26*(3), 379 – 389.

Litzelman, D.R., Westmoreland, G.R., Skeff, K.M., & Stratos, G.A. (1999). Student and resident evaluations of faculty – how dependable are they? *Academic Medicine, 74*(10), S25 – S27.

McKeachie, W.J. (1997a). Good teaching makes a difference – and we know what it is. In R.P. Perry & J.C. Smart (Eds.), *Effective teaching in higher education: research and practice* (pp. 396 – 408). New York: Agathon Press

McKeachie, W.J. (1997b). Student ratings: the validity of use. *American Psychologist, 52*(11), 1218 – 1225.

McKenzie, D. (1987). The growth of school credentialling in New Zealand 1878 – 1900. In R. Openshaw & D McKenzie (Eds.), *Reinterpreting the educational past* (pp. 82 – 106). Wellington NZ: New Zealand Council for Educational Research.

McLean, M. (2001a). Rewarding teaching excellence. Can we measure teaching 'excellence'? Who should be the judge? *Medical Teacher, 23*(1), 6 – 11.

McLean, M. (2001b). Qualities attributed to an ideal educator by medical students: should faculty take cognizance? *Medical Teacher, 23*(4), 367 – 370.

Mackersay, I. (1998). *Smithy, the life of Sir Charles Kingsford Smith*. London: Little, Brown and Company.

Marriott, D.J., & Litzelman, D.K. (1998). Teaching the teachers: is it effective? *Academic Medicine,* 73 (10), S72 – S74.

Marsh, H.W. (1987). Students' evaluations of university teaching: research findings, methodological issues and directions for future research. *International Journal of Educational Research, 11*, pp. 253 – 388.

Marsh, H.W. (1994). Weighting for the right criteria in the Instructional Development and Effectiveness Assessment (IDEA) system: global and specific ratings of teaching effectiveness and their relation to course objectives. *Journal of Educational Psychology, 86*(4), 631 – 657.

Marsh, H.W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal, 38*(1), 183 – 212.

Marsh, H.W., & Hocevar, D. (1991). Students' evaluations of teaching effectiveness: the stability of mean ratings of the same teachers over a 13-year period. *Teaching & Teacher Education, 7*(4), 303 – 314.

Marsh, H.W., & Dunkin, M.J. (1997). Students' evaluations of university teaching: a multidimensional perspective. In R.P. Perry, & J.C. Smart (Eds.), *Effective teaching in higher education: research and practice* (pp. 241 – 320). New York: Agathon Press.

Marsh, H.W., & Roche, L.A. (1997). Making students' evaluation of teaching effectiveness effective. *American Psychologist, 52*(11), 1187 – 1197.

Marsh, H.W. & Roche, L.A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology, 92*(1), 202 – 228.

Marshall, P. (2000). The British Commonwealth Air Training Plan. *Round Table,* 354, pp. 267 - 279

Martin, J.R. (1998). Evaluating faculty based on student opinions: problems, implications and recommendations from Deming's Theory of Management perspective. *Issues in Accounting Education, 13*(4), 1079 – 1095.

Merton, R.K. (1957). *Social theory and social structure.* New York: The Free Press.

Moore, P.J., Lehrer, H.R., & Telfer. R.A. (2001). Quality training and learning in aviation: problems of alignment. *Journal of Air Transportation World Wide,. 6*(1), 4 – 14.

Murray, H.G. (1997). Effective teaching behaviors in the college classroom. In R.P.

Perry, & J.C. Smart (Eds.), *Effective teaching in higher education: research and practice* (pp. 171 – 204). New York: Agathon Press.

Naftulin, D.H., Ware Jr, J.E., & Donnelly, F.A. (1973). The Doctor Fox lecture: a paradigm of educational seduction. *Journal of Medical Education, 48*, pp. 630 – 635.

Nash, R. (1985). Structures of credentialling: the flow from school to work. In J. Codd, R. Harker & R. Nash (Eds.), *Political issues in New Zealand education (pp.* Palmerston North: Dunmore Press Limited.

Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education, 27*(2), 187 – 198.

Neumann, R. (2001). Disciplinary differences and university teaching. *Studies in Higher Education, 26*(2), 135 – 146.

Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Obenchain, K.M., Abernathy, T.V., & Wiest, L.R. (2001). The reliability of students' ratings of faculty teaching effectiveness. *College Teaching, 49*(3), 100 – 105.

O'Hare, D., & Chalmers, D. (1999). The incidence of incidents: a nationwide study of flight experience and exposure to accidents and incidents. *The International Journal of Aviation Psychology, 9*(1), 1 – 18.

O'Hare, D., & Roscoe, S. (1990) *Flightdeck performance.* Ames, IA: Iowa State University Press.

O'Hare, D., Wiggins, M., Batt, R., & Morrison, D. (1994). Cognitive failure analysis for aircraft accident investigation. *Ergonomics, 37*(11), 1855 – 1869.

Oosterheert, I.D., Vermunt, J.D. & Denessen, E. (2002). Assessing orientations to

learning to teach. *British Journal of Educational Psychology, 72*, pp. 41 – 64.

Orange, V. (2001). *Park: the biography of Air Chief Marshal Sir Keith Park GCB, KBE, MC, DFC, DCL.* London: Grub Street.

Patrick, J., & Smart, R.M. (1998). An empirical evaluation of teacher effectiveness: the emergence of three critical factors. *Assessment & Evaluation in Higher Education, 23*(2), 165 – 179.

Perry, R.P. (1997) Perceived control in college students: implications for instruction in higher education. In R.P Perry & J.C. Smart (Eds.), *Effective teaching in higher education: research and practice.* Agathon Press: New York

Pomson, A. (2002). The *rebbe* reworked: an inquiry into the persistence of inherited traditions of teaching. *Teaching and Teacher Education, 18*, pp. 23 – 34.

Pozo-Munoz, C., Rebolloso-Pacheco, E., & Fernandez-Ramirez, B. (2000). The 'ideal teacher'. Implications for student evaluation of teacher effectiveness. *Assessment & Evaluation in Higher Education, 25*(3), 253 – 263.

Ramsden. P. (1992). *Learning to teach in higher education.* London; New York: Routledge.

Ramsey, D., & Ramsey, P. (1996). *Feminine and masculine values in flight instructing.* Palmerston North, NZ: Massey University.

Reason, J. (2000). Human error: models and management. *British Medical Journal, 320*, pp. 768 – 770.

Rendel, D. (1975). *Civil aviation in New Zealand.* Wellington: AH & AW Reed Ltd.

Riesenberg, L.A., Biddle, W.B., & Erney, S.L. (2001). Medical students and faculty perceptions of desirable primary care teaching site characteristics. *Medical*

Education, 35(7), 660 – 665.

Roscoe S. (1980). *Aviation psychology.* Ames: Iowa University Press.

Sander, P., Stevenson, K., King, M., & Coates, D. (2000). University students' expectations of teaching. *Studies in Higher Education, 25*(3), 309 – 323.

Santhanam, E., & Hicks, O. (2002). Disciplinary, gender and course year influences on student perceptions of teaching: explorations and implications. *Teaching in Higher Education, 7*(1), 17 – 31.

Saroyan, A., & Amunsden, C. (2001). Evaluating university teaching: time to take stock. *Assessment & Evaluation in Higher Education, 26*(4), 341 – 353.

Sexton, J.B., Thomas, E.J., & Helmreich, R.L. (2000). Error, stress and teamwork in medicine and aviation: cross sectional surveys. *British Medical Journal, 320*, pp. 745 – 749.

Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: love me, love my lectures. *Assessment & Evaluation in Higher Education, 25*(4), 397 – 405.

Shute, N. (1966). *Slide rule.* London: Heinemann Educational Books Ltd.

Snell, L., Tallett, S., Haist, S., Hays, R., Norcini, J., Prince, K., Rothman, A., & Rowe, R. (2000). A review of the evaluation of clinical teaching: new perspectives and challenges. *Medical Education. 34*, pp 862 – 870.

Spencer, K.J., & Schmelkin, L.P. (2002). Student perspectives on teaching and its evaluation. *Assessment & Evaluation in Higher Education, 27*(5), 397 – 409.

Stringer, M., & Irwing, P. (1998). Students' evaluations of teaching effectiveness: a

structural modelling approach. *British Journal of Educational Psychology, 68*, pp. 409 – 426.

Tabachnick, B.G., & Fiedell, L.S. (2001) *Using multivariate statistics* (4th ed.). Boston MA: Allyn & Bacon.

Telfer, R. A. (1993a). Effective aviation instruction. In R. A. Telfer (Ed.), *Aviation instruction and training* (pp. 219 – 236). Aldershot, UK: Ashgate.

Telfer, R. A. (1993b). What makes aviation instruction different? In R. A. Telfer (Ed.), *Aviation instruction and training,* (pp. 209 – 218). Aldershot, UK: Ashgate.

Termoehlen, J. (1987). *The role of the flight instructor – An important psychological factor in flying training.* In R.S. Jensen (Ed.), *Proceedings of 4th International Symposium on aviation psychology* (pp. 522 – 528). Columbus: Ohio State University Department of Aviation.

Thomas, M.J.W. (2001) Enhancing instructional systems: the development of a tool for evaluating instructor and student performance. In M. J.W. Thomas, & G. J. F. Hunt (Eds.), *Enhancing professionalism in aviation.* Auckland: Massey University School of Aviation.

Timpson, W.W., & Desley, A. (1997). Rethinking student evaluations and the improvement of teaching: instruments for change at the University of Queensland. *Studies in Higher Education, 22*(1), 55 – 66.

Trollip, S., & Jensen, R. (1991). *Human factors for general aviation.* Englewood,Co: Jepperson Sanderson.

Turney, M.A., Bishop, J.C., Karp, M.R., Niemczyk, M., Sitler, R.L. & Green, M.F.

(2002). National survey results: retention of women in collegiate aviation. *Journal of Air Transportation World Wide, 7*(1), 69 – 92.

Vu, T.R., Marriott, D.J., Skeff, K.M., Stratos, G.A., & Litzelman, D.K. (1997). Student attitudes in evaluating courses, faculty, and peers. *Academic Medicine, 72*(10), S7 – S9.

Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: a brief review. *Assessment and Evaluation in Higher Education, 23*(2), 191 – 212.

Watkins, D., Marsh, H.W., & Young, D. (1987). Evaluating tertiary teaching: a New Zealand perspective. *Teaching and Teacher Education, 3*(1), 41 – 53.

Watson, S. (2003). Closing the feedback loop: ensuring effective action from student feedback. *Tertiary Education and Management,* 9(2), 145 – 157.

Weimer, M. (1997) Exploring the Implications: From Research to Practice. In R.P. Perry, & J.C. Smart (Eds.), *Effective teaching in higher education: research and practice* (pp. 411 - 425). New York: Agathon Press.

Wilkinson, L. (1999). Statisical methods in psychology journals. *American Psychologist, 54*(8), 594 – 604.

Wilson, K.L., & Lizzio, A. (1997). The development, validation and application of the course experience questionnaire. *Studies in Higher Education, 22*(1), 33 – 53.

Woodhouse, M.G. (2001, August). *The Role of Flight (and Ground) Training and Assessment in Improving Flight Safety in General Aviation.* Paper presented at Category A Flight Instructor Seminar, Wellington, NZ.

Worthington, A.C. (2002). The impact of student perceptions and characteristics on

teacher evaluations: a case study in finance education. *Assessment & Evaluation in Higher Education, 27*(1), 49 – 64.