University of North Dakota
UND Scholarly Commons

Theses and Dissertations

Theses, Dissertations, and Senior Projects

12-1-2009

Major Airline Pilot Hiring, Low Cost Versus Legacy

Daniel Lang

Follow this and additional works at: https://commons.und.edu/theses

MAJOR AIRLINE PILOT HIRING, LOW COST VERSUS LEGACY

by

Daniel J. Lang
Bachelor of Business Administration, University of North Dakota 1993

An Independent Study
Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
December
2009

PERMISSION

Title   MAJOR AIRLINE PILOT HIRING, LOW COST VERSUS LEGACY

Department   Aviation

Degree   Master of Science

In presenting this independent study in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my independent study work or, in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this independent study or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my independent study.

Signature _____

Date   _____

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER

I. INTRODUCTION

There are 15 major airlines and 15 different types of initial pilot selection processes. Airlines generally publish low minimum application requirements in order to comply with non discrimination laws in the United States. However, most applicants will have far exceeded these minimums in order to be competitive. As airlines become more selective, their applicant pool becomes smaller resulting in a more homogenous group. Since all major airlines draw from this relatively small homogenous group, why does every airline have a different selection process? Most major airlines administer written and or computer tests, simulator evaluations, psychological tests, and aeronautical knowledge tests. Most, if not all, legacy airlines administer simulator evaluations during initial screening while none of the low cost airlines use them. However, no low cost airlines use simulator evaluations.

This study examined the differences in pilot selection between the low cost airline Jet Blue and the legacy major airline Continental. The results indicate that neither carrier has attempted to validate their selection process. Although various tests or phases might select pilots with desired traits neither airline can show if they are actually selection their desired pilot. Both airlines would be much better served with a structured type selection processes that is founded on proven research and statistical analysis. When asked about the hiring process at major airlines, Kit Darby, the owner of a career counseling and

information service company for commercial pilots (Air Inc.), said the hiring process is "consistently inconsistent " (Damos, 2003).

JetBlue and Continental Airlines were chosen because they have separated themselves from other airlines by consistently achieving multiple awards for excellent customer service and desirable corporate cultures. Even though both airlines focus on selecting applicants with extensive experience and superior customer service skills, essentially trying to hire the same type of pilot, they both take significantly different approaches in their pilot selection process.

The results of this study will enable future major airline pilot applicants to evaluate an airline's selection process. A thorough evaluation of airlines hiring practices will allow pilots to discover if they are the best fit for that airline and how to go about positioning themselves for future hiring opportunities. This study will also provide airlines with potential suggestions for improving their selection process.

Qualifications

All airlines have a list of minimum qualifications that all applicants must have achieved before they are allowed to apply for employment (Table 1).

Table 1. Representative airline minimum pilot hiring requirements

	JetBlue	Continental
Total hours	4000	1500
Turbine hours	1000	1000 Fixed wing
Multi Engine hours	1000	1000 Fixed wing or military single engine jet

PIC hours	1000/500 turbine	700 Fixed wing
Instrument hours	N/A	200
Certificates	Airline Transport Pilot (ATP) written	ATP written
Medical	Class I	Class I with no limitations
Vision	20/20 correctable	20/20 correctable
Education	4 year college degree preferred	4 year college degree
Other	Computer literate	N/A

Very few applicants will be hired with the minimum requirements. The pursuit of jobs at major airlines is extremely competitive since a pilot's career earnings can exceed $10 million (Darby, 2008). Therefore, the pursuit of a major airline job is extremely competitive. As a result of such competition, most pilots will achieve what the airline industry calls "competitive minimum". These minimums are usually much higher than the stated minimums but not high enough to be discriminatory. The average major airline qualifications are a reflection of the competitive minimums for that time period (Table 2).

Table 2. Average major airline qualifications (Darby, 2008)

	Military	Civilian
Average age	37	37
College Degree %	93	97
ATP certificate %	93	92
At least one type rating %	57	60

Average flight time (hours)	4,262	5,558
Pilot-in-Command (PIC) flight time	2,334	3,317
Turboprop (hours)	1,567	1,953
Jet (hours)	2,802	3,449
Multi-engine (hours)	3,427	4,779
% of total hired	32	68

Note: 33% of pilots hired have a mixed, military and civilian background.

Applicability

Rebok, Qiang, Baker, McCarthy, & Li (2005) studied 3,306 commuter air carrier and air taxi pilots from 1987-1997 in an attempt to discover if age and experience had an affect on violations rates per flight hour. The author discovered that pilots with flight times greater than 5000 hours had reduced violation rates of up to 60 percent. Childs, Spears, & Prophet (1983) conducted a two year study to assess the skill retention levels of relatively low time pilots and discovered that recency of flight time had a significantly positive affect on instrument proficiency.

Diehl (1991) studied U.S. airline accidents from 1987-1989 and verified that the majority were caused by decisional errors, not technical competency. Ree & Earles (1992) discovered that intelligence would be an appropriate criterion to measure success in either flight or training performance as opposed to technical knowledge. LeMaster & Gray (1974) investigated the use of an Air Force instrument trainer as a selection device in order to identify critical flying abilities possessed by Undergraduate Pilot Training (UPT) candidates. Attrition was not satisfactorily predicted due to causes

other than lack of flying skill. Finally, Majesty (1976) discovered that although simulator performance was significantly correlated with pass-fail in undergraduate pilot training, for ab initio students it offered no more validity than that of the current paper and pencil selection tests.

The Problem

Major airlines use unproven methods to select their future pilots. All airlines use various forms of minimum qualifications in order to measure a baseline level of competence. Most legacy airlines use a simulator evaluation and/or or technical knowledge tests to determine an applicants technical skill level. This study will challenge the current selection methods used to hire major airline pilots.

Research Questions

1. What are the hiring differences between JetBlue and Continental?
2. Does the differences in the selection process result in different types of employees?

Definitions

Ab initio Training, derived from Latin meaning "from the very beginning", Refers to a method of training students, with little or no flying experience, by emphasizing from the very beginning skills necessary in becoming a professional airline pilot. Most programs attempt to train and place a pilot in a commercial airline job with approximately 350 hours of flight time. In contrast, most general aviation pilot training programs are designed simply to achieve a particular certificate or rating and not necessarily prepare a pilot for a career in aviation. Since ab initio programs provide

experience to achieve a career as a professional pilot, training costs tend to be significantly more than general aviation programs.

Aviation informational tests often are comprised of questions about aviation in general. Psychologists often use them to measure motivation for flying.

ATP stands for Airline Transport Certificate. This is the highest level of pilot certification and is required for all captains at major airlines.

Basic Aptitude Test (BAT) is a computer administered battery that measures psychomotor skills, information processing, and attitude towards risk. It has been validated to predict a pass/fail in UPT.

Commuting is a term used to describe pilots who do not live in the same city where they are based. These pilots must commute from their resident city to their base city prior to their scheduled flights.

Competitive minimums are higher than stated minimums but not high enough to be discriminatory. For example an airline might state the minimums to apply would be 1000 hours of total flight time but have hundreds or thousands of applicants with 3000 hours which would become the competitive minimum. Although a company might not state that a pilot needs a college degree or possess multiple type ratings, the current pool of applicants might all have those qualifications thereby setting the competitive minimums higher than the stated minimums.

Flight training devices (FTD) are often exact replicas of the cockpits of an aircraft. Pilots use these trainers to practice their knowledge and skills on their assigned aircraft. There are many levels of FTD's, however none offer motion.

High fidelity flight training device are FTD's often with images and sounds of the surrounding environment.

Jumpseat refers to a seat or seats in the cockpit of all major airline aircraft that were originally intended for an FAA representative to observe the operating flight crew. However, jumpseats are most often occupied by pilots for commuting purposes.

Legacy airline refers to airlines that generally have a hub and spoke network. In the US it refers to those airlines that flew interstate routes prior to the Airline Deregulation Act of 1978.

Low cost airline, budget, or discount carrier refers to those airlines that tend to offer lower priced tickets without some of the amenities like first class or airport lounges.

Major airlines or carrier is a designation given by the US Department of Transportation (DOT) to describe airlines with more than $1 billion of revenue during the fiscal year.

Narrow body aircraft is typically an airliner with one aisle. The most common narrow body aircraft is the B737 with the largest being the B757-300 which seats up to 289.

Operations tempo (OPTEMPO) are measures that describe the tempo of a soldier's military experience, including total number of military deployments, deployment intensity (number of deployments/ number of years in the military), and training intensity (number of days on training exercises in the past six months).

Flight Operational Quality Assurance (FOQA) is a method that most major airlines use to collect, store, and analyze data from an aircrafts flight recorder in order to increase safety and reduce costs.

Flow through is a term used in the airline industry to describe a process where a major airlines commuter pilots are hired (flow) directly from the commuter (through) to the major airline with a far less rigorous selection process.

Performance consistency is a culmination of past history, garnered from a resume, paired with recent performance, interview or tests, to produce an opinion of an applicant.

Screening systems eliminate applicants who do not meet the minimum requirements (i.e. background checks, online applications) (Damos, 2003).

Selection systems Identifies the best applicants through interviews, tests, or simulator evaluations. A selection system often consists of (Damos, 2003);

1. A criterion identifying what type of pilot the company intends to hire.

2. Testing that could include; intelligence, personality, or skill.

3. Hiring model

    a. Single hurdle administers all tests to all applicants and is the most expensive.

    b. Multi hurdle administers the least expensive test first and then the more expensive tests thereafter to the remaining applicants.

    c. Progressive method administers tests throughout the selection and training process. This type of method is common in ab initio programs.

Targeted Selection is a behavioral interview approach used to collect job related behavior from an applicants past history.

Topographical fallacy is the view that because the variables look different or require a different response they must measure different constructs. Just because an

interviewer asks different questions in a different way than a paper and pencil test, it does not mean that it measures a different construct.

Unstable approach refers to aircraft approaching an airfield below a predetermined altitude, usually 1000 feet above ground level, in an unstable condition. An unstable condition would be defined as; excessive airspeed, excessive descent rate, aircraft not aligned with the runway, aircraft not properly configured for landing, or an improper power setting.

Wide body aircraft are usually twin isle aircraft with seating capacities greater than 200. The most numerous wide body aircraft is the Boeing 747 (B747) and the largest is the Airbus 380 (A380) which seats up to 853 people.

Assumptions and Limitations

The responses from the questionnaire are indicative of the organization as a whole.

Continental's flow through procedure is no longer in effect and is outside the purview of this paper because it bypasses the normal selection process.

I have only specifically analyzed the hiring processes of JetBlue and Continental Airlines. Therefore, the suggestions only apply specifically to those two airlines. However, this study would serve as a good interview preparation for all major airline applicants as well as a warning signal to other major airlines who have allowed their selection process to change over time with out any validation.

Disclosures

I have interviewed at JetBlue in 2005. I was subsequently offered a job but had secured career employment elsewhere.

Representative Airlines

JetBlue Airways was founded by Airline entrepreneur David Neeleman as a low cost domestic airline in February 1999. Initially, all service originated from New York's JFK international airport with one aircraft type, the Airbus 320 (A320). However, in the last ten years JetBlue has grown into an international airline operating over 140 aircraft (A320 and Embraer 190) with more than 2000 pilots. JetBlue has won numerous customer service awards including the 2008 Top Low Cost Airline for Customer Satisfaction by J.D. Power and Associates and the 2007 Best Low Cost/No Frills Airline by Official Airline Guide (OAG) (Jet Blue History: Awards and Accolades (n.d.)).

Continental Airlines was founded in 1934 by Walter T. Varney and Louis Mueller as Varney Speed Lines. Continental's first flight was flown under its original name Varney Speed Lines from Pueblo, Colorado to El Paso, Texas with stops in Las Vegas, Santa Fe and Albuquerque. Varney Speed Lines changed its name to Continental Airlines in 1937. The airlines first airplane was a single engine four passenger Lockheed Vega. Continental's first pressurized airplane was the Convair 240 and its first jet was the Boeing 707. In 75 years the airline's fleet has grown to over 340 aircraft including the Boeing 737, 757, 767 and the 777 with more than 4700 pilots operating in 56 countries. Continental's many awards include; No. 1 in Customer Satisfaction from Frequent Flyer/J.D. Power & Associates (June 2006) among traditional network carriers in North

America and Best Airline Based in North America from AOG in 2006 (Continental

Airlines Awards (n.d)).

Similarities

1. Both major airlines

2. Fly multiple types of jet aircraft

3. Heavy east coast presence

4. Operate internationally

5. Won multiple awards for customer service and positive corporate culture

Differences

1. JetBlue is a low cost airline and Continental is a legacy airline

2. Continental has been in operation for over 70 years and has survived many

   challenges like;

   a. Deregulation in 1978

   b. Merger with Texas International in 1982, under Frank Lorenzo

   c. Filed first Chapter 11 bankruptcy in 1983

   d. Filed second Chapter 11 bankruptcy in 1990

3. JetBlue has only been in existence for 10 years and only reported its first

   quarterly loss in February of 2006

4. Although they both operate internationally, JetBlue serves 53 cities with

   600 daily flights operating in 12 countries. Continental serves 341 cities

   with over 3000 daily flights operating into 56 countries.

5. Jet Blue has two types of narrow body aircraft where as Continental has four types of aircraft consisting of two narrow bodied and two wide bodied as well as another wide body on order (B787).

Selection Process

In order to become a pilot at JetBlue an applicant must complete a number of phases.   Initially, a pilot needs to complete an online application which allows the Pilot Interview Name Selection committee (PINS), made up of experienced Jet Blue pilots, to prescreen potential candidates. Competitive candidates are then selected to attend Phase 1. Phase 1 consists of an email requesting additional information and an invitation for an interview. The additional information requested of Phase 1 candidates consists of; PRIA information, three references who can attest to the candidate's flying skills, and a "Shining Moment" story. The Shining Moment story consists of a story 200 words or less about "a time when you went out of your way to meet the needs of a customer or a fellow employee. What did you do, and what was the response?" The Shining Moment is very important part of the interview that will enable a pilot to convey how well they will fit into Jet Blue's company culture. Candidates are invited to a one day interview where they will submit all the requested paperwork and take part in two interviews. The first interview is conducted by a panel comprised of one People Department representative and a Jet Blue pilot. This interview is conducted using the Targeted Selection method and lasts for about 50 minutes. Jet Blue uses Targeted Selection along with the Shining Moment story to determine both motivational and corporate fit. The panel interview is often concluded with questions addressing a pilot's willingness and ability to live in domicile, regional fit. After the panel interview the applicant then will interview with a

single pilot for about 30 minutes, although the purpose of this specific interview is unknown. After the two interviews, the three representatives meet to discuss the overall assessment rating (OAR) and determine if the applicant should proceed to Phase II. If the pilot is selected for Phase II, a thorough background check is done and all records are passed on to the Pilot Hiring Committee (PHC). If the applicant is successful they are called within 6 to 8 weeks. If the pilot is unsuccessful they are normally notified by mail in the same time period.

In order to become a pilot with Continental Airlines, an applicant must also complete a phased process. Initially, a pilot needs to complete a very thorough online application listing their flying experience, work history, educational accomplishments, residential history, and driving record. Continental outsources their application data collection to airlineapps.com. Continental Airlines' minimum qualifications are similar to JetBlue's minimums, however, Continental's applicants must have 1,000 hours of PIC in a turbine-powered airplane. Those candidates that meet the competitive minimums will be forwarded to the human resource department who will then determine which candidates are selected for the interview process. When a candidate is selected they are generally called to set up a time when they can meet with the representatives in Houston at Continental's training center. The interview process is generally a two day review that consists of a panel interview and a simulator evaluation. The panel interview, which lasts about one hour, is conducted by a current Continental pilot and a human resource representative. Raters ask a mix of behavioral, situational, and technical questions oriented at determining the applicant's cultural and motivational fit as well as technical aptitude. All applicants will also be asked to perform a flight profile in a high fidelity

training device, usually a Boeing 737 or MD80, by a AQP certified instructor. The final decision for each applicant is made by a selection board that attempts to determine if the pilot will be a good fit with Continental Airlines (Rudder, 2001).

LITERATURE REVIEW

History of pilot selection

Military

The Wright brothers were the first to pilot a powered airplane in 1903. The US Government utilized aircraft for war by making its first purchase for the Army in 1909 and later adding the first machine gun in 1912. The advent of WWI greatly increased the need for aircraft and for the volunteers that would pilot those airplanes. To the surprise of most, initial data from the front line indicated fighter pilots were dying not because of mechanical malfunctions but due to human error (Griffin, 1996). This revelation created an impetus to design tests to successfully determine which volunteers would possess both the physical and mental aptitude to pass the required training and more importantly not crash the airplanes.

Initial pilot selection stereotypically revolved around the "ace" pilot. In fact, one of the earliest attempts to screen for "emotional stability" was the measurement of hand tremors after firing a pistol (Griffin, 1996). However, after WWI a more rational approach developed. The English selection methods revolved around physiological parameters, such as high altitude effects and blood pressure, where as the French were using a combination including sensory testing, like the Barany chair test for disorientation and motion sickness. On the other hand, U.S. aviation psychologists were focusing on psychomotor tests because of their practical validity. WWII created

14

incredible demand for pilots who would be able to train in a short period of time while cost constraints increased the need to minimize the number of failures. A study, known as the Pensacola 1000 Aviator Study, was commissioned to evaluate 900 Navy aviators using psychological, paper and pencil physiological, and psychomotor tests. This study indicated that psychological and psychomotor tests were more valid in predicting flight training success than physiological ones. However, reliability and quality control problems led to the discontinuation of psychomotor tests in 1953. In the 1970's, computers brought about many attempts to change the means by which pilots were selected. Computer based simulators were hypothesized to increase validity by testing applicant's job sample tasks. These evaluations were found to be mostly redundant with paper and pencil tests as well as overly expensive and time consuming. Computers have allowed test presentation reliability, increased accuracy in data collection, and the ability to track response speed (latency). Psychomotor tests made a comeback after computers were developed because they improved the efficiency, reliability, validity of pilot selection.

Currently all military services select aviators based on an evaluation of their medical, mental, and physical fitness condition in addition to a number of automated and or paper and pencil tests.

Army

The Army selects pilots with the:

1. Alternate Flight Aptitude Selection Test (AFAST), an automated cognitive and psychomotor test administered to initial applicants. The AFAST is a flight aptitude selection test designed to predict attrition.

2. Multi-track Test Battery (MTTB) for helicopter task assignments. The MTTB is a paper and pencil test that classifies ab initio students into training tracks and is predictive of training performance.

Air Force

The Air Force (USAF) uses the Pilot Candidate Selection Method (PCSM) which is comprised of the (Carretta & Ree, 1994):

1. College performance

2. Previous flying experience

3. Paper and pencil Air Force Officer Qualifying Test (AFOQT) contains a mixture of cognitive and background tests intended to predict attrition and pass-fail in training.

4. Automated BAT which utilizes cognitive and psychomotor tests intended to predict pass-fail in training.

Navy

The Navy uses a paper and pencil only selection test called the Naval Aviation Selection Test Battery (NASTB). The NASTB is a valid predictor of ground school and flight training grades and a lesser predictor of attrition from training. The tests consist of:

1. The academic qualification rating (AQR) which is a valid predictor of academic performance in ground school.

2. Pilot Flight Aptitude Rating (PFAR) which is a valid predictor of flight grades in primary flight training.

3. Pilot Biographical Inventory (PBI) which is a valid predictor of attrition through primary flight training (Stricker, 2005; Talleur, Henry, Emanuel, Rantanen, & Bradshaw, 2003).

Perspectives from the History of military hiring

1. The military tests are unique from other pilot hiring applications because;

    a. Risk taking is thought to be a desired trait (Damos, 1996)

    b. Almost all initial applicants are inexperienced and need to be screened for entry into an ab initio program.

    c. The criterion for military selection is usually success in initial pilot training. Naval aviator training in 1993 was over $900,000 and is surely greater than $ 1 million in today's dollars (Griffin, 1996).

2. Military does extensive research in pilot hiring in order to reduce training failures and prevent attrition.

3. Army and Air Force have some automated tests while the Navy does not, although the Automated Pilot Examination (APEX) is in testing and due for full implementation in the near future

There is little comprehensive research available describing the history of civilian pilot selection. In fact, it was not until 1994 that civilian pilots were consistently hired in greater numbers at major airlines than military pilots (Darby, 2008). One of the possible reasons that today's airlines have such a wide variety of selection methods is likely because there seems to be little desire in the aviation community for validation of the current selection processes (Damos, 1996). After thorough research I discovered no US airlines and only two international carriers, Cathay Pacific and Qantas Airlines, which

attempted to validate their hiring process in the last 15 years. Cathay's pilot selection, although initially based on the Royal Air Force, was adapted because of the different qualities between civilian and military pilots. Second officers (S/O) might have extensive experience or be hired into their ab initio program. First officers (F/O) often have previous extensive airline experience. The selection process starts with a "sift" of an application to identify those "deemed potentially suitable" who are invited to fill out an application, although the minimums were not defined. Suitable applicants are then interviewed in their home country, which consists of MICROPAT psychomotor tests and a short medical check. Successful applicants are then invited to the final stage in Hong Kong. The final stage consists of a structured interview, a technical interview, a knowledge test, a full flight simulator evaluation, the Cattell Sixteen Personality Factor Questionnaire, and a full medical examination. The process intends to predict training outcome and prevent operational failure based on the perceived relationship with general ability, psychomotor coordination, and various personality traits (Bartram & Baxter, 1996). I was unable to analyze the Qantas hiring process because the article was not available in any research database in which I had access.

Suarez, Barborek, Nikore, & Hunter (1994) described some of the rules affecting pilot certification in the European Union. A 1991 Flight Crew Licensing Medical Group of the European Civil Aviation Conference (ECAC) adopted psychological requirements for commercial pilots. The requirements state that no holder of an aviation medical certificate shall have any psychological deficiencies. However, psychological tests were not mandated to obtain a pilot medical. Goeters (1995) felt that psychological test should be mandatory, however Murphy (1995) disagreed due to inadequate standards of

acceptable performance, a lack of an accepted test battery, and affects of cultural differences affects. A pilot representative summarized the conflict by saying "we pilots watch with considerable interest and irritation as psychologists and physicians argue over their respective competence to conduct assessments of our psychological fitness to earn a living. The fundamental inappropriateness of this has clearly struck neither party." He finished by saying, "the reality of these discussions are more concerned with power and control than with technical competence" (Johnson, 1996). Suarez suggested a broader approach that all pilots should be tested for cognitive ability, conscientiousness or integrity, and job knowledge. Johnson (1996) weighed in to the testing controversy by stating that there was no consensus of even best or correct battery of tests. Sweden psychologically tests all of its pilots and Germany uses limited testing for some licensing but Canada and the FAA require no testing. Borman & Motowidlo (1997) hypothesized that global competitiveness will increase the effort levels and adaptability of employees as well as popularize team based and customer focused organizations. Borman attempts to breakdown the performance of an employee to differentiate between task and contextual. Task performance is the effectiveness of the employees to perform tasks that contribute directly or indirectly to the core of the organization. Examples would be a pilot flying an ILS or a firefighter performing a rescue operation. Contextual activities contribute to the organizational effectiveness in social and psychological ways. Examples of contextual activities would be volunteering to perform task activities that are not a part of the job. An example would be a pilot cleaning the cabin of an aircraft in order to keep the next flight on time. If the quality of customer service is a measure of contextual performance, then contextual performance would be a criteria and personality would be

19

used as a predictor. The validity of personality tests is under debate, however if contextual performance could be measured separately, this would lead to an increase in the validity of the tests. Finally, employees should consider evidence that supervisors consider subordinate tasks and contextual performance equally when rating their subordinates.

Tests

In order to properly evaluate any future employee an employer must understand the reliability and validity of current tests. Current research for pilot selection testing revolves around:

1. Psychological measures

2. Personality

3. Cognitive ability/Intelligence/Aptitude

4. Psychomotor abilities

5. Biographical data

6. CRM skills

7. Technical skills/Simulator evaluation

Psychological measures

Most research about psychological measures in pilot selection has revolved around performance in pilot training. However, the composition of psychological tests are not universal. Psychological testing can include various aptitude, personality, and psychomotor tests. Johnson (1996) reviewed the contemporary trends in the psychological testing of pilots. Many militaries use psychological test batteries to screen

ab initio pilots (Martinussen, 1996). Those batteries of tests have proven to be successful predictors of training success but not overall performance or construct validity. Therefore, psychological tests have proven to be a legitimate screening tool for ab initio pilots however they do not accurately predict an individual's skill level. In fact, when given to large groups of experienced pilots, many have failed the tests even though they have proven to be adequate pilots. The author theorizes that these tests are often elevated and are implied to be able to predict; job skills, abilities, and suitability because of their success in correlating with ab initio training success.

<div align="center">Personality</div>

Personality tests have also been used to predict the success of flight training. Jessup & Jessup (1971) gave the Eysenck Personality Inventory to 205 pilot cadets in the Royal Air Force. His research discovered that neurotic introverts failed the most and stable introverts failed the least. The author did note a potential flaw in that an introvert might be successful in the training phase but could eventually have performance problems on the line. Siem (1992) gave a personality inventory test to 509 pilot candidates in the USAF only to discover that it did not enhance the current selection provided by the BAT. Street, Helton, & Dolgin (1992) gave the automated Pilot Personality Questionnaire to 211 Navy pilot candidates and found the competitiveness scale helped with the predictability of overall training success but was the least predictive of the dimensions. Finally, Goeters (1993) administered a Personality Research Form, self description, and the Temperament Structure Scales, personality questionnaire, to 300 pilots in training for a major airline. He discovered that personality questionnaires given to new hire

candidates are often unreliable because responses are often skewed since applicants tend

to falsify their responses as a result of their social desirability to be hired.

<center>Cognitive Ability/Intelligence/Aptitude</center>

The measure of intelligence, for pilot selection, is a broad and extensive subject. Ree &

Earles (1992) provided a general roadmap of the history of intelligence measures. The

Binet intelligence test was used in the selection of conscripts during World War I.

Following the war, personnel selection evolved by testing human abilities, aptitudes, as

well as intelligence. Specific aptitude theory involves the weighting of various abilities

like quantitative, verbal, mechanical or spatial and has been found to be superior

compared to predicting job training and performance by intelligence alone. General

ability, g, is how psychologists identify intelligence because of the positive correlation of

scores on various kinds of tests. The positive correlations were caused by general

cognitive ability, which is measured in common by all tests. Specific aptitude theory has

proven to be successful and is the foundation behind current selection tests like the

Armed Services Vocational Aptitude Battery (ASVAB) and the Air Force Officer

Qualifying Test. Intelligence measures in these tests have validated the predictability in

both training success and job performance (Ree & Carretta, 1996; Walters, Miller, &

Ree, 1993; Smith, 1994). However, McClelland (1993) disagreed with the assertion that

intelligence was the best predictor of job performance, specifically, that cognitive ability

better predicted performance levels. McClelland stated that "correlation does not equal

causation" and hypothesized that job performance could be a function of a third variable

like education.

<center>22</center>

Psychomotor skills

There is disagreement with some measures and their correlations, however, psychomotor ability is largely agreed upon to predict training success (Griffin, 1996; Carretta & Ree, 1994). Burke, Hobson, & Linsky (1997) conducted a meta-analysis of three studies involving computerized tests measuring instrument comprehension and psychomotor ability across the Royal Air Force, Turkish Air Force, and the British Armey Air Corps. The author discovered that psychomotor tests proved successful in predicting training success and would be useful in both military and civilian settings. Additionally, training success could be better predicted by adding cognitive ability (g) measures because of its dominance in the later stages of training, a position later confirmed by Carretta & Ree (1994).

Biographical Inventory

Biographical data consists of any information concerning the background of an applicant. Military selection usually involves the collection of biographical information which is then summarized into a score. Those scores are then used to predict retention in training (Stricker, 2005). Although commercial airlines collect biographical information about an applicant, I could not find any evidence that major airlines use this data in any standardized way. Most US airlines compile biographical data from:

1. PRIA and driving records to assess safety.

2. Background checks to assess liability for the corporation

3. Flight time is used to assess of safety, competence, trainability and operational performance (Li, Baker, Grabowski, & Rebok, 2001; Higgins, 2005;

Schvaneveldt, Beringer, & Lamonica, 2001; Doane, Sohn, & Jodlowski, 2004; Guilkey, 1997)

4. Type ratings are used to assess safety (Li et al.)

5. The attainment of a college degree is an assessment of competence and trainability.

Williams, Albert, & Blower (1999) studied the validity of various military tests including the biographical inventory. The Aviation Selection Test Battery (ASTB) was administered to 2852 naval aviators and student flight officers enrolled in primary flight training between November 1993 and July 1998. The test is an initial filter that is used to segment applicants into pilot or flight officers. Flight officers perform navigation and weapons systems duties. The ASTB consists of:

1. Academic Qualification Rating (AQR) which is a valid predictor of academic performance in ground school.

2. Pilot Flight Aptitude Rating (PFAR) which is a valid predictor of flight grades in primary flight training.

3. Pilot Biographical Inventory (PBI) which is a valid predictor of attrition through primary flight training.

Stricker (2005) specifically targeted the biographical inventory in order to identify the dimensions underlying the inventory's design and the relations of those dimensions to retention of students in naval aviation training. The author tested 1,819 student naval aviators who applied for training between 1986 and 1988, largely on the basis of their test and physical exam scores. The test battery consisted of:

1. Student aviator scores which are non cognitive measures of retention in primary flight training (PFT) and basic flight training (BFT).

2. Aviation Qualification Test (AQT) which measures general ability and predicts success in Naval Aviation Schools Command. Items in this section would include mathematics and vocabulary.

3. Mechanical comprehension tests which measures the ability to perceive psychological relationships and understand mechanics

4. Spatial perception test which measures spatial orientation

5. Aerospace information test which measures general aerospace knowledge

6. Biographical inventory composite score which is the sum of the student aviator score and the aerospace information score

7. Flight aptitude rating which is the sum of the; mechanical score, special apperception score, and the biographical inventory composite score. This score is a valid predictor of success in PFT and BFT.

However, Goeters et al. (1993) addressed various fallibilities when using biographical data in commercial pilot selection. Biographical data is generally used under the premise that a person's history is indicative of future behavior. The limitations of this data would include:

1. Life events are subjective in regards to their affect on someone's past

2. There is not a standardized method of interpreting the data

3. Reviews that included a person's entire life have not provided useful information (Goeters et al.)

4. There does not seem to be a link between the data and pilot success (Maschke & Hormann, 1989)

Goeters suggested companies follow the lead of the German Aerospace Research Establishment Testing Center. Pilots are scored based on a personality questionnaire, biographical data, and behavior test. This score results in a hypothesis which is checked and modified with a psychological interview in order to achieve a final assessment.

CRM Skills

"Planes don't cause accidents, pilots do." This is a provocative but unfortunately true statement. Jensen & Benel (1977) analyzed data from the National Transportation Safety Board's (NTSB) database have revealed the majority of U.S. general aviation accidents, from 1970 to 1974, were the result of a pilot's poor decision making. Diehl (1991) later validated Jensen's theories with an analysis of U.S airline accidents from 1987 to 1989. Contrary to initial stereotypes of dominant captains who 'save the day', piloting jet airplanes involves communication, resource utilization, and teamwork much more than just technical skill. A parallel can be developed with other occupations that require both skill and teamwork. Jones (1974) studied group effectiveness as a function of individual effectiveness in four professional sports; baseball, basketball, football and tennis. He discovered that 90 percent of team wins could be explained by superior athletes in sports, like baseball, where athleticism was the deciding factor. However, only 35 percent of basketball wins could be similarly explained, therefore verifying his theory that effectiveness in environments with highly skilled participants depends on the interdependency of the task. These factors led to the establishment of the Crew Resource Management concept in aviation. Helmreich (1999) conducted a qualitative research

26

project into pilot personalities by describing the origins, validation, and limitations of CRM. Although early research identified human error as a major cause in aviation accidents it was a workshop called Resource Management on the Flightdeck, put on by NASA in 1979, which coined the term Cockpit Resource Management (Jensen, 1977). A United Airlines crash in 1978 spurred that company to be one of the first to implement an organized training program aimed at addressing dictatorial captains and first officers who lacked self confidence, personality types identified as mitigating factors in the accident. A major change to CRM emerged in the late 1980's, shifting the focus from single personality traits to cockpit crew dynamics. The program was renamed Crew Resource Management which included training centered around team building, situational awareness, and stress management. The 1990's produced another significant change when training began to include additional factors Such as an organization's cultural impact on safety, flight deck automation, and the recognition and assessment of human factors issues. It was also recognized there were many other resources available and that other groups like flight attendants, dispatchers, and maintenance should be trained and included in CRM. The most current version of CRM is integrated into the training and qualification of flight crews, called Advanced Qualification Program (AQP). AQP allows airlines to decrease their training cycles by developing organization specific training addressing human factors issues in each aspect of training. AQP training is orientated around daily flight operation scenarios and usually includes a Line Orientated Flight Training (LOFT) opposed to the more traditional structured training and checking around a FAA developed flight scenario. For example, an approved airline is allowed to lengthen the normal time between proficiency evaluations, usually from 6 months to 12

months, by integrating data from their FOQA program into their training syllabus. An increase in unstable approaches, discovered in the FOQA data, would be reflected in the next year's training syllabus by providing training to correct the specific problem. CRM is mandated by the FAA for every U.S. airline but a validation of its benefits has been difficult. There are so few accidents in the U.S. that any measure of number or rate cannot be used. However, one of the accepted means used to discover if CRM is having an impact is to measure changes in behavior after CRM training. Research has concluded that if CRM training is favorably received there can be a significant and positive change in behavior. Additionally, behavior change is improved when the training is included into the line orientated flight training (LOFT) and recurrent training (Helmreich 1993). However, the future trend in CRM involves recognition that errors cannot be totally prevented and that human error is inevitable. If this premise is accepted then CRM would be a counter balance to human error with three principle lines of defense. The first would be error avoidance. The second would involve trapping incipient errors before they are committed. The final means would be to mitigate the consequences of the errors that are not trapped. However, in order for this approach to work, employees and management must develop a framework that promotes identifying the source and nature of errors. A current program that has proven successful is called Aviation Safety Action Program (ASAP). ASAP is a confidential and voluntary, non-punitive, program that allows pilots to disclose errors and safety concerns. This data is then discussed by management, the FAA, and often an employee representative to determine an educational means to either change behavior or procedure in order to increase safety. Often, ASAP data is used to develop training procedures that are then integrated into an AQP program.

For example, pilots submitted multiple ASAP reports of course deviation errors during instrument approaches. The data would be discussed during an ASAP meeting, verified with FOQA data, and integrated into the AQP training. The training might include a focus on course monitoring with crewmember deviation callouts and implemented during ground school and verified during LOFT. Selecting pilots who are able to employ CRM is essential to airline safety but requires an organization to thoroughly examine if their processes are both reliable and valid. Helmreich & Wilhelm (1991) discovered a major limitation to CRM when up to five percent of major airline pilots were not only resistant to improved crew communication techniques but even increased their opposition to its premise after training. Over 15,000 flight crews from 12 airlines and military organizations participated in a study about attitudes involving flight deck management and personal capabilities. From self-reported evaluations before and after an initial CRM seminar, some crews, called boomerangs, actually reacted negatively and consequently had a worse attitude after the training. Boomerangs often lack communication skills and rebel against attempts to correct their short comings. The problem with this type of personality is that CRM is specifically orientated at these crewmembers but is ineffective. Helmreich discovered a large concentration of boomerang personality types in what he described as the "no stuff" cluster. The author created the cluster to describe pilots with low motivation and interpersonal sensitivity traits after they completed the Cockpit Management Attitude Questionnaire (CMAQ). "No stuff" personalities showed a negative initial perception of CRM and led to negative training outcomes in a sample of military pilots. A principle finding of this study was that there should be a renewed emphasis on selection procedures in order to choose pilots with personalities accepting of

CRM concepts. Chidester, Helmreich, Gregorich, & Geis (1991) provided a look into the militaries use of CRM by conducting two samples of Air Force pilots, during recurrent CRM, to determine if personality clusters could be defined and if so were they related to either performance or training designed to increase crew coordination. The results of this study found three unique clusters of pilots.

1. Positive instrumental interpersonal pilots were characterized as goal orientated self confident, and kind.

2. Negative instrumental pilots were characterized as egotistical, dictatorial, and not aware of the feelings of others.

3. Low motivational pilots were characterized by less than average in regards to self confidence or kindness and greater than average in the areas of ego and dictatorial.

The study concluded that some personality types are stable and resistant to change and a crew's performance can be limited by just one pilot with this type of personality Furthermore, high performance of an airline crew is related to the personality types of those pilots who are interested in mastering new and challenging tasks. Low performance is often found in hostile and/or arrogant crewmembers. Although current selection methods revolve around the evaluation of technical skills there is significant evidence that personality is more indicative of performance. Chidester (1990) provided insight into how certain communication traits affect performance at US airlines. A survey of airline pilots was taken to determine the normative patterns, individual's moods, and sleep patterns during short haul flight operations. Additionally, crews performed a one and half day trip with five flight segments in a full motion simulator. The effectiveness of

the crews was judged by an experienced airline captain who had no knowledge of the crew's personality profile. The number and severity of errors was rated by two additional experienced observers. Crews that had negative instrumental captains made more errors during abnormal flights and were rated as ineffective on the first day of evaluation. This study seems to support Helmreich's & Wilhelm (1991) theory that although negative instrumental captains can create a moderately negative effect, the low motivational flight crews have the greater potential in decrease performance. With the foundation of cause and affect of communication established, current hiring methods need to be discussed. Currently, the military does not specifically select pilots based on their CRM skills. However, Hedge, Bruskiewicz, Borman, Hanson, Logan, & Siem (2000) attempted to develop and validate a CRM skills test for the selection of Air Force transport pilots called the Situational Test of Aircrew Response Styles (STARS). Seven hundred ninety two C-130 crews from the Air National Guard (ANG) and Air Force Reserves (AFR) participated in the test validation. The test consisted of a series of questions that presented participants job-relevant situations and asked them to decide which of several actions would be the most and least effective. Results indicated that a significant relationship existed between the results of the CRM test and the behavior based ratings of aircraft commander job performance.

Technical knowledge & skills (Simulator Evaluation)/Experience

Airlines attempt to measure an applicant's skill in variety of ways. The most common means to measure that skill would include a measure of one or more of the following; biographical inventory (discussed above), technical knowledge tests, simulator evaluations, and experience. As noted earlier, all U.S. legacy airlines evaluate applicants

using some form of technical knowledge and/or skills based test where as none of the low cost airline do. Damitz, Manzey, Kleinmann, & Severin (2003) noted that a poll given to a major German airline, to determine the most important pilot qualities, resulted in stress resistance being rated the highest. Other important qualities were decision making, communication, and leadership. Notably absent was any reference to aviator skill, technical knowledge, or experience. I could not find any research involving how or why U.S. legacy airlines use technical evaluations of either knowledge or skill. Commonly, some airlines will ask an applicant technical questions concerning international operations or about an approach procedure. However, those measures seem to lack both reliability and validity because the answers to those questions are often available from consultants or from many different websites involving airline hiring discussions. The most common measures used by major airlines to assess an applicant's skill level are simulator evaluations and a review of his/her flight experience. Hörmann & Maschke (1996) conducted an assessment of a personality questionnaire versus a simulator check coupled with a flight experience review. Two hundred seventy four pilots applying for employment at a European airline were the subjects of the study. Job success was determined by a review of their training records after a three year period. A flight experience review alone predicted job success 68% of the time, whereas the simulator evaluation only increased the predictability by 6% and the personality questionnaire an additional 5%. An important finding from this study was that flying hours, airline experience, age, and command experience can actually be negative predictors of job success. Also of note was that jet experience was an insignificant predictor of whether a candidate became successful on the job. Rebok et al. (2005) used data files from the

32

NTSB and the FAA to determine violation rates of 3,306 commuter air carrier and air taxi pilots. He discovered that violations decreased as total flight time exceeded 5,000 hours but only up to 10,000 hours and then again increased with plot over the age of 50. Schvaneveldt et al. (2001) studied how experienced and non experienced pilots organize information. Sixty one pilots were used in two projects to discover that experienced pilots assigned higher priorities to all phases of flight compared with less experienced pilots. However, there was little difference between the groups in regards to how they conceptually organized information. Higgins (2005) used a personal computer aircraft training device (PCATD) on 20 pilots, between the ages of 18-35, to discover how recent flight time impacted proficiency. He discovered that greater total time and recent flight experience increased the likelihood of performing procedures correctly. Li et al. (2001) studied data files from the NTSB about airline crashes from 1983-1996. The author discovered greater total time showed a significant protective effect in regards to pilot error. Furthermore, pilot error in general aviation and commuter accidents decreased as he number of ratings a pilot held increased. Finally, the U.S. military has researched the idea of using flight simulation to evaluate applicant's skill level. The result was a slight increase in the predictability of initial flight training success but with only a small increase in variance beyond the current tests (Griffin & Koonce, 1996; LeMaster & Gray, 1974). In the final analysis it was decided that the significant increase in costs and time to administer the tests was not justified.

<center>Assessment centers</center>

Do tests and interviews adequately assess the actual behavior of applicants? Most U.S. airlines use interviews and tests to exclusively to measure applicant's behavioral

<center>33</center>

tendencies. However, some European airlines feel that current research shows that assessment centers are better at determining those qualities. Damitz et al. (2003) examined the validity of assessment centers in pilot selection. A major German airline prescreened 3109 applicants with a cognitive ability test from 1994-1996. One thousand thirty six pilots passed and entered an assessment center. The assessment center conducted four exercises that intended to measure two categories; interpersonal competence and critical performance related aspects. One experienced airline captain and one psychologist rated each applicant on eighteen dimensions chosen to be representative of the two categories. Although an overall assessment rating (OAR) was calculated, the final decision was reached by comparing the scores of each applicant's two category ratings to a normative database. In order to validate the findings, the criterion validity of each successful assessment center applicant was a score achieved during their training phase. The results indicate the overall rating was a valid predictor of the criterion. However, the pilots' rating of interpersonal competence was not significant. It was hypothesized that psychologists are educated to assess behavior and that pilots lack this professional background. The author suggested more research into the validity of trained instructors versus line pilot's behavioral ratings. Jones, Herriot, Long, & Drakeley (1991) also discovered problems with assessor ratings when attempting to improve the Admiralty Interview Board assessment center for the Royal Navy. He discovered that OAR's could be improved almost 20 percent by using a multiple regression equation opposed to allowing the assessor's opinion to carry such significant weight.

Raters

Most US major airline pilot selection processes entail two phases of evaluation; an interview and a simulator evaluation. However, the role and impact of the rater is a matter of much discussion. Keenan (1978) studied the impact of interviewer training and experience on the outcome of 514 recruitment interviews. The study discovered training decreased overall bias but, when paired with experience, had little impact on overall ratings. Pulakos, Schmitt, Whitney, & Smith (1996) agreed when she discovered that consensus discussions between interviewers provided little significance to the validity of their judgments over the mechanical averages of individual ratings. The issue of improving the quality of simulator evaluations was addressed by Goldsmith & Johnson (2002) in a study that offered guidance for training evaluators. The author identified two problems, common to any type of evaluation, which needed to be addressed. The first problem involves extreme grading. Whereas the evaluator is either too lenient, the Halo Effect, or too strict which can be identified when the evaluator's mean grades vary significantly from the population mean. The second is an evaluator's failure to utilize the entire grading scale and can be identified when the evaluator's standard deviation of grades is significantly below that of the population. Both of these errors can be corrected with training, although a company would have to commit to developing a means of data keeping and monitoring in order for it to be successful.

<center>Falsification</center>

A potential challenge to the validity of personality tests and personal interviews is the ability for an applicant to falsify their responses. Goffin & Christensen (2003) discovered that interviewers are susceptible to and their decisions can be affected by applicant falsifying responses. The MMPI personality test given by some major airlines to screen

<center>35</center>

applicants for socially deviant behavior can be susceptible to falsifying. Pilots who

passed the test could simply advise other applicants to how they answered questions.

Jessup & Jessup (1971) and Goeters et al. (1993) discovered that personality tests were

especially susceptible and unless accounted for, would make these them unreliable.

Goffin & Christiansen (2003) surveyed researchers from the US, Canada, and Europe in

an attempt to test validity and determine the extent that tests were corrected for faking.

The author discovered that social desirability measures are related to staple personality

traits therefore making it difficult to discover if applicants within the high range were

faking or were simply very positive people. In fact "paradoxically, it is the most honest

and upstanding citizen that these scaled would lead us to accuse of lying." On average,

personality tests scores that have been corrected for faking have not shown greater

validity to uncorrected tests. It was his opinion that denying someone a job due to a

correction that was not based on empirical evidence was not justifiable. Levashina &

Campion (2006) provided a comprehensive literature review directly addressing faking

and deception. The author submitted "faking is a function of capacity, willingness, and

opportunity." Furthermore, "faking is an intentional distortion or falsification of

responses on measures in order to create a specific impression or provide the best

answer." However, the author tries to discern a difference between fakers and those who

attempt impression management. Fakers are those that intentionally engage in deceptive

and dishonest behavior intended to create a good impression. Impression management

often involves an applicant's attempt to look good with out being untruthful. For

example, Kit Darby is an airline pilot consultant who will help candidates prepare for

interviews. His primary philosophy is for the candidate to get the interviewer to focus on

the positive aspects of their resume and/or personality traits in order to attain the desired position. However, coaching or interview preparation services, such as this, do negatively affect the validity of an interview and could possibly lead to more faking (Stevens & Kristof, 1995; Maurer, Solamon, Andrews, & Troxtel, 2001). There are a number of problem areas, during the selection process, which could lead to a high risk of faking. Questions that are unverifiable and hypothetical as well as those that measure organizational fit or interests create the greatest potential for faking. Situational interviews are generally higher risk opposed to structural interviews that are highly standardized. One of the most effective ways management can reduce faking would be to tell all applicants that their responses will be verified or checked.

<center>Interview</center>

Although the methods and tests have changed over time the interview remains the foundation for pilot selection. Wiesner & Cronshaw (1988) defined the interview as "an interpersonal interaction of limited duration between one or more interviewers and job seekers for the purpose of identifying interviewee knowledge, skills, abilities, and behaviors that may be predictive of success in subsequent employment." The author conducted a worldwide literature review and performed a meta-analysis on the impact on validity of the format and structure of individual vs. board and unstructured vs. structured interviews. This study discovered board ratings were more predicatively valid than independent ratings and structured interviews had validity coefficients twice those of unstructured. However, in order to maintain interview validity, an interview must be reliable, structured, and use job analytic information in its development.

Many researchers have studied various ways to improve interview validity. Walters et al. (1993) hypothesized that expert pilots would be able to offer greater validity to the current selection process and that paper and pencil as well as computer administered tests were insufficient for pilot selection. Ten active duty expert pilots and four personnel research psychologists developed questions considering job relevancy and fairness. Two hundred twenty three pilot trainees with the U.S. Air Force were chosen based on their AFOQT scores to participate in a study that investigated the validity of the structured interview. Each trainee was measured with; AFOQT, BAT, and structured interview with the pass-fail in pilot training serving as the criterion. The structured interview was found to be valid but not incrementally so. The most valid predictor of performance was math knowledge on the AFOQT. However, this study had numerous potential flaws. The first was that expert pilots determined attributes they felt necessary for pilot success without conducting any job analysis. The second was a likelihood of bias because the interviewers were aware of each candidate's scores on the AFOQT and BAT prior to the interview. Damos (2003) also considers the structural interview method to be the most effective and provided a comprehensive review of pilot selection techniques for the Flight Safety Foundation. The author defined two distinct types of pilot selection systems; unstructured and structured. Unstructured systems are characterized by:

1. No interviewer training

2. No systematic method of combining information from the various parts of the selection process into any comprehensive review

3. Hiring decisions are based mostly on the decision of the hiring manager

Structured systems are characterized by:

1. Initial analysis of knowledge, skills, abilities, and personality traits desired by the company

2. Decision makers are trained on various interview techniques including bias and Halo Effect.

3. The hiring decision is made through a specific process, decision aide, which does not eliminate the selection personnel but guides them into making a valid and reliable decision.

4. Hiring results are quantifiable through a process and feedback loop which predicts future pilot performance.

Structured selection systems were determined more valid and reliable due to; training, structure, and statistical evaluation of the process.

The two most common types of interviews used at major airlines are behavioral and situational. These two types of interviews are often used in conjunction with each other. Behavioral interviews focus on past experiences and behaviors of an applicant where as situational interviews focus on the reaction of the interviewer to a given situation. Sue-Chan & Latham (2004) studied the validity of the situational interview. Seventy five managers in an executive MBA program were tested for their cognitive ability and interviewed by a panel using the situational method. The criterion variables were team playing ability and academic performance. Although the interview and cognitive ability predicted academic performance only the interview predicted team playing ability.

Bias

    Dipboye & Gehrlein (1993) studied the affect of attractiveness on interview bias.

Photographs of 506 randomly selected applicants were rated on their attractiveness.

The criterion variables examined were interview evaluations and final admission

decisions. Attractive applicants were evaluated as having higher qualifications but

this had no impact on the final admission decision. On average, attractive applicants

performed no better academically but did have greater social skills than their

counterparts. The author suggested that human resources use alternate selection

procedures, when available, and to not solely rely on interview evaluations in the

final hiring decision. Another common interview bias is the processing of the

information from the application. Macan & Dipboye (1994) studied how interviewers

observed and evaluated interviewees after forming initial impressions from the

applicant's paper credentials. The author concluded that bias existed when

interviewers were allowed to review applicant's credentials prior to the interview.

Pulakos & Schmitt (1996) confirmed this bias when her study determined that prior

interview knowledge of a candidate's cognitive ability tended to reduce the overall

interview validity.

II. METHOD

I have compiled a survey comprised of questions that intend to discover:

1. What are the qualifications of the evaluators and how are they trained?

2. How are the applicants evaluated?

3. What guidance does the company provide to the evaluators during the

   selection process?

4. What do the phases of the interview intend to evaluate?

5. How is the selection process reviewed and improved?

I have conducted qualitative research on the issues and provided a narrative of

pilot selection to determine:

1. Are airlines correctly using their selection procedures?

2. Are they hiring their intended pilot?

Finally I have provided suggestions that can be made in order for an airline to hire

their target employee?

Subjects

I have contacted a person that has or is directly involved in the hiring process at

both JetBlue and Continental Airlines. The contact has been provided with and has

answered the questionnaire in its entirety.

III. RESULTS

Pilot evaluators consist of human resource representatives, management pilots, line pilots, and instructor pilots. Continental and JetBlue require both human resource representatives and pilots to be properly trained and assessed as evaluators. Both airlines have similar AQP programs that require instructors to undergo initial, recurrent, and evaluator training. However, the only recurrent training required was for the Continental human resource representatives.

Continental uses questions randomly derived from the FAR's, Jeppesen Charts, general instrument knowledge, and current applicant's aircraft knowledge to determine technical knowledge. They also use a simulator evaluation to determine an applicant's piloting skill. JetBlue considers previous experience, derived from the online application, to determine an applicant's technical knowledge and skill. Continental uses a pass/fail grading sheet for both the interview and the simulator evaluation opposed to JetBlue's point scale for the interview. JetBlue seems to use their mission statement of "bringing humanity back to flying" as a centerpiece for determining all questions used in their Targeted Selection interview process, whereas Continental uses their mission statement in conjunction with procedures established in their flight operations, flight standards, and training mission statements as guides in determining parameters for applicants qualifications. Continental uses feedback from its evaluators, captains, as well as an eight month probationary period in an effort to determine the effectiveness of its selection process. JetBlue simply uses feedback from its evaluators and instructors to determine the

effectiveness of its selection process. Continental has not validated their selection process or changed it significantly in the last 10 years. JetBlue has not changed their interview process in the last 10 years but, according to their current application, has decreased its minimum qualifications by no longer requiring jet time. Both airlines encourage applicants to use all outside references and consultants to fully prepare for the selection process.

IV. DISCUSSION

The survey results indicate that pilot selection at major airlines varies

significantly from company to company. Selection practices are often a result of opinions

from various key decision makers within the company and not valid research. Although

both companies are attempting to hire model employees, both processes are filled with

barriers that might eliminate the best applicants. JetBlue's focus on selecting pilots who

meet their core values is likely to be the most successful. However, a lack of recurrent

rater training and potential bias because of knowledge of an applicant's background could

be limiting factors (Levashina & Campion, 2006; Macan & Dipboye, 1994). Continental

has long established hiring practices which is common among legacy airlines. However,

phases that include simulator evaluations and random technical questions are likely

ineffective and have been or are being eliminated because of their lack of validity and

reliability (Griffin & Koonce, 1996; Carretta & Ree, 1994). JetBlue will likely hire pilots

who are more culture-orientated, while Continental will likely hire someone more

focused on technical skills and knowledge. Both companies have successful corporate

cultures that prove they care about their employees but more needs to be done if they are

to select pilots with attributes that reflect their mission statements.

Suggestions

Selecting the right pilot is essential for airlines that desire to maintain safety and their

coperate culture. However, in order to produce a reliable product, airlines should develop

44

a structured pilot selection system. Many pilot selectors believe their hiring programs are structured. However, few would have an answer as to how personality types like the "boomerangs" got selected and were subsequently not eliminated prior to completing probation (Helmreich & Wilhelm, 1991). Damos (2003) suggested five steps in the development of a true structured employee selection system:

1. Conduct a job analysis to determine the knowledge, skills, abilities and personality traits essential to job success. Both airlines have AQP programs that have already developed task analyses. These analyses can be modified to determine the appropriate knowledge, skills, and abilities. Appropriate personality traits can be determined by either an outside consultant or a qualified manager.

2. Identify selection tests to identify selected criteria. Human recourse specialists are often used to select the appropriate tests that meet the needs of the airline.

3. Performance measures need to be selected in order to serve as the criteria for the selection system. Airlines often use only training scores to measure the criterion. This results in pilots that are proficient in the "school house" but lack the appropriate skills on the line. An airline could include many measures including; probationary scores, proficiency evaluation scores, disciplinary actions, and commendations to an overall measure of the criterion.

4. Administer the tests, collect the data, and determine if they actually predict job success. Data needs to be collected over a prolonged period of time before the company will know if they have selected the correct criterion. For example, an employee might complete training in the required footprint and score well during probation. However, Helmreich, Sawin, & Carsrud (1986) identified the initial

period of employment as the "honeymoon." There were significant correlations only after that period between personality predictors and performance which can be measured by a personality test. The reason for this behavior was that people wanted to do well during their initial probationary employment but often reverted to their "intrinsic achievement motives" thereafter.

5. Implement a monitoring system. The system should include statistical analysis of the alerts, notifying management of weaknesses in the criterion. An example might be an annual cumulative measure of a pilot's performance during training paired with their cancellation rate. This system will allow the company a constant evaluation of their selection system by identifying those individuals who perform below standards. Management could use this system to identify the criteria and change their selection procession in order to strengthen the criterion.

Future research opportunities

The lack of structure and consistency in pilot selection at major airlines beckons for further research.

1. A validation study of the pilot selection process at any or all major U.S. airlines could prove invaluable to furthering the industry's knowledge into how each individual test affects the overall selection method.

2. A study comparing a U.S. and an international major airline pilot selection method would provide valuable information to now global airlines as they compete for the best pilots worldwide.

3. Research focused on individual selection tests and how they affect the criterion. For example, a comparison of the training failure rates at carriers who do and do

not use pilot selection methods involving a simulator evaluation? The military discovered that a simulator evaluation is not the most effective measure to determine success in pilot training. Why do many airlines continue to expend significant resources on this method?

APPENDIX

Questionnaire

1. Human resource evaluators

    a. Do you know if the human resources evaluators are trained and assessed

       as evaluators (circle one)?

        i. Yes

        ii. No

    b. If yes, do they attend recurrent training for evaluators (circle one)?

        i. Yes

        ii. No

    c. Are HR evaluators encouraged to provide feedback to improve the process

       (circle one)?

        i. Yes

        ii. No

2. Pilot Evaluators.

    a. Are the pilot evaluator's qualified AQP instructors (circle one)?

        i. Yes

        ii. No

    b. Are pilot evaluators, who are not AQP instructors, trained and assessed as

       evaluators (circle one)?

        i. Yes

        ii. No

    c. If yes, do they attend recurrent training for evaluators (circle one)?

      i. Yes

      ii. No

d. Does your airline's AQP program require evaluator training and

   calibration sessions (circle one)?

      i. Yes

      ii. No

e. Does your airline's AQP program require annual evaluator training and

   calibration sessions (circle one)?

      i. Yes

      ii. No

f. *Does your airline have a Gold Standard session or video review (circle

   one)?

      i. Yes

      ii. No

g. For both the simulator evaluation and human resource interview, what

   type of grading sheet is used (circle and or explain)?

      i. Point scale

      ii. Pass/Fail

      iii. Other. Explain if you can _____

         _____.

h. To your knowledge, were the pilot evaluators consulted when the grading

   scale was created (circle one)?

      i. Yes

      ii. No

    i. Are pilot evaluators encouraged to provide feedback to improve the

    process (circle one)?

        i. Yes

        ii. No

3. Company

    a. Does the company have a mission statement (circle one)?

        i. Yes

        ii. No

    b. Does the company consider the mission statement when hiring new

    employees (circle one)?

        i. Yes

        ii. No

    c. If yes how so?

    _____

    _____.

4. Applicants

    a. What is the company's position on airline consultants?

        i. Applicants are encouraged to use all means in order to prepare

        themselves for the interview.

        ii. Applicants are discouraged from using interview preperation.

        iii. Other, explain.

_____

_____.

5. Interview

    a. To your knowledge how has the intervew and or application process changed over the last 10 years (be specific as possible, i.e. more or less subjective, CRM more or less evaluated, applicants must posess more or less PIC,jet time, or educational requirements to be competitive, ect.)

        i. _____

           _____

           _____

           _____

           _____

           _____

           _____.

    b. What do technical questions evaluate (Continental only)?

        i. _____

           _____.

    c. What does the simulator ride evaluate (Continental only)?

        i. _____

           _____.

6. Probation

    a. What does probation consist of?

i. _____

_____

_____.

    b. Has the evaluation of probationary employees changed in the last 10 years

(circle one)?

      i. Yes

      ii. No

    c. If yes, please explain how and if to your knowledge why.

i. _____

_____

_____.

7. Overall selection process

    a. Has the airline ever attempted to validate the hiring process (circle one)?

      i. Yes

      ii. No

    b. If yes provide details.

i. _____

_____.

* Gold Standard session or video is when a group of subject matter experts

develops a video or review for all the other instructors that defines unsatisfactory

from barely satisfactory and FAA satisfactory from company satisfactory. The

Gold Standard session is ultimately used to develop a common frame of reference

and to establish a base line for evaluators.

REFERENCES

Bartram, D., & Baxter, P. (1996). Validation of the Cathay Pacific Airways Pilot

    Selection Program. *International Journal of Aviation Psychology*, *6*(2), 149.

Borman, W., & Motowidlo, S. (1997). Task Performance and Contextual Performance:

    The Meaning for Personnel Selection Research. *Human Performance*, *10*(2), 99.

Burke, E., Hobson, C., & Linsky, C. (1997). Large Sample Validations of Three General

    Predictors of Pilot Training Success. *International Journal of Aviation*

    Psychology, *7*(3), 225.

Carretta, T., & Ree, M. (1994). Pilot-Candidate Selection Method: Sources of Validity.

    International Journal of Aviation Psychology, *4*(2), 103.

Chidester, T. (1990). Trends and individual differences in response to short-haul flight

    operations. *Aviation, Space and Environmental Medicine, 61*, 132-138.

Chidester, T., Helmreich, R., Gregorich, S., & Geis, C. (1991). Pilot Personality and

    Crew Coordination: Implications for Training and Selection. *International*

    Journal of Aviation Psychology, *1*(1), 25.

Childs, J., Spears, W., & Prophet, W. (1983). Private pilot flight skills retention 8, 16,

    and 24 months following certification (DOT Publication No. DOT/FAA/CT-

    83/34, Seville TR-83–17). Retrieved May 10, 2009 from

    http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD

    A133400

Continental Airlines Awards (n.d.). Retrieved May 10, 2009 from

    http://www.continental.com/WEB/enUS/content/company/globalcitizenship/awar

    ds.aspx

Damitz, M., Manzey, D., Kleinmann, M., & Severin, K. (2003). Assessment Center for

    Pilot Selection: Construct and Criterion Validity and the Impact of Assessor Type.

    Applied Psychology: An International Review, *52*(2), 193-212.

Damos, D. (2003). Pilot selection systems help predict performance. *Flight Safety Digest*,

    3-12. Retrieved May 10, 2009 from www.flightsafety.org/fsd/fsd_feb03.pdf

Damos, D. (1996). Pilot Selection Batteries: Shortcomings and Perspectives.

    International Journal of Aviation Psychology, *6*(2), 199

Darby, K. (2008). US Airline Pilot Careers FAA Forecast Conference DCA 3/11/08.

    Retrieved May 10, 2009 from

    www.faa.gov/news/conferences_events/aviation_forecast_2008/agenda_presentati

    on/media/kit_darby.pdf

Diehl, A. (1991). Cockpit Management Training Programs: Who, What, Why, When, and

    Where. *Flying Safety*, 47 (12). Retrieved May 1, 2009 from http://www.crm-

    devel.org/resources/paper/diehl.htm

Dipboye, R., & Gehrlein, T. (1993). Attractiveness Bias in the Interview: Exploring the

    Boundaries of an Effect. *Basic & Applied Social Psychology*, *14*(3), 317-328.

Doane, S., Sohn, Y., & Jodlowski, M. (2004). Pilot Ability to Anticipate the

    Consequences of Flight Actions as a Function of Expertise. *Human Factors*,

    46(1), 92-103.

Goeters, K., Timmermann, B., & Maschke, P. (1993). The Construction of Personality

    Questionnaires for Selection of Aviation Personnel. *International Journal of*

    Aviation Psychology, *3*(2), 123.

Goeters, K. (1995). Psychological evaluation of pilots: The present regulations and
arguments for their application. *Aviation psychology: Training and selection.
Proceedings of the 21ˢᵗ Conference of the European Association for Aviation
Psychology (EAAP)*, 2, 149-156.

Goffin, R., & Christiansen, N. (2003). Correcting Personality Tests for Faking: A Review
of Popular Personality Tests and an Initial Survey of Researchers. *International
Journal of Selection & Assessment*, *11*(4), 340-344. doi:10.1111/j.0965-
075X.2003.00256.x

Goldsmith, T., & Johnson, P. (2002). Assessing and Improving Evaluation of Aircrew
Performance. *International Journal of Aviation Psychology*, *12*(3), 223-240.

Gregorich, S., Helmreich, R., Wilhelm, J., & Chideser, T. (1989) Personality based
clusters as predictors of aviator attitudes and performance. *Proceedings of the fifth
symposium on aviation psychology*, 75, 697-702.

Griffin, G., & Koonce, J. (1996). Review of Psychomotor Skills in Pilot Selection
Research of the U. S. Military Services. *International Journal of Aviation
Psychology*, 6(2), 125.

Guilkey, J. (1997). An investigation of aviator problem-solving skills as they relate to
amount of total flight time. *Dissertation Abstracts International,* 58(5-B), 270.
(UMI No. AAM9731628)

Hedge, J., Bruskiewicz, K., Borman, W., Hanson, M., Logan, K., & Siem, F. (2000).
Selecting Pilots With Crew Resource Management Skills. *International Journal
of Aviation Psychology*, *10*(4), 377-392.

Helmreich, R., Sawin, L., & Carsrud, A. (1986). The honeymoon effect in job

    performance: Temporal increases in the predictive power of achievement

    motivation. *Journal of Applied Psychology*. 71(2), 185-188. Retrieved May 23,

    2009 from http://psycnet.apa.org/index.cfm?fa=main.doiLanding&uid=1986-

    23189-001

Helmreich, R., & Wilhelm, J. (1991). Outcomes of Crew Resource Management

    Training. *International Journal of Aviation Psychology*, *1*(4), 287.

Helmreich, R., Merritt, A., & Wilhelm, J. (1999). The Evolution of Crew Resource

    Management Training in Commercial Aviation. *International Journal of Aviation

    Psychology*, *9*(1), 19.

Higgins, J. (2005) Total and recent flight experience as predictors of pilot instrument

    performance. Unpublished thesis, University of North Dakota, Grand Forks.

Hörmann, H., & Maschke, P. (1996). On the Relation Between Personality and Job

    Performance of Airline Pilots . *International Journal of Aviation Psychology*,

    6(2), 171.

Jensen, R. & Benel, R. (1977). Judgment Evaluation and Instruction in Civil Pilot

    Training. DOT/FAA Report RD-78-24.

Jessup, G., & Jessup, H. (1971). Validity of the Eysenck Personality Inventory in Pilot

    Selection. *Occupational Psychology*, *45*(2), 111-123.

Jet Blue History: Awards and Accolades (n.d.). Retrieved May 10, 2009 from

    http://www.jetblue.com/about/ourcompany/history/about_ourhistory.html

Johnston, N. (1996). Psychological Testing and Pilot Licensing. *International Journal of

    Aviation Psychology*, *6*(2), 179.

Jones, M. (1974). Regressing Group on Individual Effectiveness. *Organizational Behavior & Human Performance*, *11*(3), 426-451.

Jones, A., Herriot, P., Long, B., & Drakeley, R. (1991). Attempting to improve the validity of a well-established assessment centre. *Journal of Occupational Psychology*, *64*(1), 1-21.

Keenan, A.(1978). SELECTION INTERVIEW OUTCOMES IN RELATION TO INTERVIEWER TRAINING AND EXPERIENCE. *Journal of Social Psychology*, *106*(2), 249.

LeMaster, W. & Gray, T. (1974). Ground training devices in a job sample approach to UPT selection and screening (Tech. Rep. No. ADHRL TR-74-86)

Levashina, J., & Campion, M. (2006). A Model of Faking Likelihood in the Employment Interview. *International Journal of Selection & Assessment*, *14*(4), 299-316. doi:10.1111/j.1468-2389.2006.00353.x

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

Macan, T., & Dipboye, R. (1994). The Effects of the Application on Processing of Information From the Employment Interview. *Journal of Applied Social Psychology*, *24*(14), 1291-1314.

Majesty, M. (1976). The making of an aircraft pilot selection system. (Tech. Rep. No. ADHRL TR-75-58).

Martinussen, M. (1996). Psychological Measures As Predictors of Pilot Performance: A Meta-Analysis. *International Journal of Aviation Psychology*, *6*(1), 1.

Maschke, P., & Hormann, H. (1989). The prediction of professional success considering licensed pilots: The validity of flight experience in comparison with standardized psychological aptitude tests. (DLR Research Rep. DLR-FB 89-53).

Maurer, T., Solamon, J., Andrews, K., & Troxtel, D. (2001). Interviewee coaching, preparation strategies, and response strategies in relation to performance in situational employment interviews: An extension of Maurer, Solamon, and Troxtel, 1998. *Journal of Applied Psychology*, 86, 709–717.

McClelland, D. (1993). Intelligence Is Not the Best Predictor of Job Performance. *Current Directions in Psychological Science*, *2*(1), 5-6. doi:10.1111/1467-8721.ep10770447

Murphy, T. (1995). JAA psychological testing of pilots: Objections and alarms. Aviation psychology training and selection. *Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP)*, 2, 157-163.

Pulakos, E., Schmitt, N. Whitney, D., & Smith, M. (1996). Individual differences in interviewer ratings: The impact of standardization, consensus discussion, and sampling error on the validity of the structured interview. *Personnel Psychology* 49, 85-102.

Pulakos, E., & Schmitt, N. (1996). An Evaluation of Two Strategies for Reducing Adverse Impact and Their Effects on Criterion-Related Validity. *Human Performance*, *9*(3), 241.

Rebok, G., Qiang, Y., Baker, S., McCarthy, M., & Li, G. (2005). Age, Flight Experience, and Violation Risk in Mature Commuter and Air Taxi Pilots. *International*

Journal of Aviation Psychology, *15*(4), 363-374.

doi:10.1207/s15327108ijap1504_4

Ree, M., & Carretta, T. (1996). Central Role of g in Military Pilot Selection. *International Journal of Aviation Psychology*, *6*(2), 111.

Ree, M., & Earles, J. (1992). Intelligence Is the Best Predictor of Job Performance. *Current Directions in Psychological Science*, *1*(3), 86-89. doi:10.1111/1467-8721.ep10768746

Rudder, R. (2001) Continental Airlines: Maintaining the Momentum. *Airline Pilot Careers*, 6 (7), 18-25.

Schvaneveldt, R., Beringer, D., & Lamonica, J. (2001). Priority and Organization of Information Accessed by Pilots in Various Phases of Flight. *International Journal of Aviation Psychology*, *11*(3), 253-280.

Siem, F. (1992). Predictive Validity of an Automated Personality Inventory for Air Force Pilot Selection. *International Journal of Aviation Psychology*, *2*(4), 261.

Smith, M. (1994). A theory of the validity of predictors in selection. *Journal of Occupational & Organizational Psychology*, *67*(1), 13-31.

Stevens, C., & Kristof, A. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology*, 80, 587–606.

Street, D., Helton, K., Dolgin, D. (1992) The unique contribution of selected personality tests to the prediction of success in naval pilot training. (Rep. No. NAMRL-1374). Retrieved May 10, 2009 from http://handle.dtic.mil/100.2/ADA258144

Stricker, L. (2005). The Biographical Inventory in Naval Aviation Selection: Inside the Black Box. *Military Psychology*, *17*(1), 55-67. doi:10.1207/s15327876mp1701_5

Suarez, J., Barborek, S., Nikore, V., & Hunter, D. (1994) Current trends in pilot hiring and selection (FAA Memo. No. AAM-240-94-1).

Sue-Chan, C., & Latham, G. (2004). The Situational Interview as a Predictor of Academic and Team Performance: A Study of the Mediating Effects of Cognitive Ability and Emotional Intelligence. *International Journal of Selection & Assessment*, *12*(4), 312-320. doi:10.1111/j.0965-075X.2004.00286.x

Talleur, D., Henry L., T., Emanuel Jr., T., Rantanen, E., & Bradshaw, G. (2003). Personal Computer Aviation Training Devices: Their Effectiveness for Maintaining Instrument Currency. *International Journal of Aviation Psychology*, *13*(4), 387.

Walters, L., Miller, M., & Ree, M. (1993). Structured Interviews for Pilot Selection: No Incremental Validity. *International Journal of Aviation Psychology*, *3*(1), 25.

Wiesner, W., & Cronshaw, S. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology*, *61*(4), 275-290.

Williams, H., Albert, A., & Blower, B. (1999) Selection of Officers for U.S. Naval Aviation Training. Paper presented at the RTO HFM Workshop on "Officer Selection", held in Monterey, CA. Retrieved May, 09, 2009 from ftp://ftp.rta.nato.int/PubFullText/RTO/MP/RTO-MP-055/MP-055-18.pdf