2013

Development of a Pilot Selection System for a Midwestern University Aviation Program

Kathryn Wilson
Minnesota State University - Mankato

Development of a Pilot Selection System for a Midwestern University Aviation Program

By

Kathryn Wilson

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

In

Industrial/Organizational Psychology

Minnesota State University, Mankato

Mankato, Minnesota

May 2013

Abstract

The research for the following paper titled, Development of a Pilot Selection System for a Midwestern University Aviation Program and authored by Kathryn Wilson was conducted at Minnesota State University, Mankato located in Mankato, Minnesota. This study was a requirement of the Industrial/Organizational Psychology Master's Program and was conducted during the 2012-2013 academic school year.

This paper discusses the evaluation of an existing selection system for a Midwestern University's Aviation Program and attempts to find significant predictors of pilot performance using personality measures including the Five Factor Scale, Cockpit Management Attitudes Questionnaire, Self Monitoring Scale, an Integrity Scale, and cognitive measures including Block Counting, Rotated Blocks and Numerical Reasoning. Data from 24 student pilots was examined with bivariate correlations and stepwise regression and results indicate personality plays a role in predicting successful pilot performance. The CMAQ, extraversion and block counting measures were positively correlated with facets of performance including decision making in-flight, consistently arriving on-time for lessons, and situational awareness in-flight, respectively. Also, agreeableness predicted a negative relationship with situational awareness in-flight and instructor rating of performance. Based on the results, it is suggested the current selection measure consist of the Five Factor Scale, CMAQ, Block Counting and Rotated Blocks measures and be validated in the future to evaluate reliability.

Table of Contents

Development of a Pilot Selection System for a Midwestern University Aviation Program

Chidester, Helmreich, Gregorich and Geis (1991) stated, "The performance of pilots can be construed as a product of skill, attitude, and personality factors." (p.25). Research has found that a structured selection system can accurately identify applicants who meet the pilot requirements effectively (Damos, 2003). The high cost of training and concern for pilot and passenger safety coupled with the knowledge that major airlines have found predictive validity with intelligence tests as an indication of future training performance for experienced pilots, makes an effective selection system even more necessary. Several major United States airlines require a battery of written or computer-based tests including measures of cognitive ability, aeronautical knowledge, simulations and personality assessments before an offer of employment. Starting this process in Aviation school can help future pilots learn the expectations and rigorous demands placed on aviators in addition to refining the abilities and qualifications of the future talent pool. This paper will examine the existing selection system of an Aviation program at a Midwestern University, and will continue data collection from a previous thesis study that spanned Fall 2011- Spring 2012. The goal is to develop a selection program that successfully identifies behaviors and predictors consistent with performance. In addition, the present study intends to identify certain predictors that may signal problem behaviors in the program.

Personality characteristics have been found to differentiate successful pilots from the general population. Personality is defined as "the characteristic way in which a person thinks, feels and behaves; the ingrained pattern of behavior that each person evolves, both consciously and unconsciously, as the style of life or way of being in

adapting to the environment." (Shahrokh, Hales, Phillips & Yudofsky, 2011, p.189). Personality can also be defined as, "the relatively enduring patterns of thoughts, feelings, and behaviors that distinguish individuals from one another" (Roberts & Mroczek, 2008, p.31). Based on the idea that personality is enduring and consistent across situations (McCrae & Costa, 1994), certain personality questions may be asked as a part of a selection battery. Selection tests are designed to identify specific predictors of performance in addition to assessing if an individual's knowledge, skills and abilities are in line with the position for which they are applying. In the current study, students being selected for an Aviation Program should have a base of comprehension, mechanical and spatial skills in addition to certain personality traits that may make some individuals more equipped to handle the tasks and situations that a pilot will encounter.

A 2004 study by the National Aeronautics and Space Administration (NASA) used the NEO-PI-R with 93 pilots to determine if the pilots had certain personality traits in common (Fitzgibbons et al.). The NEO-PI-R assesses the Five Factor Model of personality originally developed by Costa and McCrae (1985), including the dimensions of neuroticism, extraversion, openness to experience, agreeableness and conscientiousness. Neuroticism can be defined as the tendency to experience negative affect (i.e. anxiety, depression), extraversion identifies the amount and degree of interpersonal interactions, openness to experience identifies how proactive an individual is in seeking out new experiences, agreeableness identifies how an individual's interpersonal interactions would fall on a continuum from compassion to hostility and last, conscientiousness identifies the amount of persistence and motivation in terms of goal-oriented behaviors (Piedmont & Weinstein, 1994).

Results found the majority (60%) of pilots scored low or very low on the neuroticism scale with only 13% reporting a high level of neurotic behavior. Forty two percent of the pilots reported high scores on the extraversion scale, with only 23% reporting low scores. Openness to experience had a relatively normal distribution with 29% scoring high and 37% scoring low on the scale. Agreeableness was similar, with 27% of pilots scoring high and 32% scoring low on the scale and last, the majority scored high (58%) on the conscientiousness scale, with only 7.5% scoring low. These results indicate that as a collective, pilots tend to be emotionally stable, outgoing and very motivated and organized when it comes to accomplishing goals. The *personality profile* as developed by Fitzgibbons et al. (2004) describes a pilot as someone who is emotionally stable and low in anxiety, hostility and impulsiveness with high competence and achievement-striving behaviors. A pilot also tends to be trusting and straightforward, with a high level of assertiveness. The authors noted that their results could be used as convergent validity for previous pilot models of personality (i.e. Hormann & Maschke, 1996; Picano, 1991).

In a study of 1,301 U.S. Air Force student pilots, Callister et al. (1999) observed that male pilots tended to have higher levels of extraversion and lower levels of agreeableness when compared with adult male norms and females tended to have higher levels of extraversion and openness to experience and lower levels of agreeableness when compared with adult female norms. Callister also described the average male pilot as achievement oriented, highly competent, responsible and capable of handling high levels of stress. The average female pilot was defined as having similar characteristics of being competitive and tough-minded, but also showing more openness to experience, due to

breaking traditional gender roles in aviation. As a collective, U.S. Air Force pilots tend to score high in extroversion, low in agreeableness and average for neuroticism and conscientiousness (Callister, 1999).

A comprehensive program of research that assessed the structure of male and female personalities in various performance situations identified two core dimensions critical to pilot performance (reviewed by Helmreich, 1986):

1. Instrumental traits relating to achievement and goal seeking
2. Expressive traits relating to interpersonal behaviors, sensitivity and orientation

Both dimensions were found to be important predictors of team performance in aerospace environments. Chidester et al. (1991) noted that superior pilot performers on multiperson crews showed high scores on positive, instrumental traits (i.e. mastery of new and challenging tasks) and low scores on negative instrumental attributes (i.e. arrogance and hostility). High scores on expressive traits were also related to superior performance. This shows that operating an aircraft is a complex process that requires coordination and cooperation of crewmembers.

A study on military pilots in the context of crew coordination identified three different personality profiles through a cluster analysis (Chidester et al., 1991). In this study, flight-crew effectiveness was defined as a product of technical skills, attitudes and personality characteristics. The first profile was the positive instrumental/interpersonal cluster where individuals showed high levels of instrumental and expressive traits. The second profile was the negative instrumental cluster where individuals showed low levels of positive expressive traits. The third profile was the low motivation cluster where

individuals had below-average scores on positive instrumental and expressive traits in addition to elevated levels of verbal aggressiveness. It was noted that pilots in the observed program with the positive instrumental/expressive profile appeared to benefit the most from training. Low motivation pilots appeared to benefit the least and some extreme cases even rejected the attitudes of the program. This research provides support for the idea that clusters of personality traits tend to have positive implications for pilot training performance.

Cooper, White and Lauber (1979) created a detailed review of 10 years of air transport accidents and discovered that accidents rarely resulted from a lack of knowledge or technical skills, but from breakdowns in communication and workload distribution. Chidester et al. (1991) speculated that variations in crew performance may be more reliably predicted by personality characteristics and attitudes regarding appropriate flight-crew behavior than by knowledge or skills. An interesting perspective brought by Rose (2001) observed that pilots as a whole have good social skills and reasoning, can deal with complex information and make decisions while remaining calm. He also observed that while pilots seem to act very quickly, they are actually very slow and methodical to make well-informed decisions in a crisis situation. In an effort to improve communication skills, airlines have been implementing Cockpit Resource Management (CRM) programs that address the “people skills” of flying an aircraft by training pilots, flight attendants, mechanics, dispatchers and anyone else involved in the flight process on communication, decision making and other team-related skills (Baron, 1997).

After addressing the personality characteristics that have been found to create a "pilot profile," an overview of selection testing and history will contribute to understanding the usability and effectiveness of the current measure. Employee selection tests were first developed with pioneers including Walter Dill Scott and Hugo Munsterberg advocating the use of applying psychology to problems in business (Thomas & Scroggins, 2006). The National Research Council was created before World War I to assist in the selection and placement of troops and continued through World War II, where psychologists developed the application of tests for selection, training and performance evaluation (Driskell & Olmstead, 1989). While cognitive testing has been rather widely accepted throughout history, the use of personality tests in selection has been more controversial (Thomas & Scroggins, 2006). Recent research into personality testing has been more positive with psychologists suggesting that the combination of personality and cognitive testing may enhance validity while reducing adverse impact (Ryan, Ployhart, & Freidel, 1998). The development of personality constructs has been around since the 1930s, when Thurstone may have been the first to identify five separate personality components (Thomas & Scroggins, 2006). A large quantity of research supports personality as a predictor of individual cognitive, attitudinal and behavioral variables that can affect employee success in the workplace (Thomas & Scroggins, 2006).

The ability to predict performance is an important tool in selection, especially when the employee is flying an aircraft of people, cargo or equipment. In addition, flight training programs are expensive to both the trainer and trainee and accurate selection would reduce costs incurred from these programs. One of the first selection tests used by the U.S. Navy during World War II was a combination of a Mechanical

Comprehension Test and Kelly's Biographical Inventory, which was referred to as the Flight Aptitude Rating and had a correlation with flight training success of .43 (Bartram & Dale, 1982). The Biographical Inventory concentrated on historical data from each participant and the Mechanical Comprehension Test assessed logic and reasoning. However, researchers argued this was not a true personality test (Ellis & Conrad, 1948) and future personality inventories did not prove to be effective at predicting performance. The Guilford-Martin Personality Inventory had the highest predictive validity and displayed biserial coefficients from .10 to .14 (Bartram & Dale, 1982), however, the use of pass/fail of training as the criterion could have masked differences existing between participants. When using the Eysenck Personality Inventory as a selection test for pilots, the inference from the data was that successful military pilots tend to have distinctly lower neuroticism and higher extraversion from the general population, although there were differences between the two forms of the test that were administered (Bartram & Dale, 1982).

Another study attempting to assess the predictive validity of an automated personality inventory for Air Force pilot selection by Siem (2002), found the use of self-report personality scores did not enhance the predictive validity of a selection system that included operational tests. The self-report measures covered several areas, but hostility, self-confidence and values flexibility were the only scales related to personality. These relationships were not strong enough to indicate predictive validity and Siem suggested that personality characteristics are more predictive of job performance rather than training because of the "honeymoon effect." This describes the time during initial

training when students are on their best behavior and certain personality characteristics that could be predictive are masked until the individuals settle into their positions.

In contrast to the poor results found by past military attempts at pilot selection, Hormann & Maschke (1996) conducted a validation study of a personality questionnaire for the prediction of job success of commercial airline pilots. In the study, 938 pilots applied for employment with a European airline and were assessed with a multidimensional personality questionnaire and a flight simulation. After 274 pilots were hired, they were measured again 3 years later and found that job success could be predicted with 73.8% accuracy by previous scores on the flight simulation and prior flight experience. When the personality questionnaire (Temperament Structure Scales/TSS) was added into the equation, the prediction of job success increased to 79.3%. Overall, the successful pilots tended to score much higher on interpersonal scales (i.e. extraversion, dominance, aggressiveness, empathy) and lower on emotional scales (i.e. emotional instability, aggressiveness, empathy) of the personality questionnaire. In total, 84% of the hired pilots were selected correctly and stayed with the company without any major difficulties (Hormann & Maschke, 1996).

An additional study by Chidester et al. (1991) found concurrent results that indicated high performing commercial pilots tend to have high scores on traits including expressivity and the need for mastery and low scores on hostility and arrogance. While it is noted that some other studies have failed to find confirmatory support for these findings, Chidester argued this may be due to varying methods and samples.

Based on the literature, the current study puts forth the following hypotheses:

H1 High scores on extraversion and conscientiousness and low scores on neuroticism will be positively correlated with performance.

H2 High scores on professionalism will be positively correlated with performance.

H3 High scores on spatial and numerical reasoning will be positively correlated with performance.

Method

Participants

Forty-two pilot students from the Aviation Department of a Midwestern university participated in research during the school year from September 2012 to May 2013. The majority of the participants are male, with an average age of 20 years old. The majority are also native English speakers with a few international students. Ages range from 18 to 23 with previous hours of flight experience ranging from none to over 27. One respondent with 27.5 hours was an outlier, as the average among the other respondents was 2.6 reported previous flight hours.

Measures

The measures in the present study were chosen through background research into previous pilot selection tools and subject matter expert (SME) interviews with individuals in the Aviation Department. The following measures were divided into a two-part pencil and paper assessment. Part I was timed and consisted of a block counting measure and a numerical reasoning measure. Part II was untimed and consisted of the NEO-PI-R Five Factor Scale, the Cockpit Management Attitudes Questionnaire (CMAQ), a Self

Monitoring Scale and an Integrity Scale. Additional demographic, past flying experience, hours completed, type of ground school completed and other past performance information were also collected.

Block Counting and Rotated Blocks. A 20-item Block Counting scale and 12-item Rotated Block scale adapted from Peterson's Military Practice Tests (Wiener, 2005) were used to analyze participant spatial reasoning. Both measures were timed, with the first being 3 minutes and the second being 11 minutes.

Numerical Reasoning. A 22-item Numerical Reasoning scale adapted from a practice test bank (Newton & Bristoll) was used to analyze participant logic ability. The measure was timed for a total of 20 minutes.

IPIP Five Factor Scale. A 50-item scale with items from the International Personality Item Pool (IPIP) was used to measure where applicants fall on the scales of openness to experience, conscientiousness, extraversion, agreeableness and neuroticism. This is based on the original Five Factor Model developed by Costa & McCrae (1985). The ratings are on a 5-point Likert-type scale from *Strongly Disagree* (1) to *Strongly Agree* (5). Sample items from each scale include:

1. I often feel blue. (Neuroticism)
2. I feel comfortable around people. (Extraversion)
3. I have a vivid imagination. (Openness)
4. I don't see things through. (Conscientiousness, reverse scored)
5. I suspect hidden motives in others. (Agreeableness, reverse scores)

CMAQ. The 8-item Cockpit Management Attitudes Questionnaire was developed by Gregorich, Helmreich & Wilhelm (1990) and was used to assess participant

feelings toward other members of a crew. The ratings are on a 5-point Likert-type scale from *Strongly Disagree* (1) to *Strongly Agree* (5). Sample items include:

1. My decision-making ability is as good in emergencies as it is in any other situation.
2. Even when I am tired, I can perform effectively.
3. Good communication is more important for flying than technical skill.
4. Captains should encourage crew members to questions procedures during flight operations.

Self Monitoring Scale. The 25-item Self Monitoring Scale developed by Snyder (1974) was used to determine the degree to which each participant varies their reactions based on a particular situation or group of individuals. The ratings are on a dichotomous *True/False* scale. Sample items include:

1. I find it hard to imitate the behavior of other people.
2. I can only argue for ideas I already believe.
3. I would probably make a good actor.
4. I may deceive people by being friendly when I really dislike them.

Integrity Scale. A 10-item Integrity Scale adapted from a variety of generic integrity-type questions was used to assess individual attitudes toward stealing, cheating, etc. by asking how often the individual feels certain statements could be justified. The ratings are on a 4-point Likert-type scale from *Never* (1) to *Always* (4). Sample items include:

1. Avoiding paying the fare on public transport.
2. Cheating on taxes if you have a chance.

3. Throwing litter in a public place.
4. Driving under the influence of alcohol.

Instructor Ratings. Last, instructor ratings were used to assess performance based on competencies including situational awareness, preparedness, vigilance and decision making. The scale consists of four items, with questions 1 and 2 on a 5-point Likert-type scale from *Never* (1) to *All of the Time* (5) and questions 3 and 4 on a 5-point Likert-type scale from *Poor* (1) to *Excellent* (5). A copy of the performance measure with full rating scales may be found in the appendix.

1. In your experience with [Name], how often does this student show up for lessons on-time?
2. In your experience with [Name], how often does this student have a flight plan prepared?
3. In your experience with [Name], how would you rate this student's ability to make decisions in-flight (i.e. ability to make the appropriate choice for the situation)?
4. In your experience with [Name], how would you rate this student's situational awareness in-flight (i.e. ability to manage multiple tasks and adapt based on changing conditions)?

Procedure

The test was administered in two equal parts during the first three weeks of class. Part I, including spatial and mechanical reasoning was timed, while Part II assessing personality characteristics, teamwork attitudes, self-monitoring behaviors and integrity was untimed. Part I and II were administered on the same day of the week, but Part II

was given one week after Part I. This design was requested by the professors in an attempt to conserve existing class time. Before completing Part I, participants filled out the demographic questionnaire. The test was proctored in two different classrooms with students who had just entered the program and had not taken flight classes with this particular university. The separation was due to having two different class sections of incoming first-year aviation students.

After the initial assessment battery, performance data was collected during the first week in March, which is around the time the majority of students finished their first stages check and spent at least twelve hours flying. The instructor ratings of student performance were administered after the completion of fall semester, around the middle of spring semester.

Results

Preliminary Analyses

From the original 39 responses, a selected number (n=15) were deleted. These responses included individuals who either left the program or provided incomplete data. After excluding these cases, the final number of valid responses collected was (n=24). Reliability was examined for the personality scale items of the IPIP Five Factor Scale and the Cockpit Management Attitudes Questionnaire. Reliability was respectable for neuroticism (α=.79) and agreeableness (α=.73), and very good for conscientiousness (α=.84) and extraversion (α=.89). Reliability was low for openness (α=.64), but past studies have also found the openness scale to be the least reliable (John, Naumann & Soto, 2008) and this was not integral in the final analyses. The Cockpit Management

Attitudes Questionnaire had unacceptable reliability initially (α=.60), but after dropping an item, reliability rose to minimally acceptable (α=.64). Cronbach's alphas for the measures are presented in Table 1. Reliability was not conducted for the self monitoring, integrity, block counting, rotated blocks or numerical reasoning scales, because the scales had to be graded for correct answers (block counting, rotated blocks, numerical reasoning) or certain responses were given weights which led to a composite score for the entire scale (self monitoring, integrity). The procedure for grading certain scales or weighting others was based on the methods past researchers have used with the measures. Complete correlation tables are presented in Tables 2, 3 and 4.

Table 1. *Cronbach's alpha reliability table*

Measure	Reliability
Neuroticism	.79
Extraversion	.89
Agreeableness	.73
Conscientiousness	.84
Openness	.64
CMAQ	.64

Table 2. *Correlation matrix of personality and performance*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Neuro	1															
Extra	-.13	1														
Open	-.34	-.15	1													
Agree	-.41*	.10	.17	1												
Consc	-.28	.45*	-.08	.19	1											
CMAQ	-.19	.01	.26	-.22	.51*	1										
SM	-.03	.53**	.11	.03	.40	.12	1									
Integrity	.29	.18	-.39	-.40	.30	.04	.39	1								
IR	-.06	.09	.02	-.49*	.04	.20	-.06	.38	1							
On Time	.08	.49*	-.01	-.32	-.01	-.07	.16	.16	.12	1						
FPP	.08	.09	-.32	-.18	-.07	-.08	-.12	.21	.31	.51	1					
DM	-.28	.14	-.07	-.26	.35	.57**	-.14	.17	.59**	-.03	.37	1				
SA	-.01	-.02	-.24	-.57**	.21	.32	-.11	.32	.44*	.24	.39	.58**	1			
Exam 1	.17	-.18	.08	-.27	.15	.45*	-.01	-.14	-.12	.32	.18	-.05	.08	1		
Exam 2	.22	.10	.21	.12	.06	-.22	.10	.03	.06	.09	.19	-.20	-.28	.07	1	
Exam 3	.01	-.05	.25	-.21	.15	.54**	.21	-.08	.20	.10	.11	.18	-.08	.71**	.10	1

Notes. N=24 for all variables. * Denotes significance at the p<.05 level and ** denotes significance at the p<.01 level. Neuro=Neuroticism, Extra=Extraversion, Open=Openness, Agree=Agreeableness, Consc=Conscientiousness, CMAQ=Cockpit Management Attitudes Questionnaire, SM=Self Monitoring, IR=Instructor Rating, FPP=Flight Plan Prepared, DM=Decision Making, SA=Situational Awareness.

Table 3. *Correlation matrix of cognitive ability and performance*

Variable	1	2	3	4	5	6	7	8	9	10	11	12
ACT	1											
BC	-.06	1										
RB	-.17	.18	1									
NR	.05	.09	.43*	1								
IR	.23	.12	.14	-.15	1							
On time	-.04	.34	-.20	.05	.12	1						
FPP	.03	.32	-.12	.07	.31	.51*	1					
DM	.04	.09	.27	-.08	.59**	-.03	.37	1				
SA	.19	.45*	.22	.15	.44*	.24	.39	.58**	1			
Exam 1	.14	.36	-.38	-.15	-.12	.32	.18	-.05	.08	1		
Exam 2	.08	-.22	-.41*	.00	.06	.09	.19	-.20	-.28	.07	1	
Exam 3	.15	.12	-.29	-.29	.20	.10	.11	.18	-.08	.71**	.10	1

Notes. N=24 for all variables. * Denotes significance at the p<.05 level and ** denotes significance at the p<.01 level. BC=Block Counting, RB=Rotated Blocks, NR=Numerical Reasoning, IR=Instructor Rating, FPP=Flight Plan Prepared, DM=Decision Making, SA=Situational Awareness.

Table 4. *Correlation matrix of background information and performance*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Age	1																
HrsFlight	-.19	1															
#GS	.02	.20	1														
AddExp	.51*	-.17	.36	1													
Fam/Fri	.03	-.23	-.28	.16	1												
Activity	.29	-.29	.05	.27	.06	1											
DaysAb	-.29	.68*	.18	-.34	-.08	-.22	1										
ClassAtt	.30	-.66**	.77	.35	.07	.21	-1.0**	1									
StageCh1	.68**	-.03	.38	.46*	-.01	.01	-.31	.32	1								
IR	-.12	.42*	.09	-.12	-.17	-.09	.57**	-.57**	.04	1							
On Time	.20	.20	.17	-.03	-.45*	-.02	.10	-.10	-.09	.12	1						
FPP	.24	.07	.04	.11	-.04	-.09	-.10	.12	.20	.31	.51*	1					
DM	.11	.18	-.10	.14	.08	-.07	.17	-.17	.09	.59**	-.03	.37	1				
SA	.22	.19	.04	.30	-.21	.14	.08	-.07	.22	.44*	.24	.39	.58**	1			
Exam 1	-.45*	.20	.11	-.28	-.39	.06	.04	-.04	-.34	-.12	.32	.18	-.05	.08	1		
Exam 2	-.02	-.01	.40	.24	-.07	-.10	-.05	.06	.32	.06	.09	.19	-.20	-.28	.07	1	
Exam 3	-.60**	.16	.06	-.41*	-.21	-.06	.20	-.20	-.42	.20	.10	.11	.18	-.08	.71**	.10	1

Notes. N=24 for all variables. * Denotes significance at the p<.05 level and ** denotes significance at the p<.01 level. HrsFlight=Hours of flight experience prior to starting the program, #GS=Number of ground schools completed prior to entering program, AddExp=Additional life experience related to the field of aviation, Fam/Fri=Family or close friend in the field of aviation, Activity=Involvement activities outside program, DaysAb=Number of days absent from class, ClassAtt=% of class student attended, StageCh1=Hours to pass stage check 1, IR=Instructor Rating, FPP=Flight Plan Prepared, DM=Decision Making, SA=Situational Awareness.

Test of Hypotheses

Pilot Profile

Hypothesis 1 stated students that more closely resembled the "pilot personality profile" would perform better overall than those that did not fit. The pilot personality profile refers to an individual with high levels of conscientiousness, high extraversion and low neuroticism. This was tested with bivariate correlations and was partially supported. There was a significant correlation between extraversion and a component of the performance measure (r=.49, p<.05), indicating students that have higher levels of extraversion tend to consistently show up to their lessons on-time. In this sample, neuroticism and conscientiousness were not related to the performance measure. However, a linear regression observed conscientiousness to have a substantial beta weight when predicting decision making (β=.31, p=.17) and situational awareness (β=.26, p=.26). This effect size indicates there could potentially be a relationship, but the present sample could have been too small to observe a significant relationship.

Professionalism

Hypothesis 2 stated students with high professionalism would perform better overall than those with low professionalism. This analysis was not possible, due to the FAA-required professionalism measure being removed from the standardized lesson forms for each student. However, in place of professionalism, results from the Cockpit Management Attitudes Questionnaire (CMAQ) scale are substituted. The CMAQ was viewed as a comparable substitute because the measure is intended to assess leadership, coordination and communication (Helmreich, 1984). These attitudes contribute to how pilots and members of the flight crew approach various situations and the amount of

professionalism involved could affect various outcomes. There was a significant correlation between the CMAQ and a component of performance (r=.57, p<.01), indicating students with more positive attitudes towards leadership and communication tend to be more effective at decision making in-flight. In addition, the CMAQ was also correlated with conscientiousness (r=.51, p<.05), in-class exam 1 (r=.45, p<.05) and in-class exam 3 (r=.54, p<.01). This hypothesis was tested with bivariate correlations and results suggest it was supported with the measure substitution.

Spatial and Numerical Reasoning

Hypothesis 3 stated spatial and numerical reasoning would be positively correlated with performance. This hypothesis was tested with bivariate correlations and was partially supported. There was a significant correlation between the block counting measure and in-flight situational awareness (r=.45, p<.05). Block counting was also related to students being involved in activities outside of the program (r=.42, p<.05), indicating those with higher scores on the measure tend to be involved in sports or other membership-type communities outside the aviation program. The additional two measures of rotated blocks and numerical reasoning were not significantly related to the measures of performance. However, a linear regression observed the rotated blocks measure to have a substantial beta weight when predicting decision making (β=.36, p=.12) and the instructor rating (β=.26, p=.28). This effect size indicates there could potentially be a relationship with a larger sample.

Additional Relevant Analyses

Exploratory analyses were conducted to provide additional support to the hypotheses. To evaluate the predictors that were most valuable for predicting instructor

ratings, step-wise regression was conducted using all personality variables (five factor, self monitoring, integrity and CMAQ). The results of this analysis indicate that agreeableness is the only significant predictor (β=-.51, p<.05). When predicting whether students consistently arrived to lessons on-time, the significant personality predictors include extraversion (β=.53, p<.01) and agreeableness (β=-.42, p<.05). No personality predictors significantly indicated whether students consistently had their flight plans prepared. A step-wise regression of these predictors on student decision making revealed that the only statistically significant predictor was CMAQ, as noted by the results of hypothesis 2 (β=.57, p<.01). Finally, agreeableness was found to be a significant predictor of student situational awareness in-flight (β=-.58, p<.01).

Discussion

Hypothesis 1 stated students with elevated levels of conscientiousness and extraversion and low levels of neuroticism, or those that more closely resemble the pilot personality profile would have higher performance than those whose personality characteristics do not align. This was partially supported, as extraversion tends to predict students being consistently on-time for their lessons. Perhaps a reason extraversion tends to predict student flight performance is that extroverted behavior is often associated with leadership and assertiveness. In addition, extroverted students being timelier may indicate a tendency toward proactivity. While a significant relationship was not observed with neuroticism and conscientiousness, the effect size of conscientiousness in predicting decision making and situational awareness suggests that with continued data collection, a significant effect may be obtained.

Hypothesis 2 stated students with high levels of professionalism would have higher performance than those with low professionalism. While the original analysis was not possible due to the deletion of the original measure from the FAA-standardized lesson forms, the substitution of the CMAQ was fitting in this setting. The CMAQ intends to assess leadership, coordination and communication attitudes in a team flight situation. In this sample, hypothesis 2 was supported, indicating students with more positive attitudes towards these attitudes tended to perform higher on decision making and in-class exams 1 and 3. An additional relationship between the CMAQ and the personality construct of conscientiousness also suggest those with more positive attitudes toward teamwork and communication tend to be disciplined and have a need for achievement.

Hypothesis 3 stated students with a proficient ability in spatial and numerical reasoning would have higher overall performance compared to those with lower ability in this area. This was partially supported as the block counting measure was the only cognitive ability measure that was able to identify students to have more in-flight situational awareness. While the numerical reasoning and rotated blocks measures did not have significant relationships with performance in this sample, the effect size of rotated blocks in predicting both decision making and the instructor rating suggests that the lack of statistically significant findings is likely related to the limited sample size. Hence, future research should explore this variable in more detail.

The research also indicated some unexpected relationships between a component of the Five Factor Model, agreeableness, and the measures of performance. Agreeableness was found to be an important predictor in the instructor rating, arriving

on-time for lessons, and situational awareness in-flight. Because agreeableness had negative relationships with these three components of performance, it suggests the more agreeable an individual is, the worse their performance is in terms of aviation. A possible explanation for these relationships is that while pilots need to be communicative and team-oriented, they must also be leaders and assert themselves. Highly agreeable individuals tend to be more likely to engage in pro-social and helping behaviors, but also have a tendency to be more dependent.

Overall, results did not fully support the original personality-performance hypotheses in the aviation setting. However, personality characteristics do play a role in understanding student performance. This research is applicable because the sample was selected from the population for which it is intended to generalize. The Aviation Program at a Midwestern University will use this information to identify students who are likely to perform well in the current setting and students who may struggle with particular areas and require additional training. This research has successfully identified personality characteristics that will aid in the identification of those students.

Limitations of this study include the small sample size and the incomplete data in certain areas. The original sample included 39 students but was reduced to 24 after data on several students was incomplete or missing. Some missing data was due to students dropping out of the program and some was due to inconsistencies in the recorded data. Most notably, a measure of professionalism was dropped from the lesson pages, which led to a substitution of the CMAQ measure to test the second hypothesis.

In the next iteration of this selection measure, I recommend including the five factor scale, CMAQ, block counting and rotated blocks measures. These measures

predicted proficient student behavior including, consistently showing up on-time for lessons and above average decision making and situational awareness in-flight. The five factor scale also predicted deficient student behavior including below average situational awareness and a low instructor rating. In future research, it would be beneficial to examine the relationship between assertiveness and flight performance. Proactive personality would also be a complimentary measure to further understand the relationship between extraversion and timeliness and replace the integrity and self monitoring scales. The addition of assertiveness and proactive personality measures may contribute additional variance and shed more light on the individual characteristics that contribute to overall flight performance.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Baron, R. (1997). *The cockpit, the cabin, and social psychology.* Retrieved from http://airlinesafety.com/editorials/CockpitCabinPsychology.htm

Barrick, M. R., & Mount, M. K. (1991). The Big-Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology,* 44, 1-26.

Bartram, D. & Dale, H. C. A. (1982). The Eysenck personality inventory as selection test for military pilots. *Journal of Occupational Psychology*, 55, 287-296.

Callister, J.D., King, R.E., Retzlaff, P.D. & Marsh, R.W. (1999). Revised NEO personality inventory profiles of male and female U.S. Air Force pilots. *Military Medicine*, 164, 12, 885-90.

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

Cooper, G. E., White, M. D., & Lauber, J. K. (1979) Resource management on the flight deck (NASA Conference Publication No. 2120; NTIS No. N80-22083). Moffett Field, CA: NASA-Ames Research Center.

Damos, D. L. (2003). Pilot selection system help predict performance. *Flight Safety Digest,* 22, 2, 1-12.

Driskell, J. E., & Olmstead, B. (1989). Psychology and the military research applications and trends. *American Psychologist,* 44 (1), 43-54.

Ellis, A. & Conrad, H. S. (1948). The vailidity of personality inventories in military practice. *Psychological Bulletin,* 45, 385-426.

Eysenck, H. J., & Eysenck, S. B. G. (1967). On the unitary nature of extraversion. *Acta Psychologica,* 26, 383-390.

Fitzgibbons, A., Davis, D. & Schutte, P.C. (2004). Pilot personality profile using the NEO-PI-R. *NASA Center for Aerospace Information*.

Helmreich RL. Cockpit management attitudes. *Hum Factors* 1984; **26**:583–9.

Helmreich, R. L. (1986). Theory underlying CRM training: Psychological issues in flight crew performance and crew coordination. In H. W. Orlady & H. C. Foushee (Eds.), *Cockpit Resource Management Training: Proceedings of the NASA/MAC Workshop. NASA Conference Publication, 2455.* Moffett Field, CA: NASA-Ames Research Center.

Hormann, H. J. & Maschke, P. (1996). On the relation between personality and job performance of airline pilots. *The International Journal of Aviation Psychology*, 6, 2, 171-178.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (pp. 114-158). New York, NY: Guilford Press.

McCrae, R.R., & Costa, P.T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science,* 3, 173-175.

Newton, P., & Bristoll, H. (n.d.). *Numerical reasoning practice test 1*. Retrieved from http://www.psychometric-success.com/practice-papers/Psychometric Success Numerical Ability - Reasoning Practice Test 1.pdf

Picano, J. (1991). Personality types among experience military pilots. Aviation, Space, and Environmental Medicine, 62(6), 517-520.

Piedmont, R. & Weinstein, H. (1994). Predicting supervisor ratings of job performance using the NEO personality inventory. The Journal of Psychology, 128(3), 255-265.

Roberts, B.W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science,* 17, 31-35.

Rose, R. G. (2001). Practical use of the pilot personality profile. *Aviation Publishing Group*. Retrieved from: http://www.avweb.com/news/aeromed/181606-1.html.

Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology,* 83 (2), 298-307.

Shahrokh, N. C., Hales, R. E., Phillips, K. A., & Yudofsky, S. C. (2011). *The language of mental health: A glossary of psychiatric terms*. (1st ed.). Arlington: American Psychiatric Publishing, Inc.

Siem, F. M. (2002). Predictive validity of an automated personality inventory for Air Force pilot selection. *The International Journal of Aviation Psychology*, 2, 4, 261-270.

Thomas, S. L. & Scroggins, W. A. (2006). Psychological testing in personnel selection: Contemporary issues in cognitive ability and personality testing. *Journal of Business Inquiry*, 5, 28-38.

Wiener, S. (2005). Military flight aptitude tests. In T. Moore (Ed.), (6th ed.). United States: Thomson Peterson. Retrieved from http://www.usarec.army.mil/hq/warrant/download/militaryflight.pdf

Appendix

Aviation Student Performance Survey

1. In your experience with this student, how often does he/she show up for lessons on time?

☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ All of the Time

2. In your experience with this student, how often does she/he have a flight plan prepared?

☐ Never
☐ Rarely
☐ Sometimes
☐ Often
☐ All of the Time

3. In your experience with this student, how would you rate his/her ability to make decisions in-flight (i.e. ability to make the appropriate choice for the situation)?

☐ Poor
☐ Below Average
☐ Average
☐ Above Average
☐ Excellent

4. In your experience with this student, how would you rate his/her situational awareness in-flight (i.e. ability to manage multiple tasks and adapt based on changing conditions)?

☐ Poor
☐ Below Average
☐ Average
☐ Above Average
☐ Excellent

5. Additional comments about this student's performance: