INVESTIGATING PREDICTORS OF PRIMARY FLIGHT TRAINING IN THE CANADIAN FORCES

by

Dawn Herniman

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University
Ottawa, Canada

Your file Votre référence
ISBN: 978-0-494-94338-0

Our file Notre référence
ISBN: 978-0-494-94338-0

Abstract

The present research examined cognitive (general, specific, and executive functions), psychomotor, and personality predictors of Primary Flight Training performance in the Canadian Forces (CF). Measures included current CF pilot selection measures as well as the Royal Air Force Aircrew Aptitude Test (RAFAAT) and ExamCorp battery (ExamCorp, 2012). Results indicate that psychomotor ability and information processing, sub-tests of the RAFAAT, were significant predictors of pilot training performance. Three ExamCorp executive functioning sub-tests measuring attention control and associated learning were also significantly related to pilot training outcomes demonstrating potential utility for use in the CF pilot selection model. Higher levels of conscientiousness and lower levels of extroversion were associated with better training outcomes. To improve the top-down pilot selection process, a weighted regression model based on these results was tentatively proposed. Future longitudinal studies are recommended to overcome sample size limitations and to provide longer-term validation of the findings here.

Acknowledgements

I am grateful to the Canadian Forces for sponsoring my Master's studies, and to both the Director of General Military Personnel Research and Analysis (DGMPRA) and to the Director of Air Personnel Strategy for sponsoring and funding my research.

There are several individuals who contributed to this thesis and provided guidance and support throughout my academic endeavors. First, I would like to acknowledge my supervisors, Dr Avi Parush and Dr Janet Mantler, for their patience, guidance, support, and constructive feedback. For their input and guidance I also extend my appreciation to my committee members, Dr Chris Herdman, Dr Tzvetanka Dovreba-Martinova, and my external reviewer, Dr Damian O'Keefe.

To my parents, Charles and Diane, I cannot adequately express how much I appreciate your continual encouragement, support, love, and belief in me despite my self-doubt. For her research prowess, motivational talent, and unyielding support, I thank my friend and role-model Karen Rankin. Thanks to Karen, Michelaine, and Kim for reminding me of life outside of academia with good food, wine, laughter, and friendship. Thanks, too, to my academic cohorts, namely Jeff Gauger and Ashley McCulloch, for your friendship, support, laughter, and study group fun.

Finally, I would also like to thank the staff at Allied Wings Aviation Training Centre and 3 Canadian Forces Flying Training School in Portage La Prairie, Manitoba, for their encouragement, enthusiasm, and support which contributed to the completion of this research project, including LCol Rob Kamphuis, Bill Ryan, Lisa Klassen, Chrissy Perry, Andrew Kline, and Rose Malenchak. I am also grateful to Colin Kemp, my DGMPRA research contact, for his insights and suggestions.

Table of Contents

List of Tables

List of Figures

List of Appendices

Investigating Predictors of Primary Flight Training in the Canadian Forces

"The quality of the box matters little.

Success depends upon the man who sits in it."

Baron Manfred von Richthofen,

The Red Baron

Military history has shown that a pilot's skill is a critical factor in ensuring air supremacy and maintaining both a tactical and strategic advantage on the ground and in the air (Werrell, 2007). In fact, identifying potential pilots with the necessary skills to succeed in training has been a concern to militaries as early as the First World War (North & Griffin, 1977; Skinner & Ree, 1987). Creating an accurate pilot selection system is also essential to reduce military expenditures; the cost of training one military pilot through to wings qualification (i.e., completion of all pilot training) is estimated between $1 million and $3 million, depending on the aircraft type (i.e., jet, multi-engine, or rotary wing; Lieutenant Colonel J. C. Gallego, personal communication, July 28, 2011; Spangler, 2009; United States General Accounting Office, 1999). At the same time, pilot training has historically been plagued with high attrition rates in training (Hunter & Burke, 1994) and such attrition can result in empty billets in successive training courses and delays in meeting operational demands for qualified pilots, ultimately reducing operational readiness (Lieutenant Colonel J. C. Gallego, personal communication, July 28, 2011; Hunter & Burke, 1994).

Given that training attrition is caused by numerous factors ranging from voluntary withdrawal to ability deficiencies, a solid selection system should reduce training costs, maximize training successes, and mitigate attrition problems (Society for Industrial Organizational Psychology, 2012). Indeed, when training is complex, costly, and dependent on specific abilities, such as is the case with military pilots, "Poor selection will result in increased training attrition, training requirements, and costs, and lead to poor job performance and poor organizational effectiveness" (Carretta & Ree, 2003, p. 359). Therefore, predicting pilot training success early has pragmatic and real world benefits.

The goal of this thesis is to contribute to improving the predictive capacity of early pilot selection, thus reducing training costs and improving operational capabilities. The thesis specifically aims at assessing several pilot selection tests with a particular emphasis on cognitive abilities and their relations to success in early flight training in the Royal Canadian Air Force. This document starts with a brief review of the history of pilot selection to provide the reader with background information concerning the development of pilot selection and the psychological domains that have typically been studied. The historical background leads to reviewing the history and current pilot testing in the Royal Canadian Air Force along with some critique of the current situation. The reviews help identify several themes that are relevant to pilot tasks. These themes, which are the focus of investigation in this thesis, are further reviewed. Finally, the goals, objectives, and hypotheses of the study are presented.

Historical Review of Pilot Selection

The selection of military pilots has been a concern as early as the First World War. According to Dockeray and Isaacs (1921), only Italy and France had investigated pilot selection measures prior to the First World War when Italians tested potential pilots for emotional reaction, reaction time, equilibrium, attention, and perception of muscular effort while the French investigated emotional stability and reaction time (Carretta & Ree, 2003). Conversely, pilot selection in the United States, Canada, and Britain at the beginning of the First World War involved subjective interviews conducted by serving pilots assessing a candidate's fitting the right "type" (i.e., the right school, tall, smart, and sporty; Bailey, 1999). In Canada and England the successful pilot candidate was expected to be educated, young, and strong enough to manoeuvre the aircraft in combat, thus representing the "twentieth century cavalry officer mounted on Pegasus" (Air Ministry notes; as cited in English, 1996). However, this form of pilot selection was not very effective, and intervention was required in order to reduce training attrition, and consequently reduce training costs (Damos, 2007). Whereas Canada and England opted to reduce training costs and training attrition by increasing physical standards, the United States focused on developing additional ability measures (English, 1996).

During the First World War the importance of improving pilot selection measures became apparent as pilot training attrition increased, resulting in accompanying increases in training costs. In fact, in 1915, records show that 90% of aircraft losses were attributable to pilot performance, and of that, pilots who were assessed with physical defects that affected flying performance caused 60% of the losses (English, 1996). Investigation into what made a good pilot revealed the importance of personal qualities, as well as cognitive and psychomotor abilities (Martinussen & Hunter, 2010). Psychometric measures were subsequently developed during this era based on Spearman's (1904) two-factor theory of intelligence that involves a general cognitive factor as well as test-unique factors (Carretta & Ree, 2003). Most notably, in the United States, Thorndike (1919) developed an intelligence test (the Mental Alertness Test) that was added to the pilot selection system in 1918, and paved the road for the psychometric measurement of intelligence (Damos, 2007). In addition to the general cognitive ability measure, the United States Army Air Service pilot selection test battery included measures of reaction time, emotional stability, and sense of equilibrium (Martinussen & Hunter, 2010). Many countries also included a psychomotor ability measure which was also assessed using various apparatus to simulate aspects of a "flying machine" (Martinussen & Hunter, 2010). By the end of the First World War several European countries had developed ability measures similar to the United States in addition to apparatus measures of psychomotor abilities (Martinussen & Hunter, 2010).

Although pilot selection research dropped off significantly between the First and Second World Wars for most countries, during that time Germany developed a comprehensive test battery that included measured of general intelligence, psychomotor ability, perceptual abilities, leadership, and character (Fits, 1946; as cited in Martinussen & Hunter, 2010). Although the United States, Canada, and England commenced the Second World War with only a few pilot selection measures, by the end of the war a number of selection measures had been developed and implemented including psychometric measures of intelligence, psychomotor ability, spatial ability, and mechanical comprehension (Bailey, 1999; Hussey, 2004; Martinussen & Hunter, 2010). In fact, the pilot selection measures employed by Germany were similar to those used by the Allies, whereas the Japanese used tests based on the American Army Alpha battery (Martinussen & Hunter, 2010). Interestingly, in addition to incorporating cognitive measures, Canada was the only Allied nation to employ the Visual Link Trainer in the aircrew-selection process (English, 1996).

Since the Second World War, pilot selection research has investigated personality factors and further explored the cognitive and psychomotor ability domains. In the 1950s, the United States conducted comprehensive research of 26 personality measures to find a suitable personality test to be used in pilot selection (Martinussen & Hunter, 2010). Using long-term pilot performance measures as the criteria, the researchers found that personality tests were better predictors of long-term criteria than were ability measures (Martinussen & Hunter, 2010). Although some countries (e.g., Denmark, Norway, and Sweden) have found predictive validity for personality tests, research on personality measures for pilot selection in other countries (e.g., the United States and England) has been inconclusive and personality tests are not used during the selection process (Martinussen & Hunter, 2010). However, research into personality as a predictor of pilot performance continues to this day as researchers attempt to explain the variance in pilot performance not accounted for by either cognitive or psychomotor ability. In addition to personality, other domains have also seen progress since the Second World War, mainly due to advances in computer technology.

Computer technology has had a significant impact on pilot selection systems (Martinussen & Hunter, 2010). Indeed, advances in computer technology have not only resulted in cheaper and more efficient testing systems, but have also allowed for testing of more complex domains and abilities that was not previously possible (Martinussen & Hunter, 2010). For example, most Air Forces have included computer-based measures of complex abilities such as general cognitive ability (i.e., *g*), attention, reaction time, multi-tasking, and psychomotor ability (BaseOps, 2012; Martinussen & Hunter, 2010; Tsang & Vidulich, 2008). Conversely, few countries rely solely on paper-and-pencil tests (Martinussen & Hunter, 2010). In addition to facilitating multi-aptitude selection test batteries, advances in computer technology have also permitted work sample tests in the form of computer-based moving flight simulators (Tsang & Vidulich, 2008).

Work sample tests are described as artificially created situations in which individuals are required to perform either the same or similar tasks that are required to perform the job (Hunter & Burke, 1995). Long and Varney (1975) initially created a computer-based moving simulator (i.e., work sample test) for the U.S. Air Force called the Automated Pilot Aptitude Measurement System (APAMS). Due to the high cost and difficulty implementing the APAMS in a decentralized selection system, the APAMS was never adopted by the U.S. Air Force. However, the Royal Canadian Air Force, which uses a centralized pilot selection system, developed the Canadian Automated Pilot Selection System (CAPSS) modelled after the APAMS (Okros, Spinner, & James, 1991; Spinner, 1991). Germany also developed and implemented a flight simulator, called the Flight Psychological Selection System (i.e., FPS-80), as part of their pilot selection system (Gnan, Flynn, & King, 1995). While candidates fly specified missions in the FPS-80 they are observed and assessed by psychologists on a number of different cognitive (time sharing, attention, perceptual speed, reaction time, learning speed, precision control), behavioural (stress resistance, self-control, adaptability), and psychomotor abilities (Gnan, Flynn, & King, 1995).

This review of pilot selection research, and the domains historically associated with pilot training performance, has identified some common themes, mainly the relevance of measures of general mental ability, specific cognitive abilities, psychomotor ability, personality, and work sample or previous experience. These specific themes are the focus of this thesis when assessing some of the currently used pilot selection tests in the Canadian Air Force.

The Current Canadian Forces Pilot Selection Model

The current Canadian Forces pilot selection model includes pre-screening measures that are applied to all candidates as well as pilot specific requirements. The pre-screening requirements for Canadian Forces officer applicants includes academic screening (i.e., completed a university degree or enrol in a degree-granting program), initial medical screening, meeting the minimum officer cut-off score on the Canadian Forces Aptitude Test (CFAT; a measure of general mental ability), and having a positive outcome from the structured interview. Candidates who successfully complete the pre-screening requirements then move on to the pilot specific screening, which includes the Canadian Automated Pilot Selection System (CAPSS) simulator, an aircrew medical evaluation. The CAPSS simulates an Instrument Flying Rules (IFR) environment in a light aircraft cockpit and is able to move in three axes (Woycheshin, 1999). Woycheshin (2002) demonstrated that there is a strong positive relation between CAPSS performance and Primary Flight Training performance, and Phase II performance for multi-engine and jet. The Royal Air Force Aircrew Aptitude Test (RAFAAT) battery was added to the pilot selection phase in 2008 to validate the measures with a Canadian Forces pilot population. As of November 2011, only one measure (psychomotor ability) has been included in the pilot selection model in 2011 as the RAFAAT has yet to be validated with a Canadian Forces pilot population. The RAFAAT battery includes measures of domains that have been associated with pilot training performance, including psychomotor ability, critical reasoning (i.e., fluid intelligence), and various cognitive ability measures. Despite the use of CAPSS simulator and RAFAAT battery in the Canadian Forces pilot selection system, some issues remain that must be explored.

Issues with Current Canadian Forces Selection Measures

Despite the longstanding use of the CAPSS simulator in the Canadian Forces pilot selection system, previous studies have identified some issues with its use. For example, studies have shown that CAPSS is positively biased towards candidates with fixed-wing previous flying experience and negatively biased towards both females and candidates with rotary wing previous flying experience (Darr, 2009; Pelchat, 1999; Woycheshin, 2001, 2002). For candidates with fixed-wing previous flying experience, those with more flying time achieved a higher CAPSS score than those with less flying time (Darr, 2009; Johnston & Catano, 2010; Woycheshin, 1999, 2001, 2002). As the CAPSS simulator represents a small engine aircraft with all appropriate instruments, it is logical that people with previous flying experience would have an advantage, given their familiarity with the instruments and knowledge of the sensitivity of the equipment, particularly the lag response to control inputs. However, there is a question of selection fairness of the CAPSS as candidates with rotary wing experience performed as well as candidates with no previous flying experience; rotary-wing experience did not elevate the chance of success on CAPSS (Darr, 2009). Furthermore, previous research has shown that while CAPSS performance was positively related to multi-engine training performance (flying and academic), CAPSS performance was negatively or unrelated to jet and rotary wing training performance criteria (Skomorovsky & Donohue, 2010). Nevertheless, the fact that previous flying experience has such a significant impact on CAPSS performance is not the only issue with this selection measure, as gender differences are also a concern.

The study of gender differences in Canadian Forces pilot selection measures has been difficult due to the limited number of female applicants compared to male applicants (Darr, 2009; Woycheshin, 2001). Regardless, Darr (2009) recommended the investigation of alternate predictors of pilot performance as she identified gender differences in CAPSS performance that indicates the presence of adverse impact. As women are considered a protected group under the Canadian Employment Equity Act of 1995, it behoves the Canadian Forces to investigate alternate non-biased predictors (Catano, Wiesner, Hackett, & Methot, 2005). Unfortunately, due to the restriction on the number of primary flight training courses available during the data collection period of the present study, which restricts the number of pilot trainees, it is not expected that the number of female pilot trainees will be sufficiently large to perform a meaningful analysis.

In addition to the issues of bias and adverse impact, the CAPSS simulator is nearing the end of its life cycle, further prompting interest in investigating alternate predictors of pilot training performance in the Canadian Forces. The CAPSS simulator is based on twenty-year-old computer technology (Spinner, 1988) and maintaining the antiquated computer equipment is difficult due to the high cost of replacement parts and difficulty in acquiring such outdated technology. The combination of these issues (i.e., bias, adverse impact, end of life cycle) prompted the addition of the Royal Air Force Aircrew Aptitude Test (RAFAAT) to the current selection system in 2010.

The Canadian Forces purchased the RAFAAT battery in 2008, basing the decision on previous research that indicated the measures were valid predictors of pilot training performance (Darr, 2010b). However, the RAFAAT validation used a Royal Air Force population, and before it is fully implemented in the Canadian Forces it must be validated with a Canadian Forces pilot applicant population. A preliminary test of the RAFAAT with a Canadian Forces pilot candidate population used CAPSS performance as a proxy of primary flight training performance (Darr, 2010b) and resulted in the inclusion of the RAFAAT battery in the Canadian Forces pilot selection system using a top-down selection strategy (i.e., rank ordering results with preference given to top performers; Darr, 2010b). Validation of the RAFAAT using Canadian Forces pilot training performance as the criteria will permit the calculation of norms for the individual tests, thus permitting consideration of alternate selection decision strategies (e.g., cut-off scores; Society for Industrial Organizational Psychology, 2003). The inclusion of the RAFAAT measures in the Canadian Forces pilot selection system will allow the measurement of other domains associated with pilot training performance in addition to performance on the CAPSS simulator.

The issues outlined above with CAPSS, the requirement for the validation of the RAFAAT, and an interest in investigating additional predictors of pilot training performance stimulated interest in pursuing the current study. Moreover, the current state of testing suggests looking further into pilot job components to determine how they are related to current selection testing.

Pilot Tasks

More recently, pilot selection has been based on analyses of what pilots are expected to do while flying. Such job analyses have identified knowledge, skills, abilities, and other attributes that are required to perform to a high standard and these have been classified into major categories. This section illustrates the relation between pilot tasks and the psychological domains associated with the performance of such tasks.

In order to provide a more complete picture of the job requirements and the demands associated with flying an aircraft, it is necessary to consider specific flying tasks. Wickens (2007) classified tasks into major categories that are also used as a prioritization system pilots commonly refer to as Aviate, Navigate, Communicate, and Systems Management (ANCS). Aviating tasks include maintaining stable flight control and lift, as well as control of the altitude parameters (Wickens, 2007). Navigating tasks require the pilot to navigate through three-dimensional space to reach a specified desired point while avoiding hazards such as bad weather, terrain, obstacles, and other aircraft (Wickens, 2007). Communicating tasks require the pilot to communicate with air traffic control, with other aircraft in the airspace, and with the co-pilot in order to construct an accurate representation of the situation, anticipate issues, and ensure air safety (Morrow & Rodvold, 1998; Morrow, Menard, Ridolfo, Stine-Morrow, Teller, & Bryant, 2003; Wickens, 2007). Pilots must also have knowledge of the aircraft and its associated equipment and systems (e.g., digital, mechanical, electrical, and hydraulic), in order to quickly take action when systems malfunctions are detected (Wickens, 2007). Military pilots also have mission-specific tasks, including threats posed by flying into hostile theatre of action, flying during adverse weather conditions, using austere runways, threat of enemy attack (i.e., small arms fire, ground-to-air missiles, air-to-air combat), and actions on target, which have their own associated concerns (i.e. collateral damage, civilian casualties, friendly fire incidents, etc.).

The tasks involved in performing the ANCS and mission-specific functions place numerous demands on the pilot that are associated with specific psychological domains. Hunter and Burke (1994) conducted a meta-analysis of 68 studies in which they ranked the domains that have typically been associated with pilot performance, based on to their average correlations with pilot training performance. As such, the highest correlation with pilot training performance was job samples (e.g., CAPSS and other simulator-based measures; $r = 0.34$), followed by gross dexterity ($r = 0.32$), reaction time ($r = 0.28$), perceptual speed ($r = 0.20$), general intelligence ($r = 0.13$), verbal ability ($r = 0.12$), personality ($r = 0.10$), and education ($r = 0.06$).

Recent job analyses have also supported inclusion of the domains that have historically been associated with pilot training performance. Byrdorf (1993) conducted a NATO-wide study to determine predictors of fast-jet pilots, and identified nine critical aptitudes: situation awareness (perceptual closure and reaction time combined), spatial orientation, divided and selective attention, time sharing, perceptual speed, psychomotor coordination, visualization, and aggressiveness. The results of the pilot job analysis for commercial and airline pilots posted on O-Net OnLine (2010a; 2010b), as well as the recent pilot job analysis conducted by the Canadian Forces (Darr, 2010a), provided similar information concerning the knowledge, skills, abilities, and other characteristics (KSAOs) required for pilot job performance. The military job analysis (Darr, 2010a) classified the KSAOs into ten major groups including cognitive, psychomotor, stress management, time management, analytical ability, decision-making, communication, people skills, motivation, and self-confidence.

Although these job analyses are informative, it is thought that they are not entirely complete. For example, a task commonly found in pilot research, namely "maintain situational awareness" (Carretta, Perry, & Ree, 1996), was not mentioned in the CF Pilot Job analysis. Therefore, it is suggested that tasks and KSAOs relating to situational awareness, as well as more specific cognitive demands experienced by pilots, may have been overlooked. In the absence of a cognitive task analysis, which would likely reveal more specific cognitive factors associated with pilot performance (Seamster, Redding, and Kaempf, 1997), it is proposed that complex cognitive factors associated with pilot performance may be revealed by measures of executive functioning. Notwithstanding the proposed inclusion of executive functioning measures, a combination job analysis method (CJAM; Levine, 1983) should be performed, in which the results of the functional job analysis are augmented with a more thorough Cognitive Task Analysis (CTA; Wei & Salvendy, 2004) followed by an object-oriented cognitive task analysis and design (OOCTAD; Wei & Salvendy, 2004).

Conceptual Framework

Although there is a rich history of pilot selection measures, it is not clear that such selection has strong associations with the domains and KSAOs required for pilot performance as identified through job analyses. More research is needed to connect pilot selection measures with the known performance domains. Nevertheless, the job analyses have identified many KSAOs applicable to pilot performance, the most important of which were categorized into cognitive, psychomotor, and personality domains (see Appendix A). These domains and their association with pilot training performance were investigated in this study following the conceptual framework presented in Figure 1. The following sections provide background on the various domains and explain their association with job and pilot specific performance. The first domain to be reviewed is the cognitive domain, which includes general mental ability, fluid intelligence, specific cognitive abilities, and executive functioning.

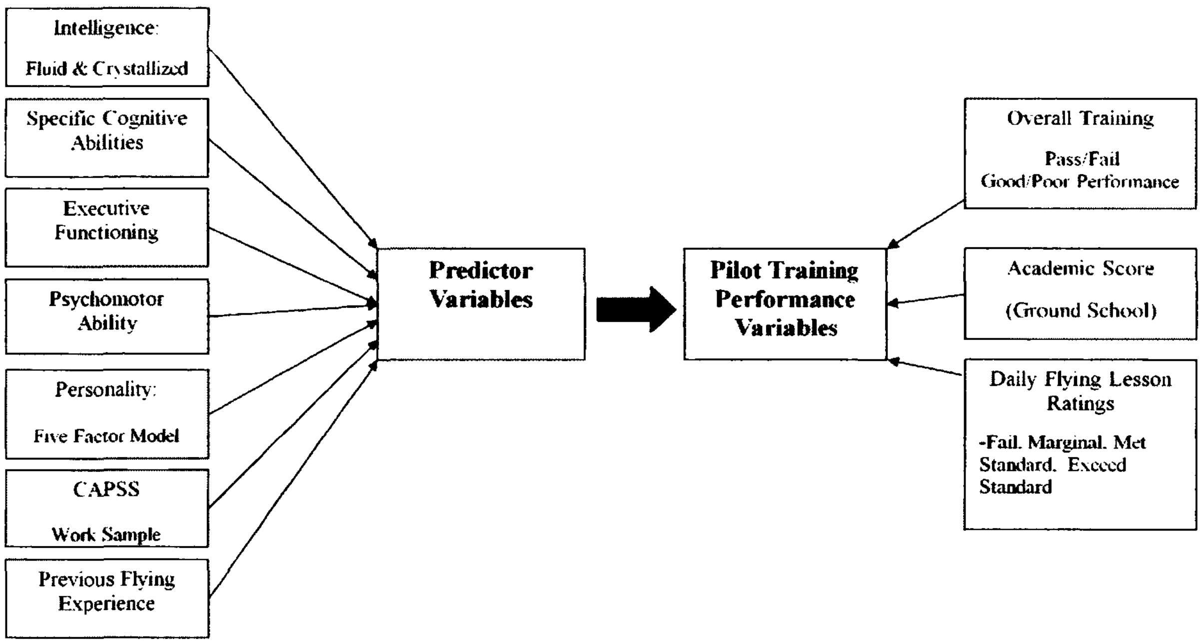


Figure 1. Proposed Canadian Forces Pilot Selection Framework

Domains of Interest in Pilot Selection

Previous Flying Experience. A number of studies have investigated the relation between previous flying experience and performance in Primary Flying Training (i.e., initial flying training in the military), and found that previous flying experience has been shown to predict performance in military primary flight training (Carretta & Ree, 2003; Darr, 2009; Johnston & Catano, 2011; Ree, Carretta, & Teachout, 1995; Woycheshin, 1999). While *g* has been shown to predict Primary Flight Training performance (r = .19), a study by Carretta and Ree (1994) also found a significant relation between previous flying experience and Primary Flight Training (r = .17). Furthermore, the combination of *g* and previous flying experience was deemed to be the best overall predictor of Primary Flight Training performance ($r = .24$; Carretta & Ree, 1994; Johnston & Catano, 2010). This relation has been demonstrated in studies specific to the CF, with correlations between previous flying experience and Primary Flight Training performance ranging from $r = .30$ (Woycheshin, 2002) and $r = .39$ (Deitcher & Johnston, 2004) for academic performance, to even higher correlations for flying performance, ranging from $r = .52$ (Deitcher & Johnston, 2004) to $r = .55$ (Woycheshin, 2002). Additionally, a meta-analysis of 50 studies conducted by Martinussen (1996) found that previous flying experience was the best predictor of pilot performance, followed by cognitive ability and psychomotor ability. Ree, Carretta, and Teachout (1995) found that *g* indirectly influenced work-sample performance through prior job knowledge as well as through job knowledge acquired during training. They also found that the influence of *g* on work samples, such as the CAPSS simulator, was entirely mediated by job knowledge (i.e., previous flying experience). The review of research involving previous flying experience has demonstrated its importance as a predictor of flying training performance.

General and specific cognitive ability. There are several conceptualizations for the structure of intelligence and the cognitive ability domain. Although Spearman (1904) is typically identified as the founder of the psychometric measurement of human intelligence, various theories and models of intelligence have since been proposed and investigated. Today, the most widely researched and accepted theory of intelligence is the Cattell-Horn-Carroll (CHC) integrated model (McGrew, 2005, 2009), which incorporates Carroll's (1993) tri-stratum model of cognitive abilities with the Cattell-Horn Gf-Gc model (Cattell & Horn, 1987). The CHC model, as explained by McGrew

(2005), includes an overall intelligence factor, g, which represents the highest level of the hierarchy, and several second order abilities, including (a) fluid intelligence or reasoning (Gf), (b) crystallized intelligence (Gc; e.g. comprehension-knowledge), (c) cognitive processing speed (Gs), (d) visual-spatial abilities (Gv), (e) auditory processing (Ga), (f) short-term memory (Gsm; includes working memory), (g) long-term storage and retrieval (Glr), and (h) quantitative knowledge (Gq; e.g. math knowledge and achievement). While the CHC model tentatively includes both psychomotor ability (Gp) and psychomotor speed (Gps), for the present study these factors will be addressed and discussed separately under a specific psychomotor domain (McGrew, 2005). The CHC model has a strong empirical basis, as well as a theoretical basis because it was developed through the factor analyses of hundreds of ability studies (Beier & Ackerman, 2007).

General intelligence (*g*) associated with the CHC model represents general cognitive ability that includes both fluid and crystallized intelligence. Fluid intelligence represents reasoning ability independent of prior learning and experience; it involves an individual's ability to think and act abstractly and solve novel problems (Cattell, 1963; Horn, 1976). Crystallized intelligence reflects acquired knowledge, prior learning, and intellectual achievement (Cattell, 1963; Horn & Cattell, 1966). Information processing speed is related to and influences executive functions and cognitive ability, and, because it modulates cognitive efficiency, it represents a reliable measure of general intellectual performance (Causse, Dehais, & Pastor, 2011). In fact, researchers have found that general cognitive ability (*g*) is a significant predictor of both training (Ree & Earles, 1991) and job performance (Hunter, 1986; Hunter & Hunter, 1984; Hunter & Schmidt, 1996; Ree & Earles, 1992; Ree, Earles, & Teachout, 1994). Additionally, for a wide range of jobs, studies have shown that while general intelligence/ability accounts for approximately 50% of common variance, other abilities such as quantitative, spatial, and verbal, each account for roughly 8%-10% of the remaining common variance (Fatolitis, Jentsch, Hancock, Kennedy, & Bowers, 2010).

Cognitive ability and pilot performance. Many of the cognitive abilities covered above have typically been used in military pilot selection models. Numerous studies have identified general cognitive ability (*g*) as being the best predictor of pilot training performance (Carretta, 1992; Carretta & Ree, 1995; Darr, 2010b; Martinussen, 1996; Ree, Carretta, & Teachout, 1995; Yerkes, 1919). Information processing has also been shown to predict pilot training performance (Wickens & Carswell, 2007). Given that pilots work in a high-paced environment that requires quick decision-making abilities, the ability to quickly process information is critical (Ackerman, 1986; Burke, Hobson, & Linsky, 1997; Fitts & Posner, 1967). Studies have also identified visual spatial ability, short-term memory, and quantitative ability as predictors of pilot training performance (Burke, Kokorial, Lescreve, Martin, Van Raay, & Weber, 1995; Carretta & Ree, 1997; Hunter & Burke, 1994; Siem, 1992).

Indeed, this study investigated the newly acquired measure of cognitive ability in the Royal Canadian Air Force, namely the Royal Air Force Aircrew Aptitude Test (RAFAAT) battery. Based on the results of a task analysis, the RAFAAT sub-tests reflect a combination of Carroll's theory of general intelligence and a review of past literature (Taspfield, 2003). As such, RAFAAT sub-tests were created to assess the following five domains: acquired knowledge, reasoning ability, short-term memory, spatial ability, and work rate (i.e., perceptual speed; Southcote, 2007). Although intelligence and specific cognitive abilities are significant components of the cognitive domain that have been shown to predict pilot training performance, another major component of the cognitive domain, namely executive functioning, must also be investigated.

Executive functioning. Executive functions are a significant component of the cognitive domain. Whereas executive functions belong in the cognitive domain, and may overlap with some specific cognitive abilities, executive functions differ in that they coordinate and apply one's cognitive abilities, knowledge, and skills while pursuing a goal (Skoff, 2004). Although there is no agreed-upon operational definition of executive functioning (see Barkley, 2012 for examples), for this study executive functions are defined as higher order or complex cognitive processes that control and regulate the more automatic lower level cognitive processes typically associated with the prefrontal cortex (Alvarez & Emory, 2006; Higgins, Peterson, Lee, & Pihl, 2007; Miller & Wallis, 2009; Peterson, 2010; Skoff, 2004). Baddeley (1986) describes executive functioning of the central executive as the mechanism by which performance is optimized during situations that require the involvement of a number of cognitive processes (Barkley, 2012). Additionally, there are a number of separate theories of executive functioning, which are generally based on studies involving damage to the frontal lobes, or prefrontal cortex (see Barkley, 2012 for examples). Indeed, neuropsychological investigation of the central executive has linked damage to the frontal lobes with disorders of the central executive, also called executive functioning (Baddeley, 1996; Shallice, 1982, 1988). Higgins (2009) offer the following description of the relation between the prefrontal cortex and executive functioning:

> "From its position at the top of the motor system hierarchy, the prefrontal cortex uses incoming information from the sensory systems of the brain (to which it is richly interconnected), analyses this information in the context of the current goals of the organism, determines an appropriate course of action designed to achieve these goals and steers the organism along a path of successful adaptation to the environment (p. 105)."

One model related to executive functioning is Baddeley's (1986, 2000; Baddely & Hitch, 1974) model of working memory, which includes the central executive (i.e., executive functions), two slave systems for short-term storage, namely the phonological loop (i.e. verbal information) and the visuo-spatial sketchpad (i.e., visual and spatial information), and the more recent addition of the episodic buffer. According to Baddeley (2000), the episodic buffer represents the temporary interface between the two slave systems and long term memory, and is controlled by the central executive through conscious thought. Baddeley (1996) describes the central executive as an attention controller, based on the Supervisory Activating System (SAS) component of Norman and Shallice's (1986) attentional control model. To this end, the central executive manages how to focus on task, as well as how to divide attention between two tasks and how to switch attention between tasks.

According to Baddeley (1996), the central executive of the working memory model involves four major components, including 1) dual-task performance (i.e., timeshare between verbal and visuo-spatial), 2) random generation (i.e. constant switching of retrieval plans), 3) selective attention and inhibition, and 4) the activation of long-term memory (Baddeley, 1996). The difference between Baddeley's model and other theories of executive functioning is that Baddeley advocates for a separation between the psychological functions of the executive and the anatomical location of the executive functioning (Baddeley, 1996). Baddeley (1996) also suggests a difference exists between the psychometric approach to intelligence (i.e., separate cognitive abilities), such as Carroll's model on which the RAFAAT was based, and neurological measures of executive functioning, mainly due to the type of material, method of processing, and use of measures of frontal tasks in the case of the neurological studies. Despite the apparent relevance of Baddeley's central executive, one objective of this study was to determine if traditional neuropsychological measures of frontal lobe performance could account for individual differences in pilot training performance. As such, a more thorough examination of the neuropsychological view of executive functioning follows.

Executive functioning has been associated with the adaptation to novel and complex situations and goal-directed behaviour (Causse, Dehais, & Pastor, 2011; Miyake, Friedman, Emerson, Witzki, Howeter, & Wager, 2000; Royall et al., 2002). As such, executive functions "allow the inhibition of automatic responses in favour of controlled and regulated behaviour, in particular when automatic responses are no longer adequate to the new environmental contingencies" (Causse et al., 2011, p. 219). A meta-analysis of 27 studies conducted by Alvarez and Emory (2006) identified that executive functioning is generally assessed through measurement of its components, including working memory, sustained and selective attention (i.e., attention control), and inhibition and switching.

Zelazo, Carter, Reznick, and Frye (1997) suggested these three executive function sub-processes work together to solve complex problems and process complicated decisions. Working memory involves storing, sequencing, updating and monitoring information in working memory, and requires the active manipulation of relevant information in working memory, rather than passively storing information (Miyake et al., 2000; Stuss, Eskes, & Foster, 1994). Working memory is responsible for storing goal-relevant information, including the retention of task relevant information (Miller & Wallis, 2009). Additionally, as reported by Higgins (2009), the dorsolateral prefrontal cortex, through its working memory processes, facilitates the temporal organization of behaviour (Fuster, 1997) and facilitates both the planning and organization of behaviour in the pursuit of goals (Duncan, 1995). Selective attention refers to the ability to shift between multiple tasks, operations, or mental sets (Miyake et al., 2000). Miyake and colleagues (2000) suggest that the shifting ability also includes the ability to engage and disengage appropriate tasks while performing new operations in the face of proactive interference or negative priming. Inhibition refers to the ability to deliberately control and suppress automatic, dominant, and prepotent responses when required (Miyake et al., 2000). Kimberg and Farah (1993) provide an alternate conceptualization of inhibition in which a high level of activation is maintained for a weaker, to-be-selected process, as opposed to the direct suppression of a dominant process.

While research addressing executive functions has typically focused on the neuropsychological approach involving lesion studies and abnormal cognitive functioning (Miyake et al., 2000; Shallice & Burgess, 1991), recent studies have also linked executive functioning with various educational and job outcomes (Higgins, Peterson, Lee, & Pihl, 2007). However, it appears that the relation of executive functioning (associated with the prefrontal cortex) and pilot training performance has not yet been investigated. As flying an aircraft presents one of the greatest cognitive challenges to human capabilities (Wickens, 2007) based on the complexity of the tasks pilots are required to perform, it is expected that executive functioning will play a significant role in the prediction of pilot training performance.

Executive functioning and pilot performance. Although the relation of general cognitive ability as a predictor of pilot training performance has been studied extensively, the influence of executive functioning has not. The majority of the tasks performed by pilots require the pilot to constantly update and maintain his or her spatial and situation awareness and impose heavy demands on a pilot's cognitive processes, including working memory, attention control to deal with multi-tasking interruptions, and decision-making ability. Indeed, a recent study, Causse, Dehais, and Pastor (2011) found that executive functions were fundamental in the majority of tasks required of pilots performance, which they categorized into aviate, navigate, communicate, and systems management tasks. It is expected that the working memory updating, attention control, and inhibition components of executive function are required to perform basic flying tasks associated with the aviate category. Such tasks include handling aircraft flight, monitoring engine parameters, planning navigation, maintaining accurate and current situation awareness, correctly adapting to traffic and environmental changes, and inhibiting wrong behavioural responses in order to perform accurate decision-making such as deciding to abandon or postpone landing during adverse meteorological conditions (Causse et al., 2011). It is also expected that the various components of executive functioning are required to perform the tasks involved in the communicate category. Indeed, previous research has identified a relation between executive functions and radio communication performance; specifically, poor working memory performance was associated with degraded ability to follow Air Traffic Control (ATC) radio communications (Causse et al., 2011; Morrow et al., 2003; Taylor, Noda, O'Hara, Mumenthaler, Kraemer, & Yesavage, 2000). Additionally, the various executive functioning components are also involved in the navigate category of tasks. Previous studies have proposed that the pilot employs working memory and specific spatial abilities when updating the aircraft's location against the ATC communication or the navigation chart route (Endsley & Bolstad, 1994). Heavy cognitive demands are also imposed on pilots' working memory while they are involved in updating the flight model situation (Morrow et al., 2003). Taylor and colleagues (2000) found that a combination of processing speed and working memory predicted errors made by pilots during simulated flight. Wickens (2007) suggested that a pilot's poor executive functioning due to issues with attention control likely explains cognitive tunnelling which occurs when a pilot focuses attention on one specific activity longer than is optimal, thus failing to switch to tasks requiring attention.

In summary, the review thus far has demonstrated that the various constructs involved in the cognitive domain have either demonstrated predictive validity with pilot training performance, or show promise. Both general (i.e., *g*) and specific (i.e., information processing speed, work rate) cognitive abilities have demonstrated predictive validity with job performance, and pilot training performance specifically. Additionally, fluid intelligence and other aspects of cognition, such as executive functioning (i.e., working memory, attention control, and inhibition/switching), also show promise with predicting pilot performance. Thus, cognitive abilities are critical to the fundamental flying tasks. However, these tasks (i.e., aviate, communicate, and navigate) also include a significant psychomotor component. The following section addresses this component and its relation to flight performance.

Psychomotor Ability. Psychomotor ability (PA) is defined as "capabilities of the motor system to plan, coordinate and execute movements" (Ghez & Krakauer, 2000, p. 653). This definition can be expanded (O-Net, 2011) to include gross motor coordination (multi-limb coordination, manual dexterity), fine motor coordination (finger dexterity), and speed of movement/response (reaction time, speed of limb movement, control precision, rate control). Carretta and Ree (1997) suggested a hierarchical model of PA with a higher order general psychomotor factor, *p* (like *g*), which underlies all psychomotor ability tests. However, researchers have also noted that although general mental ability and psychomotor ability are somewhat correlated, the relevance of psychomotor ability decreases with more complex and cognitively demanding jobs (Carretta & Ree, 2000; Hunter & Hunter, 1984; Levine, Spector, Menon, & Narayanan, 1996). The correlation between psychomotor ability and intelligence likely results from the employment of reasoning and learning abilities required to complete the tests (Carretta & Ree, 1997).

Psychomotor ability and pilot performance. Psychomotor abilities (hand-eye or hand-eye-foot coordination) are required to smoothly control the aircraft in three-dimensional space, while simultaneously scanning instruments and airspace for potential issues and hazards (Wickens, 2007). Investigation of psychomotor ability as a predictor of performance has long been associated with pilot training performance. Indeed, psychomotor ability was included in pilot selection as early as the First World War (Carretta & Ree, 1997; Wheeler & Ree, 1997). Numerous studies have continued to demonstrate the importance of psychomotor ability as a predictor of pilot training performance and its usefulness in pilot selection, as psychomotor ability explains incremental validity in performance over measures of cognitive ability (Carretta, 1989; Carretta & Ree, 2000; Griffin & Koonce, 1996; Hunter & Burke, 1994; Johnston & Catano, 2010; Ree & Carretta, 1994). Recent advancement in computer-testing have further advanced the measurement of psychomotor ability (Martinussen & Hunter, 2010). Martinussen's (1996) meta-analysis of 50 studies found that psychomotor ability was a good predictor of training pass-fail outcomes. In fact, Burke and colleagues (1995) recommended that a NATO computer-based pilot assessment battery should include the domain of psychomotor ability, as well as spatial ability, attention, and information processing. Additionally, pilot job analyses support the inclusion of psychomotor ability as these studies have generally identified the necessity of hand-eye-coordination for pilot performance (Damos, 2007; Darr, 2010a; O-Net, 2011). It would be expected, therefore, that in assessing the measures of pilot selection for the Canadian Forces, that the measures of psychomotor ability would explain a significant proportion of successful pilot performance.

This section has presented a review of the literature concerning psychomotor ability and demonstrated the importance of psychomotor ability in the prediction of pilot training performance. Although it is expected that psychomotor ability, in addition to general and specific measures of cognitive and executive functioning are important predictors of pilot training performance, it is clear that these constructs do not fully explain good pilot performance. As such, this suggests that additional factors need to be considered when selecting potential pilots. Thus, the next section will explore personality and investigate whether some personality traits also predict pilot training performance.

Personality. Personality, which is "individual differences in characteristic patterns of thinking, feeling and behaving" (APA, 2011), can be classified into two areas, 1) the study of individual differences in specific personality characteristics, and 2) the investigation of how the various parts of a person come together as a whole. Commencing with Cattell's (1943) early work, the study of personality found footing with the investigation of language, and eventually resulted in Cattell's factor analysis that identified 16 primary personality factors and eight second-order factors (Cattell, Eber, & Tatsuoka, 1970). In subsequent re-analysis of Cattell's early work, Fiske (1949) and Tupes and Christal (1961) were unable to identify more than five personality factors. Further research (Borgatta, 1964; Norman, 1967; Smith, 1967; Wiggins, 1968) eventually solidified the Five Factor Model (FFM) of personality (Digman, 1990). Other researchers also identified different structures of personality, basing their theories on two factors (Wiggins, 1968), three factors (Eysenk, 1991), six factors (Hogan, 1986), seven factors (Jackson, 1984), eight factors (Hough, 1998a, b) or nine factors (Hough, 1992). However, the use of the Five Factor Model has dominated research investigating the relation between personality and job performance (Rothstein & Goffin, 2006). Indeed, Barrick and Mount (2005) reported that the Five Factor Model of personality was used to investigate the personality-job performance relation in all sixteen meta-analytic studies published since 1990 (Rothstein & Goffin, 2006).

The Five-Factor Model (FFM), also known as the "Big Five" theory of personality, is the most widely used model of personality in contemporary research (Steel, Schmidt, & Shultz, 2008). The Big Five categorizes personality traits into five broad dimensions, namely Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, each of which also consists of lower, underlying facets (Costa, 1996). Openness to experience (McCrae & Costa, 1985), also known as Intellect (Borgatta, 1964; Hogan, 1983) or Culture (Norman, 1963), includes being imaginative, adventurous, curious, intelligent, and have an appreciation for art (Barrick & Mount, 1991; Digman, 1990). Conscientiousness includes the characteristics of dependability, achievement-orientation, perseverance, and reflects self-discipline, organization, responsibility, and thoroughness (Barrick & Mount, 1991; Borgatta, 1964; Costa & McCrae, 1988; McCrae & Costa, 1985; Norman, 1967). The extraversion trait features attributes such as assertiveness, initiative, surgency, energy, ambition, positive emotionality, and sociability (i.e., good interpersonal skills; Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Costa & McCrae, 1992; Norman, 1963). The trait of agreeableness includes being courteous, cooperative, good natured, and tolerant (Barrick & Mount, 1991; Borgatta, 1964; McCrae & Costa, 1985; Norman, 1963). Finally, the neuroticism factor, also labelled emotional stability, includes characteristics of anxiety, depression, anger, and insecurity (Barrick & Mount, 1991; Borgatta, 1964; McCrae & Costa, 1985; Norman, 1963).

Various studies have shown the association of personality, based on the Five Factor Model, and job performance. Two major meta-analyses, by Barrick & Mount (1991) and Tett, Jackson, and Rothstein (1991), are generally considered to be the foundation of the research concerning the use of personality to predict job performance (i.e., work-related behaviour; Rothstein & Goffin, 2006). Barrick and Mount (1991) found that each of the Big Five could predict at least one aspect of job performance with some degree of accuracy, for different occupations. For example, they found that

Conscientiousness was associated with all performance criteria (job proficiency, training proficiency, and personnel) across all jobs; extraversion predicted all performance criteria for managers and sales occupations, which involved social interaction; openness to experience and extraversion predicted training proficiency across all occupations. Additionally, Catano, Cronshaw, Wiesner, Hackett, and Methot, (1997) suggest that several different aspects of job or training performance could be moderately predicted by a measure of conscientiousness alone. As such, there is considerable evidence that personality measures are useful in personnel selection systems to predict job performance (Barrick, Mount, & Judge, 2001; Tett, Jackson, & Rothstein, 1991; Tett, Jackson, Rothstein, & Reddon, 1994). More specifically, these studies generally supported the use of emotional stability (i.e., neuroticism) and conscientiousness as valid predictors of job performance in all occupations. In fact, research has shown that the combination of general intelligence and conscientiousness is the most parsimonious predictor of job performance (Anderson, Born, & Cunningham-Snell, 2002). Although studies show less support for the use of agreeableness, openness to experience, and extraversion in predicting work-related variables, there appears to be a relation between extraversion, openness and training proficiency, and between agreeableness and predicting teamwork (Barrick, Mount, & Judge, 2001).

Although these meta-analyses generally show smaller correlations for personality characteristics as predictors of job performance (Rothstein & Goffin, 2006), Cascio (1991) underscores the utility of personality measures in personnel selection processes. Most notably, researchers have found personality adds incremental validity to other selection measures for the prediction of job performance (Goffin, Rothstein, & Johnson, 1996; Schmidt & Hunter, 1998). In their meta-analysis, Tett and colleagues (1991) found much higher validity coefficients (almost double) for validation studies in which a confirmatory research strategy was employed over exploratory studies, and when job analysis was used. According to Rothstein and Jelly (2003), linking personality using job analysis is critical as personality measures are situation specific and subject to situational moderators. This will be particularly important in pilot selection research.

Personality and pilot training performance. Since the early days of military pilot selection, there has been interest in the personality characteristics that define good pilots, with movie stereotypes fostering an idealized pilot who possesses "the right stuff" (Campbell, Castaneda, & Pulos, 2009). In an early attempt to identify the personality characteristics associated with pilot performance, Dockeray and Isaacs (1921) used participant observation to determine that the best pilots were calm and methodical, could quickly adapt to new situations, and showed good judgement. Although this study was based more on subjective assessment, since then various studies (Street, Helton, & Dolgin, 1992) and meta-analyses (Campbell et al., 2009; Martinussen, 1996) reported that later studies incorporated more robust personality assessments. Some personality measures that have been used in pilot selection research include NEO PI-R (Costa & McCrae, 1992), Cattell's Sixteen Personality Factor Questionnaire (16PF; Cattell, Eber, & Tatsuoka, 1970), Norway's Defense Mechanism Test (Kragh, 1960), the Minnesota Multiphasic Personality Inventory (Framingham, 2011), and the Eysenk Personality Inventory (Eysenk & Eysenk, 1964). However, these studies only found small effect sizes for the relation between personality characteristics, in which an aggregate personality scale was calculated, and flying training criteria (i.e., pass/fail r_{mean} = .14;

flying ratings $r_{mean} = .11$; Campbell, Castaneda, & Pulos, 2009; Hunter & Burke, 1994; Martinussen, 1996).

Although the validities are typically low for personality measures (i.e., around $r = .14$; Martinussen, 1996), it is believed that personality characteristics will add incremental validity to the CF pilot selection system (Darr, 2009). Indeed, the CF pilot job analysis (Darr, 2010a) identified several areas related to personality, including stress management (i.e., mental stability, handle baggage interference, handle complexity under pressure), people skills (i.e., ability to work with others, share information, work/learn with others, and interpersonal skills), motivation (motivation to succeed, perseverance towards career goals, accepting criticism, and handling setbacks), and self-confidence (i.e., self confident, but not overconfident). For the stress management area identified in the job analysis, it would be expected that the neuroticism trait should predict pilot training performance; worrisome, stress-prone, and unrealistic individuals would be expected to perform poorly in high-stress military pilot training programs (Campbell, Castaneda, & Pulos, 2009). For the people skills aspect identified by the job analysis, it is expected that the extraversion trait would predict training performance; various studies have linked extraversion to performance in jobs that require significant interaction with others (Barrick & Mount, 1991; Mount, Barrick, & Stewart, 1998). Interaction with others is particularly important in a military aviation training environment that requires teamwork, camaraderie, and significant interaction with others (Campbell, Castaneda, & Pulos, 2009). In this respect, the cooperative aspect of agreeableness is also expected to predict pilot training performance as it is related to people skills. Robins, Tracy, Trzesniewski, and Potter (2001) found that individuals with high self-esteem were emotionally stable (i.e., inverse of neurotic), conscientious, extraverted, and somewhat agreeable and open to experience. Therefore, it is expected that all the Big Five personality traits will be positively related to higher levels of training performance, with neuroticism having a negative relation.

This section reviewed research on personality, and more specifically, how personality characteristics relate to flying training performance. This section on personality adds to the sections on the cognitive and psychomotor influences of pilot training performance. The next section will summarize all the domains applicable to this study, including previous flying experience.

Summary

In summary, this review introduced general and specific cognitive and executive functioning, psychomotor ability, personality, and previous flying experience as predictors of pilot training performance. Although a great deal of research for pilot selection has examined psychomotor ability and cognitive ability, the future of this line of cognitive research extends to executive functioning (Causse, Dehais, & Pastor, 2011). Additionally, previous research shows promise for including a measure of personality into pilot selection systems. Experience has also been examined due to the influence of previous flying experience on pilot training performance outcomes and current selection measures such as the Canadian Automated Pilot Selection System (CAPSS). Despite research that has demonstrated the association of these domains to pilot training performance, not all these domains have been considered in the Canadian Forces pilot selection system.

It is evident that the current Canadian Forces pilot selection model may not be as effective as it could be with the inclusion of all the domains reviewed in this paper. Although the current CF selection system includes validated measures of general cognitive ability and a work sample, recently added measures of specific cognitive ability, reasoning, and psychomotor ability have not yet been validated with a Canadian Forces pilot training sample. Further, the current system does not include measures of executive functioning or personality, which have shown promise as predictors of pilot selection. As indicated earlier, the conceptual framework (Figure 1) on which this paper is based involves the association of the identified domains to various pilot training performance criteria, which are examined in greater detail in the following sections.

The Current Study – Purpose, Objectives, and goals

Given the importance of ensuring a proper pilot selection system to reducing training costs and optimizing operational effectiveness, it is essential to ensure the proper domains and associated measures are incorporated into the selection tools. The purpose of the present study was to assess whether individual differences in cognitive ability, general intelligence, fluid intelligence, executive functioning, psychomotor ability, personality, and a work sample, as measured by various military and commercially available tests, are significant predictors of pilot training performance. Such tests included measures currently used in the Canadian Forces pilot selection system, including the Canadian Forces Aptitude Test (CFAT), the Canadian Automated Pilot Selection System (CAPSS, see Figure 2 in Appendix B), and the Royal Air Force Aircrew Aptitude Test battery (RAFAAT; see Figures 3 and 4 in Appendix C). Additionally, a commercial test battery, ExamCorp, was included as a measure of both executive functioning and personality. Within the purpose of this study, the objective was to link each individual domain (i.e., specific cognitive ability, general mental ability, executive functioning, psychomotor ability, and personality factors) to the primary flight training performance outcomes. This study also investigated whether the addition of the RAFAAT and ExamCorp measures added incremental validity[1] over and above the current measures (i.e., CFAT, CAPSS), to determine if such measures should be added to the Canadian Forces pilot selection system.

Hypotheses

The hypotheses tested in this study explore the associations of the cognitive, psychomotor, and personality domains, via their associated measures (including a work sample), to pilot training performance outcomes.

Previous Flying Experience

Previous flying experience in many cases includes civilian ground school instruction and basic visual flying training, including instruction on aircraft instruments and controls. As Primary Flight Training in the Canadian Forces teaches similar skills, it is expected that previous flying experience will be strongly associated with Primary Flight Training. Indeed, numerous studies have found that previous flying experience is a good predictor of performance on initial flying training (i.e., Primary Flight Training; Carretta & Ree, 2003; Darr, 2009; Johnston & Catano, 2011; Ree, Carretta, & Teachout, 1995; Woycheshin, 1999). Thus, hypothesis 1 is presented:

H1: It was hypothesized that pilot trainees who had completed pervious flying experience would perform better on Primary Flight Training than pilots trainees

[1] Incremental validity is defined as "the degree to which a measure explains or predicts a phenomenon of interest, relative to other measures" (Haynes & Lench, 2003, p. 456), or the unique variance of the measure with the criterion, controlling for all other measures (i.e., partial correlations; Bracket & Mayer, 2003).

who had no previous flying experience.

Cognitive Domain

Three components of the cognitive domain have demonstrated either predictive validity or promise of predictive validity of pilot performance, which for the purposes of this study will be performance on Primary Flight Training (PFT). Specifically, the hypotheses related to cognitive ability and pilot performance are:

General intelligence (*g*). Research has repeatedly demonstrated a strong positive relation between General Mental Ability (GMA or *g*) and pilot training performance (Carretta, 1992; Carretta & Ree, 1995; Darr, 2010b; Martinussen, 1996; Ree, Carretta, & Teachout, 1995; Yerkes, 1919). Indeed, previous research has identified an indirect link between *g* and the acquisition of flying skills, and a direct link between *g* and the acquisition of job knowledge (Carretta & Ree, 1997; Johnston & Catano, 2010; Skomorovsky & Donohue, 2010). Based on this research, I predicted:

> H2: Pilot trainees who had higher scores of general intelligence (*g*; as measured by the CFAT and RAFAAT Critical Reasoning Battery) would have significantly higher levels of pilot training performance (as assessed through academic average, daily flying lesson rating, check ride average) than those with lower levels of general intelligence.

Specific cognitive abilities. Given the fact that military pilots fly in a fast-paced environment that requires quick reaction to multiple inputs as well as speedy decision-making abilities, it is expected that measures of specific cognitive abilities may provide additional predictive power over *g*. To this end, previous research has demonstrated significant relations between specific cognitive abilities, such as information processing, attention, and short-term memory, and pilot training performance (Burke, Kokorial, Lescreve, Martin, Van Raay, & Weber, 1995; Carretta & Ree, 1997; Hunter & Burke, 1994; Siem, 1992; Wickens & Carswell, 2007). In light of this, I predicted:

H3: Pilot trainees who had higher scores of work rate (i.e., information processing speed) and digit recall as measured by RAFAAT would have significantly higher levels of pilot training performance than those with lower levels of work rate.

Executive functioning. Flying tasks pose heavy cognitive challenges to pilots by requiring them to continually update and maintain their spatial and situation awareness, and include additional demands to working memory and attention control to deal with multi-tasking interruptions and decision-making ability. Indeed, the aviate, navigate, communicate, and systems management tasks associated with flying involve executive functioning abilities such as working memory updating, attention control, and inhibition components. Research has demonstrated the potential for measures of executive functions to predict pilot performance (Causse, Dehais, & Pastor, 2011). Therefore, I predicted:

H4: Pilot trainees who had higher levels of executive functioning, such as working memory, attention control, associated learning, and verbal organization (as measured by ExamCorp) would have higher levels of pilot training performance than pilot trainees with lower levels of executive functioning.

Given the high degree of cognitive demands placed on pilots, it is hypothesized that executive functioning will predict more variance in pilot performance than general intelligence and the specific cognitive abilities investigated earlier in this study.

Therefore:

H5: Executive functions (as measured with ExamCorp) would account for more unique variance in pilot performance than general intelligence (*g*; as measured by the CFAT and RAFAAT Critical Reasoning Battery), and specific cognitive abilities (as measured by RAFAAT).

Psychomotor Ability

Psychomotor ability is required to smoothly pilot an aircraft within the three-dimensional environment. Psychomotor ability has been included in pilot selection systems since the First World War, and has demonstrated a strong positive relation with pilot training outcomes over and above cognitive ability (Carretta, 1989; Carretta & Ree, 1997, 2000; Griffin & Koonce, 1996; Hunter & Burke, 1994; Johnston & Catano, 2010; Martinussen, 1996; Ree & Carretta, 1994; Wheeler & Ree, 1997). Major pilot job analyses (Damos, 2007; Darr, 2010a; O-Net, 2011) support the inclusion of psychomotor ability measures to identify necessary multi-limb coordination, a domain that is not specifically assessed in the current Canadian Forces pilot selection model. As such, I predicted:

H6: Pilot trainees with higher levels of psychomotor ability (as measured by RAFAAT) would have higher levels of pilot training performance than those with lower levels of psychomotor ability.

H7: Psychomotor ability (as measured by RAFAAT) would account for unique variance in pilot training performance over and above CAPSS performance.

Personality

Determining the typical personality characteristics of successful pilots that define "the right stuff" has been a concern since the early years of military flying (Campbell, Castaneda, & Pulos, 2009; Dockeray & Isaacs, 1921). Although previous studies have typically found small effect sizes for personality predictors of pilot training and job performance (Campbell, Castaneda, & Pulos, 2009; Hunter & Burke, 1994; Martinussen, 1996), a recent Canadian Forces pilot job analysis (Darr, 2009) identified several key competencies related to personality. To this end, Big Five personality factors have shown promise in predicting both pilot training performance (Barrick & Mount, 1991; Hogan & Ones, 1997; Mount & Barrick, 1995, Salgado, 1997; Tett, Jackson, & Rothstein, 1991) and aviator-specific occupations (Carretta, Rodgers, & Hansen, 1996; Martinussen, 1996; Siem & Murray, 1994). As such, Hypotheses 8 is presented:

H8: It was hypothesized that pilot trainees with higher levels of openness, conscientiousness, extraversion, agreeableness, and emotional stability (the Big Five factors as measured by ExamCorp) would have higher levels of pilot training performance than those with lower levels of these traits.

Integrative Hypothesis

Research has shown that cognitive, psychomotor, and personality domains predict pilot training performance. As the purpose of this study is to improve the Canadian Forces pilot selection system, and determine the gain in validity from the addition of the new domains over the current measures, based on the previous research hypothesis 9 is presented:

H9: The addition of the new domains of fluid intelligence (G_f, measured by the RAFAAT Critical Reasoning Battery), specific cognitive abilities, executive functioning, psychomotor ability, and personality (as measured by the RAFAAT and ExamCorp batteries) would account for incremental validity in the prediction of Canadian Forces pilot training performance over and above the current measures (i.e., CFAT and CAPSS).

Methods

Participants

Participants were Canadian Forces pilot students attending the Primary Flight Training course conducted at the Canadian Forces Flying Training School[2], in Portage La Prairie, Manitoba. Data were collected from eight Primary Flight Training courses between 20 July 2011 and 10 March 2012, each of which accommodated approximately 10-12 students.

Sample size. The original sample ($N = 78$) was reduced because three participants were administratively removed from Primary Flight Training ($N = 75$). Timing issues[3] for the first two administrations resulted in a 30% response rate for those sessions; however, the addition of a test-specific day to the training syllabus resulted in a 90% to 100% response rate for the remaining six administrations. The issues with the early administrations noted previously (i.e., some participants chose to complete either the RAFAAT or ExamCorp, not both), coupled with computer network glitches during the RAFAAT administration (i.e., data were lost for three participants), resulted in unequal sample sizes for RAFAAT ($N = 70$), ExamCorp ($N = 73$), and missing both RAFAAT and ExamCorp (N = 68). Therefore, the sample sizes varied depending on the

---

[2] The official name of the school is 3 Canadian Forces Flying Training School.
[3] As discussed in procedures, the first two administrations were conducted during the flying phase.

analysis being conducted.

Gender and official language. The majority of participants were male (*n* = 69, 92%,) and Anglophone (*n* = 65; 87%). Although there were few women in the sample, this is representative of the actual distribution of female pilots in the Canadian Forces (i.e., 5%; R. Godlewski, personal communication, May 16, 2012). Therefore, no comparisons were made by gender because such comparisons would likely misrepresent the true relation between variables (Tabachnick & Fidell, 2007).

Age. For age, $M = 25.43$, $SD = 4.44$. Visual examination of the continuous age variable revealed that age is positively skewed (skewness z-score = 5.01) for this sample, with two outliers attributed to participants older than 38 years of age (ages 39 and 41). Due to the deviations to normality, it was decided to transform age into a categorical variable; the majority of the participants were 20 to 25 years of age (*n* = 43, 57%), whereas 28% of participants were 26 to 30 years of age (*n* = 21), and the remaining 15% of participants were 31 years of age or older (*n* = 11).

Education. Table 1 shows that most participants had completed an undergraduate degree (*n* = 60, 80%).

Table 1
Highest Level of Education Completed (n = 75).

Type of Diploma/Degree	n	%
High School	8	11
College or part of undergraduate degree	6	8.0
Undergraduate degree	60	80.0
Graduate	1	1

Previous flying experience. The majority of participants (n = 59; 79%) had previous flying experience. Further, of those participants who had previous flying experience, most (*n* = 56, 95%) had experience on fixed wing aircraft, with only 1 participant (2%) having only rotary wing experience and 2 participants having only glider experience (3%). The number of hours of previous flying experience ranged from 0 to 400 hours (M = 46.94, SD = 72.32). To mitigate the effects of skewness, the raw scores for this variable were adjusted to be within 2 standard deviations of the mean, resulting in an adjustment of the range of previous flying experience to be from 0 to 180 hours (M = 40.36, SD = 45.97).

Video game experience. Table 2 presents the frequency, per week, of participants' video game experience by genre. First Person Shooter games were played more frequently and for longer periods of time than any other genre of video game. Moreover, participants played video games an average of M = 6.19 hours per week. Due to the unequal sample size between males (n = 69) and females (n = 6), the investigation of video game experience as a moderator of the relation between gender and performance on CAPSS and the psychomotor ability measure were not pursued.

Table 2

Frequency and Percent of Video Game Hours Played Per Week by Genre (n = 75)

	Hours of Video Play per Week					
	0	Less than 1	1-3	3-5	5-7	Greater than 7
First Person Shooter	14	24	14	9	7	7
Role Playing Game	44	12	5	7	3	4
Puzzle	27	30	11	7	0	0
Strategy	29	27	4	9	3	3
Flight Simulator	23	31	16	1	1	3

Measures

Predictor variables. Assessment instruments are presented based on the following selection constructs:

Intelligence (g). Intelligence was assessed using both the CF Aptitude Test (CFAT) and the Critical Reasoning Battery (CRB) of the Royal Air Force Aircrew Aptitude Test (RAFAAT). AFAAT Critical Reasoning Battery. These two test batteries include items that measure both crystallized and fluid intelligence.

Canadian Force Aptitude Test (CFAT). The CFAT is a 60 item timed measure of general cognitive ability including subtests of verbal ability (15 items), spatial ability (15 items), and problem-solving ability (30 items; (Boswell & Kushnereit, 2009; Ibel & Cotton, 1994). Previous research (Carter, Boswell, Kushnereit, & Ebel-Lam, in press; Donohue, 2005) indicates adequate internal consistency (Cronbach's alpha) for the CFAT and its associated sub-tests, where Cronbach's alpha for Verbal Skills = 0.80, Spatial Ability = 0.70, and Problem Solving and total CFAT = 0.90.

Critical Reasoning Battery (CRB). The RAFAAT Critical Reasoning Battery includes three sub-tests measure verbal reasoning (16 items), numerical reasoning (16 items), and diagrammatic reasoning (16 items; Southcote, 2004). Additionally, the CRB diagrammatic reasoning measure represents a measure of fluid reasoning that is not captured in the CFAT. As item-level measures were not available for all sub-tests, more in-depth psychometric examinations (i.e., Cronbach's alpha) were not able to be performed using this data set. However, previous research (Southcote, 2004) indicates adequate reliability and validity for the RAFAAT critical reasoning measures.

Specific cognitive abilities. The Royal Air Force Aircrew Aptitude Test (RAFAAT; Southcote, 2004) includes measures of work rate (table reading, 86 items; visual search, 74 items), vigilance (attention control; scored using an error score), and short-term memory (digit recall, 15 items). These measures have typically shown low (i.e., $r = 0.07$ to $r = .0.18$) predictive validity with Royal Air Force pilot samples (Southcote, 2004). Previous research indicates adequate internal consistency for Vigilance (Cronbach's alpha = 0.91) and Table Reading (MATF; split half internal $r = .88$; NATO Training Group, 2008). As item-level measures were not available for all sub-tests, more in-depth psychometric examinations (i.e., Cronbach's alpha) were not able to be performed using this data set.

Executive Functions. Executive functioning was assessed with ExamCorp, which is a proprietary computer-based test battery that measures both executive functioning and personality (Peterson, 2010). The ExamCorp battery has not been validated with a pilot training population. The ExamCorp executive functioning measures include three working memory tasks (i.e., random letter span, random object span, and recency judgement), three associative learning tasks (i.e., acquired non-spatial association, acquired association, and go/no-go), two attention control tasks (i.e., response inhibition and sustained attention), and one verbal organization task (i.e., word fluency). As item-level measures were not available for all the ExamCorp sub-tests, more in-depth psychometric examinations (i.e., Cronbach's alpha) were not able to be performed using this data set. However, previous research (Peterson, 2003, 2010) indicates adequate reliability and validity for the ExamCorp measures, the specifics of which cannot be included in this document as they are protected by the test developer.

Working Memory. Working memory represents the ability to simultaneously access and manage many variables and large amount of information (ExamCorp, 2012). Individuals with excellent working memory capacity can simultaneously track multiple events while regulating themselves and their surrounding, tend to be efficient and fast in general reasoning tasks, and plan events in the absence of explicit instruction. The three working memory measures included Random Letter Span, Random Object Span, and Recency Judgement (Peterson, 2003). The Random Letter Span score represents the maximum span of letters that was successfully randomized (Peterson, 2003). The Random Object Span score represents the ratio of unique pictures (objects) correctly clicked from a randomized group of pictures, divided by the time required to successfully identify the 12 objects in successive, randomized arrays (Peterson, 2003). The Recency Judgement task score represents the number of correct trials in which the candidate correctly identified the most recent objects (Peterson, 2003).

Associated Learning. Associated learning refers to the ability to learn novel, situation-specific standards of importance (ExamCorp, 2012). Associated learning is an indicator of the ability to learn from error feedback, remember newly learned ideas and functions, and attention focus (i.e., focus on important items while disregarding trivial information; ExamCorp, 2012). There were three measures on associated learning, including Acquired Spatial Association, Non-Spatial Association, and Go/No-go (Peterson, 2003). The Acquired Spatial Association score represents the total number of trials the candidate required to correctly link a set of cards to randomly distributed lights (Peterson, 2003). The Non-Spatial Association score represents the number of trials required to learn the arbitrary associations between words and non-words (Higgins, 2009). The Go/No-go score represents the number of trials the candidate required to learn the go/no-go rule (Peterson, 2003).

Attention Control. Attention control refers to the ability to voluntarily direct attention to a given task (ExamCorp, 2012). Attention control involves sustained attention (i.e., concentrate on task-relevant details, without distraction, for long periods) and response inhibition (i.e., appropriately adjust behaviour and focus to the situation and control interfering impulses; ExamCorp, 2012). In this study, attention control was measured using the ExamCorp sub-tests Response Inhibition and continuous Sustained Attention. The Response Inhibition score represents the ratio of correct times the candidate inhibited their response to target, to the total number of times the target was presented (Peterson, 2003). The continuous Sustained Attention score represents the correct number of target the candidate clicked in a string of distract objects (Peterson, 2003).

Verbal Organization. Verbal organization represents the ability to efficiently and rapidly verbally sequence information in short-term memory in order to remember items in the order they appear (ExamCorp, 2012). Individuals scoring high on verbal fluency are expected to be good at ordering complex processes, knowing when things should be done in specific situations, think and act independently (ExamCorp, 2012). The ExamCorp battery measures verbal fluency with the Word Fluency measure. The score represent the total number of words generated, beginning with a specific letter, over a 5 minute period (Peterson, 2003).

Psychomotor ability. Psychomotor ability will be measured using two sub-tests from the Royal Air Force Aircrew Aptitude Test (RAFAAT). The psychomotor measures of the RAFAAT are computer-based, and involve the use of a joystick, or a joystick and foot pedals, depending on the test. As item-level measures were not available for all the RAFAAT psychomotor ability sub-tests, more in-depth psychometric examinations (i.e., Cronbach's alpha) were not able to be performed using this data set. However, previous research (NATO Training Group, 2008; Southcote, 2004) indicates adequate reliability for both CVT (Cronbach's alpha = 0.94; test-retest = .80) and SMA (Cronbach's alpha = 0.94; test-retest = 0.78). Uncorrected validity coefficients for Basic Flight Training were generally moderate for both CVT (r = .21) and SMA (r = .23). The battery includes two psychomotor measures:

Control of Velocity Test (CVT). The Control of Velocity Test is a pursuit-tracking test the measures hand-eye coordination using a joystick. The score represents the maximum number of dots successfully targeted.

Sensory Motor Apparatus (SMA). The Sensory Motor Apparatus test is a compensatory tracking test that measures hand-eye-foot coordination (Southcote, 2004), using the joystick and foot pedals. The score is an error score that reflects the degree the target dot deviated from the crosshairs. SMA was reverse coded for the analyses for ease of interpretation, so that a higher score represents better psychomotor ability.

Work sample. The work sample was measured using the Canadian Automated Pilot Selection System (CAPSS), which simulates an Instrument Flying Rules (IFR) environment in a light aircraft cockpit and is able to move in three axes (Woycheshin, 1999). While the CAPSS is accepted as a work sample measure (Woycheshin, 1999), a recent study (Darr, 2010b) found psychomotor ability to load highly on CAPSS performance. The CAPSS is administered at the Canadian Forces Aircrew Selection

Centre in Trenton, Ontario. The CAPSS testing involves four, one-hour long testing sessions, during which time applicants receive instructions and practice on the flight manoeuvres to be assessed during that session. Each session builds on the preceding sessions, with the final assessment of CAPSS success based on that applicant's performance on the final session. Previous research generally indicates correlations between CAPSS session 4 and Primary Flight Training outcomes ranging from .23 to .30 (Spearman's rho; Darr, 2010b; Skomorovsky & Donohue, 2010; Woycheshin, 1999, 2001b)

Personality. Personality was measured by a scale contained in the ExamCorp battery, Five Dimensional Temperament Inventory (FDTI). The FDTI is based on the Five Factor Model of personality, and includes the personality factors of Conscientiousness Achievement, Extraverted Sociability, Emotional Stability, Agreeable Amiability, and Intellectual and Cultural Interest (Peterson, 2003). It is a 50-item measure with 10 items for each of the five factors. The items are presented on a sliding, bipolar analogue scale in which three adjectives load on each end of the poles, representing positive or negative aspects of the underlying factor (Higgins, 2009). As item-level measures were not available for all the FDTI (ExamCorp) sub-tests, more in-depth psychometric examinations (i.e., Cronbach's alpha) were not able to be performed using this data set. However, previous research (Higgins, Peterson, Lee, & Pihl, 2007; Peterson, 2003, 2010) indicates adequate reliability and validity for the RAFAAT specific cognitive ability measures, the specifics of which cannot be included in this document as they are protected by the test developer.

Demographics questionnaire. The demographics questionnaire included questions of the candidate's gender, age, education, previous flying experience, and video game expertise (see Appendix E). Studies have found gender differences for performance in psychomotor ability (Carretta, 1997), and age differences in pilot's general mental ability (Hardy, Staz, D'Elia, & Uchiyama, 2007), executive functions (Higgins, Peterson, Lee, & Pihl, 2007), information processing speed (Hardy et al., 2007), and psychomotor ability (Hardy et al., 2007). Previous flying experience has also been associated with higher scores on CAPSS (Woycheshin, 1999), and Primary Flight Training (Carretta & Ree, 1994; Martinussen, 1996; Woycheshin, 2002). Video game expertise was included to investigate the relation between type of video game experience as a moderator of the relation between gender and performance on both the Sensory Motor Apparatus and CAPSS performance.

Criterion variables. The criterion variables used in this study pertain to Primary Flight Training (PFT) performance. The Phase I of Canadian Forces pilot training, Primary Flying Training course, is the common training required for all pilot trainees regardless of aircraft type (i.e., jet, multi-engine, and rotary wing); candidates with a commercial pilot's license bypass Primary Flight Training directly to Phase IIA (Basic Flying Training). This introductory flying course teaches visual flight (i.e., clearhood) aviation. The PFT course is approximately 3 months in duration and consists of two phases: the ground school or academic phase (15 training days) and the flying phase (23 training days, including 21 dual and 3 solo flights; 29.8 hrs flying). The flying phase was conducted on a single piston engine aircraft (Grob G120A) that has side-by-side seating to accommodate both candidate and instructor. The PFT training performance criterion are assessed by Pass/fail, Good/poor performance, academic grades, and flying average (daily flying lesson ratings), which are described in more detail below.

Training pass/fail. This is a dichotomous variable that indicates pilot training performance in terms of pass/fail. Training success and failure for the PFT course will be recorded as a dichotomous variable 0 (*training failures*) and 1 (*successful trainees*). To achieve a successful outcome, students must successfully complete the academic phase, meet daily flying lesson plan requirements, and succeed on both check flights. Upon completion of PFT, students are assigned to continue to advanced flying training in a specific aircraft stream (i.e., jet, multi-engine, and rotary wing); the airframe assignment is based on PFT performance combined with student preference and military requirements.

Good/Bad performance. This variable was created by creating a dichotomous variable from the four levels of the overall grade (F = fail, C, B, and A), to represent a weak (poor) performance group and a strong (good) performance group. As such, fail and C performers were combined into the poor performance group and the B and A performers were combined into the good performance group.

Academic average. The academic average reflects the overall combined average of all academic tests (e.g., aircraft operating instructions, regulations, flight operations, meteorology, and aerodynamics) administered during the academic phase.

Flying average. The flying average is a weighted average based on the sum of the grade achieved by the candidate on each assessed flight, divided by the maximum number of points possible, multiplied by 100 to determine a percent score[4]. The use of this variable was based on a similar variable computed by Carretta (2011), which accounts for flights not completed due to early failure of the course. For each flight a flying instructor assessed the student's performance on a variety of tasks required to be performed on that specific flight, and the score reflects the student's performance on the flying manoeuvres and procedures performed during the flight. The flying assessments are based on ratings of 0 (*unsatisfactory*), 1 (*marginal*), 2 (*achieved standard*), 3 (*exceeded standard*). For 65 of the 75 students, the total number of assessed flights conducted on the course was 21; therefore the maximum number of points was 63 (3 points for exceeded standard multiplied by 21 flights). However, for ten candidates a new training syllabus reduced the number of assessments for flights and simulators to 19; therefore the maximum number of points was 57.

Procedure

Ethics approval was obtained from both the Carleton University Psychology Ethics Committee (Reference # 11-101) and the Canadian Forces Social Science Review Board (Approval # 997/11-F, in accordance with CANFORGEN 198/08). At the beginning of training Primary Flight Training students were asked to participate in the study through a participant recruiting brief (see Appendix D). Once the purpose of the study was explained, participants were informed that their participation is voluntary, and they will be provided the informed consent (see Appendix E). The informed consent confirmed that the choice to participate would not be shared with the chain of command.

---

[4] That is (∑flight score/maximum points)*100

Furthermore, both the briefing and the informed consent confirmed that participation would also involve the candidate granting access to their selection tests results (i.e., Canadian Forces Aptitude Test [CFAT], Canadian Automated Pilot Selection System [CAPSS], and Royal Air Force Aircrew Aptitude Test [RAFAAT]), as well as access to their final course results (i.e., course report and progress cards). Upon completion of the informed consent, participants were asked to complete the demographics questionnaire (see Appendix F), and then asked to complete the computer-based test batteries. Participants had not completed the RAFAAT during Aircrew Selection were asked to do so, and all participants were asked to complete the ExamCorp battery. Verbatim instructions were created for the entire administration process (see Appendix G). Upon completion of all testing, participants were provided a Participant debrief form (see Appendix H). Archival data for the CFAT, CAPSS, and RAFAAT was retrieved from the treasury board database and Director General Military Personnel Research and Analysis.

The RAFAAT and ExamCorp testing was set up in 11 test stations; each station included a laptop, RAFAAT keypad, RAFAAT foot pedals, RAFAAT joystick, and mouse for the ExamCorp battery. The testing session for the RAFAAT required approximately 3.5 hours to complete, which included two 10 minute breaks. The ExamCorp required approximately 1.5 hours to complete, including one 5 minute break. The programs for both ExamCorp testing and the RAFAAT test were loaded on each laptop. Candidates had already completed the Canadian Automated Pilot Selection System (CAPSS) as part of the Canadian Forces aircrew selection phase.

Results

Preliminary Analyses

Data screening. All variables were screened for accuracy of data, out of range values, missing data, and univariate and multivariate outliers. Data were screened for both ungrouped variables (based on continuous outcome variables) and grouped variables (for dichotomous outcome variables) data. Univariate descriptives were run for all variables (independent and dependent variables) to check for out-of-range values, plausible means and standard deviations, and univariate outliers, as suggested by Tabachnick and Fidell (2007). The following results were obtained:

Out of range variables. One of the independent variables, Scheduling (from the Royal Air Force Aircrew Aptitude Test battery), had minimum values that were outside the range specified in the test manual. Further examination failed to reveal a reason these variables were out of range. Additional investigation with archived data revealed that this variable exhibited numerous below-minimum values. Previous studies also identified similar issues with the Scheduling sub-test. As the validity of the Scheduling sub-test was identified as being less than .1 (Southcote, 2004), the Director General Military Personnel Research and Analysis recommended the removal of the Scheduling sub-test from the Canadian Forces Pilot Selection model. The decision to remove the sub-test occurred after the commencement of this current study.

Reverse coding. Review of the preliminary results also revealed negative associations between the Sensory Motor Apparatus variable and other variables due to the fact that the score was based on an error score. Therefore, to facilitate interpretation, a new variable was computed using the inverse Sensory Motor Apparatus score. The new inverse Sensory Motor Apparatus variable was used in all analysis.

Dichotomous variables. Initial investigation of the Pass/Fail Grade variable revealed a group split of 64 to 11, representing roughly a 6 to 1 ratio (85% to 15%), which approached the extreme threshold of a 90 to 10 split suggested by Tabachnick and Fidell (2007). As this degree of split would create difficulties with additional analysis, such as deflated associations with other variables, it was decided to investigate other outcome variable options. The course report for each candidate included a categorical descriptor of the candidate's overall performance; F = fail, C, B, and A. Although the categories (F, C, B, and A) were assessed separately, violations of the assumptions due to many cells having frequencies below 5 meant invalid results. To remedy this situation, a dichotomous variable was created by combining the fail and C performers in a poor performance group and the B and A performers in a good performance group. This also addressed a concern expressed by the Royal Canadian Air Force, specifically Brigadier-General Galvin (Commander 2 Canadian Air Division, personal communication, May 23, 2012) who wants weak performers weeded out early, to reduce training attrition and the associated training costs on successive courses. As a result, the new Good/Poor Performance variable group size difference improved somewhat to a 60 to 15 (80% to 20%) split, representing a 4 to 1 ratio. Although these ratios don't exceed the 10 to 1 ratio, or 90 to 10 percent split, identified by Tabachnick and Fidell (2007) as cause for serious concern, the degree of split suggested there would likely be deflated relations with other variables.

Missing data. As indicated in the Methods section, the sample sizes varied due to missing data on different variables which were typically due to computer glitches and administration timing issues. Specifically, a computer glitch occurred during the RAFAAT administration; participant data was lost as a result of an unexpected disconnection of the student laptop from the network supervisor laptop during administration on two separate dates, using two different computers. As for the administration timing issues, early administrations of the computer-based testing occurred during the very stressful flying phase of training, which resulted in substantially fewer participants than later administrations that dedicated one specific day to testing. As such, the missing values were designated as missing (i.e., 999) within the dataset. The results of the missing data are represented in Table 3.

Table 3
Number of Missing Cases and Total n for Study Variables

Variables	Missing	Total n
CFAT – sub-tests and total score	0	75
Demographics – age, gender, education, etc.	0	75
RAFAAT – all sub-tests minus vigilance	5	70
RAFAAT – vigilance sub-test	6	69
ExamCorp – executive functioning	2	73
ExamCorp – personality	2	73
Course report data	0	75

Univariate and multivariate outliers. All variables were examined for univariate and multivariate outliers. *Z*-scores were calculated for all variables, and then all cases were examined for values exceeding $z = 2.58$ ($p < .05$), which follows a more stringent criterion for the small sample size suggested by Field (2009) and Tabachnick and Fidell (2007). Based on the $z = 2.58$ criterion, one univariate outlier was found on different cases on each of the following variables: CAPSS, CFAT Total Score, CFAT Spatial Ability, Control of Velocity, Recall, Vigilance, Visual Search, Conscientiousness, Emotional Stability, Intellect/Culture, Non-Spatial Association, Academic Grade, and

Flying Average. Additionally, two univariate outliers exceeding $z = 2.58$ were found on different cases on Critical Reasoning Battery Verbal, Spatial Association, Recency Judgement, and Word Fluency. The outliers were dealt with by adjusting the corresponding raw score to be within 2.58 standard deviations from the mean, a method suggested by Field (2009) and Tabachnick and Fidell (2007).

Univariate outliers on the predictor variables were also checked by group, based on the dichotomous training outcome variable Good/Poor Performance. Due to the small group size for the smallest group split (i.e., fail or poor performance), the data set was examined for standardized scores exceeding $z = 1.96$, $p < .05$ (Field, 2009). Based on this criterion, univariate outliers were identified on different cases on the Good/Poor Performance; the poor performance group had one outlier on 15 variables (all cases were different) and two outliers on different cases on one variable, whereas the good performance group had univariate outliers on different, non-overlapping cases on all variables.

The data were checked for multivariate outliers by examining Mahalanobis distances (Tabachnick & Fidell, 2007). The computed Mahalanobis Distances did not exceeded the critical value, $\chi^2(34) = 65.18$, $p < .001$, thus no multivariate outliers were present in this data. Results by grouped variables also indicate that no multivariate outliers existed on the Good/Poor Performance variable.

Summary. In summary, all variables were examined for accuracy and out-of-range scores. This resulted in the deletion of one variable from the dataset, the RAFAAT sub-test Scheduling, due to unexplained out-of-range values. On account of the severe group split on the Pass/Fail Grade (85% to 15%), a new variable with a slightly better group split (80% to 20%) was created, resulting in the Good/Poor Performance variable. Missing scores were found on the ExamCorp and RAFAAT sub-tests; as these missing scores were attributed to experimental issues (i.e., scheduling, computer glitches) the missing values were retained and identified as 999. Univariate outliers were identified on numerous variables, but were not attributed to only a few cases. Univariate outliers were dealt with by adjusting the corresponding raw scores to be within 2.58 standard deviation of the mean. No multivariate outliers were identified. After the data were screened, the next step involved the testing of assumptions for correlation and multiple regression analyses.

Testing of Assumptions

Normality. All variables were assessed for deviations to normality by calculating the standardized skewness and kurtosis values, which were then compared to the recommended threshold value for small samples of $z = 1.96$ ($p < .01$; Field, 2009). After adjusting for outliers, as identified above, CFAT Spatial Ability, CFAT Problem Solving, and Non Spatial Association exceeded standardized skewness scores of 1.96, but were less than 2.58. The Flying Average outcome variable, with a standardized skewness of -3.25, approached a standardized skewness value of almost 3 standard deviations above the mean. These variables were investigated further to determine whether transformations would significantly improve the distributions.

Transformations of the original CFAT Spatial Ability, CFAT Problem Solving, Non Spatial Association, and Flying Average were subsequently investigated, prior to cleaning and adjustment of outliers. Due to the lack of improvement in bivariate correlations as well as the difficulty in interpreting analysis results, it was decided to use the untransformed CFAT Spatial Ability, CFAT Problem Solving, Non Spatial Association, and Flying Average variables in subsequent analyses. Consequently, regression analysis involving CFAT Spatial Ability, CFAT Problem Solving, Non Spatial Association, and Flying Average were interpreted with caution as the validity of the findings are likely to be biased due to the influence of the skewness on the parametric tests. As the Flying Average variable was extremely skewed, having a standardized score that exceeded 3 standard deviations, and in recognition of the small sample size, Kendall's correlation coefficient (τ_b) was calculated (Tabachnick & Fidell, 2007).

The distribution of the group means of the grouped outcome variable Good/Poor Performance were also assessed for normality. After adjusting the raw scores of outlying cases by group, the standardized skewness and kurtosis values for all variables in the poor performance group were less than $z = 1.96$. Similarly, with the exception of one variable (CFAT problem Solving; $z_{skewness} = -2.17$), the standardized skewness and kurtosis values for the good performance group were less than $z = 1.96$. Therefore, both groups on all variables met the assumption of normality and were used with confidence in further grouped analyses, such as correlations and independent samples t-tests.

Linearity. Based on the non-normal distributions of a number of the study variables, assessing the linearity of relations among variables was of even greater importance for linear/multiple regression analysis (Tabachnick & Fidell, 2007). For linear regression, bivariate linearity was observed for all variables in this data set through examination of the bivariate scatterplots, for both ungrouped and grouped data. However, examination of the residual scatter plots indicated minor deviations of the Lowess line from the 0-line for most plots, which suggested the relations between the independent variables and dependent variables violated the assumption of linearity (Cohen, Cohen, West, & Aiken, 2003). Given the conflicting scatterplot results, and recognizing the deviations from linearity were minor for the residuals scatterplot, the assumption of linearity was considered to be met for linear/multiple regression.

Homoscedasticity. Examination of the bivariate scatterplots indicated the assumption of homoscedasticity has been violated with many of the variables. This was supported by examination of the residuals scatterplots, which indicated that the conditional variance of the residuals around the regression line was not constant (Cohen, Cohen, West, & Aiken, 2003). Therefore, the assumption of homoscedasticity has been violated. Consequently, the results of the all analyses were weakened, but not invalidated, and the confidence intervals may be incorrect (Cohen, Cohen, West, & Aiken, 2003; Field, 2009).

Homogeneity of variance. Homogeneity of variance was assessed by examination of Levene's test for all variables grouped by Good/Poor Performance. Significant Levene's test results were identified for variables Non-Spatial Association, Spatial Association, and Sustained Attention grouped by Good/Poor Performance. Therefore, the assumption of homogeneity of variance was violated for these variables and the variances of these variables may not be stable at all levels of the outcome variable (Cohen, Cohen, West, & Aiken, 2003). The result could be inaccurate associations between variables, and analyses involving these variables (specifically Hypotheses 5 and 9) should be interpreted with caution.

Multicollinearity. Data were checked for the presence of multicollinearity by examination of the bivariate correlations for values greater than $r = .90$. No correlations were found to exceed the $r = .90$ guideline (Tabachnick & Fidell, 2007). This was followed up with the examination of collinearity diagnostics (conditioning index and variance proportion), and VIF and tolerance values (Field, 2009; Tabachnick & Fidell, 2007) when each hypothesis was tested. The average VIF near 1 and tolerance values greater than .2 indicated that multicollinearity did not bias the regression model. Tolerance values were all acceptable as all values were greater than .2. Although the last root of the condition index was large for some of the hypotheses, no row had more than one variance proportion greater than .50. Therefore, all collinearity diagnostics were within acceptable values, and multicollinearity is not present in this dataset.

Independence of errors. The independence of errors was assessed with the Durbin-Watson statistic (Field, 2009). The Durbin-Watson values were between 1 and 3 for all hypotheses, indicating that the assumption of independence had not been violated.

Residuals and influence. To assess the regression model and whether the model fits the observed data well, the residuals were assessed for influential cases. Examination of the standardized residuals for each hypothesis indicated that only hypotheses 2 and 9 did not have standardized residuals that exceeded 2 on more than 5% of the cases. Therefore, the model is a good representation of the actual data only for those two hypotheses. The maximum value of Cook's Distance was significantly less than 1 for all hypotheses, suggesting there were no influential data points using this statistic (Field, 2009). However, leverage values exceeding 3 times the average leverage[5] (Stevens, 2002), were identified on hypotheses 3 and 8. Examination of the standardized DFBetas

---

[5] Average leverage = $(k + 1)/n$, where k = number of predictors, n = sample size. For this data set, average leverage is .0882 (Field, 2009)

for all variables identified absolute values greater than 2 for only 2 cases in the entire data set on the CAPSS variable in hypothesis 7, which had a large influence on the model's parameters (Field, 2009). Overall, the examination of the influence statistics indicated there may be issues with the results of all hypotheses with the exception of hypothesis 2; therefore some caution should be used when interpreting the results.

Summary. The tests of the assumptions of multiple linear regression analysis for this data set identified a number of violations. Although the results indicated that the assumptions of independence of errors and multicollinearity were not violated for any data, the other assumptions (i.e., normality, linearity, homoscedasticity, and homogeneity of variance) were violated to some degree on most variables. Additionally, with the exception of hypothesis 2, influence statistics indicate there were cases that have influenced the results of the analysis. Based on this information, the results of this study should be interpreted with caution due to the violations to the assumptions and influential cases that may have biased the results of the multiple regression analysis.

Main Analyses

Due to the large number of variables of interest in this study, the analyses were organized by hypothesis, including the descriptive statistics (e.g., means and standard deviation) for both grouped and ungrouped data, as well as correlations and multiple regression analysis. Pearson correlations were used to determine the associations between the continuous predictor variables and the academic outcome variable, whereas Kendall's correlation coefficient (τ_b) was used for the correlations between the predictors and the skewed flying average outcome variable. Kendall's correlation coefficient is the preferred correlation for non-parametric (skewed) data when the sample size is small

(Field, 2009). Biserial correlations were computed to determine the relations between the continuous predictor variables and the dichotomous outcome variable, Good/Poor Performance (Field, 2009). In addition to identifying the type of correlations to be performed based on the characteristics of the variables in this study, the issue of range restriction was also considered and addressed.

One expected issue in this study was the restriction of range in the scores of a variable, which often occurs as a result of a sample having been preselected based on the scores on a variable. As the pilot candidates in this study had been preselected based on CFAT and CAPSS scores, these scores were considered to be restricted in range. Of particular note is the CAPSS variable, which was the only pilot-specific selection criteria used by the Canadian Forces, which has a minimum cut-off score of .70 (from range of 0 to 1). Unfortunately, it was impossible to perform the calculations to correct for range restriction for the CAPSS variable as the research and cross-validation of the CAPSS system was also performed on a sample that was already selected using the selection system of the time, and therefore no data is available for a non-restricted sample. According to Tabachnick and Fidell (2007), the consequence of a restricted range of scores in the sample can be too small (i.e., insignificant) correlations involving the restricted variable. Therefore, small or insignificant results were anticipated for correlations involving the CAPSS variable.

Previous flying experience as a predictor of flying training performance. Hypothesis 1 proposed that higher levels of previous flying experience, measured by fixed wing flying hours, are associated with higher levels of pilot training performance. The descriptive statistics and bivariate correlations for previous flying experience and the training outcome variables are presented in Table 4. The correlation analysis suggested there was a significant positive relation between previous flying experience and both the ground school academic grade and the flying average. Therefore, higher levels of previous flying experience (i.e., more flying hours) were associated with higher ground school academic grades and higher averages for the flying assessment (i.e., flying average). However, the correlation analysis revealed that there was no significant association between previous flying experience and either of the two dichotomous predictors of overall training performance, specifically Pass/Fail and Good/Poor Performance. The biserial correlations indicated there were no significant associations between previous flying experience and either the Pass/Fail outcome or the Good/Poor Performance outcome.

Table 4
Descriptive Statistics and Correlations for Previous Flying Experience and All Training Performance Outcome Variables

Variables	*n*	*M*	*SD*	1	2	3	4	5
1. Previous flying experience[a]	75	40.36	45.97	-				
2. Academic grade[a]	75	96.52	1.99	.38*	-			
3. Flying average[b]	75	67.48	15.39	.21**	.23**	-		
4. Pass/Fail[cd]	75	0.85	0.356	.14	.22*	.79**	-	
5. Good/Poor Performance[ce]	75	0.80	0.403	.11	.26*	.76**	.83**	-

[a] Pearson correlations (1-tailed)
[b] Kendall's tau correlations (1-tailed)
[c] Biserial correlations (1-tailed)
[d] Pass/Fail: 0 = fail; 1 = pass
[e] Good/Poor Performance: 0 = poor; 1 = good
* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Despite the non-significant result of the Chi-square analysis (Table 5), the odds ratio indicated that the odds of passing Primary Flight Training were 2.48[6] times higher for candidates who had previous flying experience than for those who did not have previous flying experience.

Table 5

Chi-square Test of Pass/Fail Grade and Previous Flying Experience (PFE; n = 75)

Previous Flying Experience	Pass/Fail Grade fail	pass	*Pearson* χ^2 *(p)*	*Likelihood ratio (p)*	*Fisher's exact test (p)*	*Odds ratio*
No PFE	4	12	1.73	1.56	n/a[a]	2.48
	(1.1)	(-.4)	(.19)	(.21)	(.18)	
PFE	7	52				
	(-.6)	(.2)				

Note. Adjusted standardized residuals appear in parentheses below group frequencies.
[a] n/a = not applicable; Fisher's exact test is a measure of probability

Summary of previous flying experience predictor. Hypothesis 1 predicted that higher level (hours) of previous flying experience would be associated with better performance on pilot training outcomes. The results indicate that the hypothesis was supported, as previous flying experience was positively associated with performance on both the academic grade and the flying average. Additionally, the results of the odds ratio indicate that individuals with previous flying experience are almost 2.5 times more likely to pass Primary Flight Training than individuals who do not have flying experience.

[6] Field (2009) recommends using the odds ratio for 2x2 tables as a more useful measure of effect size. The odds ratio was calculated as follows: $OddsRatio = \frac{odds_1}{odds_2} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$

(*www.stat.ufl.edu/~winner/sta6934/CDA1.ppt*)

Intelligence predictors of flying training performance. Hypothesis 2 proposed that higher scores on the Canadian Forces Aptitude Test and the RAFAAT Critical Reasoning Battery are associated with higher levels of pilot training performance (Good/Poor Performance, Academic Grade). The descriptive statistics and bivariate correlations for the CFAT, CRB, and training outcome variables are presented in Table 6. The correlation analysis suggested the following predictors had a significant positive relation with the academic grade (i.e., ground school); CFAT Total Score, CFAT Spatial Ability, and CFAT Problem Solving. Therefore, as performance on CFAT Total Score, Spatial Ability and Problem Solving improves, the performance on the academic portion of Primary Flight Training also improves. Additionally, the correlation analysis revealed that CFAT Spatial Ability and Problem Solving were also positively associated with the Flying Average. However, the biserial correlations between the predictors and the grouped Good/Poor Performance variable indicated there were no significant relations. Conversely, the correlation analysis indicated that the CRB variables were not significantly associated with any of the training outcome variables.

Table 6

Descriptive Statistics and Correlations for Canadian Forces Aptitude Test (CFAT), Critical Reasoning Battery (CRB), and Training Performance Outcome Variables

	n	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. CFAT Total Score [a]	75	48.64	5.54	-										
2. CFAT Verbal Skills [a]	75	11.23	2.20	.52**	-									
3. CFAT Spatial Ability [a]	75	12.17	2.09	.70**	.17	-								
4. CFAT Problem Solving [a]	75	25.23	3.55	.83**	.13	.41**	-							
5. CRB Total Score [a]	70	25.66	5.86	.54**	.09	.39**	.56**	-						
6. CRB Diagrammatic [a]	70	8.00	2.12	.28**	-.08	.27*	.32**	.67**	-					
7. CRB Numerical [a]	70	6.94	3.50	.59**	.13	.40**	.61**	.90**	.43**	-				
8. CRB Verbal [a]	70	10.81	1.70	.27*	.14	.16	.23*	.63**	.17	.41**	-			
9. Academic Grade [a]	75	96.52	1.99	.40**	.12	.22*	.40**	.18	.11	.18	.16	-		
10. Flying Average [b]	75	67.48	15.39	.11	-.13	.17*	.18*	.07	.12	.04	-.01	.23**	-	
11. Good/Poor Performance [cd]	75	0.80	0.40	.04	-.23	.20	.05	.001	.12	-.04	-.02	.37*	1***	-

[a] Pearson correlations (1-tailed)

[b] Kendall's tau correlations (1-tailed)

[c] Biserial correlations (1-tailed)

[d] Good/Poor Performance: 0 = poor; 1 = good

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Intelligence predictors of good/poor performance. An independent samples t-test was conducted in order to confirm the results of the biserial correlations by comparing the means of good and poor training performance on each of the general cognitive ability predictors. There were no significant differences in the scores of the predictors grouped by good and poor performance (see Table 7). However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 7
Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for Canadian Forces Aptitude Test (CFAT) and Critical Reasoning Battery (CRB) Total Score and Sub-Tests for Good and Poor Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
CFAT Total Score[a]	48.72 (5.71)	48.33 (4.98)	-0.24	73	.81	-3.59	2.82
CFAT Verbal Skills[a]	11.05 (2.27)	11.93 (1.83)	1.40	73	.17	-0.38	2.14
CFAT Spatial Ability[a]	12.32 (1.95)	11.6 (2.56)	-1.19	73	.24	-1.91	0.48
CFAT Problem Solving[a]	25.28 (3.78)	25.00 (2.51)	-0.27	73	.78	-2.34	1.77
CRB Total Score[b]	25.66 (6.13)	25.64 (4.80)	-0.01	68	.99	-3.54	3.50
CRB Diagrammatic[b]	8.09 (2.23)	7.64 (1.60)	-0.70	68	.48	-1.71	0.82
CRB Numerical[b]	6.89 (3.59)	7.14 (3.21)	0.24	68	.81	-1.85	2.35
CRB Verbal[b]	10.80 (1.76)	10.85 (1.46)	0.10	68	.92	-0.96	1.07

Note. Levene's test was not significant for all variables; results presented are based on equal variance assumed. The assumption of homogeneity of variance was not violated.
[a] CFAT variables: Good, $n = 60$; Poor $n = 15$
[b] CRB variables: Good, $n = 56$; Poor $n = 14$

Intelligence predictors of academic grade. A backward stepwise multiple regression analysis was performed using the sub-tests of the Canadian Forces Aptitude

Tests and Critical Reasoning Battery (i.e., CFAT Verbal Skills, Spatial Ability, and Problem Solving; Critical Reasoning Battery Diagrammatic, Verbal and Numerical) as the predictors and the academic grade as the dependent variable. The composite total score variables for both CFAT and CRB were not included in the analysis as it was thought that the sub-tests would provide more meaningful results. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 8.

The regression with all the variables in the model (Model 1) was not significant, $R = .41$, $R^2 = .17$, $F(6, 63) = 2.17$, $p = .06$. The Adjusted R^2 value of .09 indicated that only 9% of the variability in academic grade was predicted by a model with all the CFAT and Critical Reasoning Battery sub-test variables. Moreover, in the initial model, only one of the six variables exhibited a regression coefficient that was significantly different from zero; CFAT Problem Solving, $\beta = .41$, $t\,(63) = 2.76$, $p < .01$. Therefore, higher levels of problem solving are associated with higher academic grades. There were no statistically significant linear relations detected between the mean of academic average and any of the remaining predictors. Although all subsequent models progressively strengthened in significance with the gradual removal of the weakest predictors, the only variable that demonstrated a significant relation with the academic grade was CFAT Problem Solving, which it maintained through all models. The final and optimal model (Model 6), included only CFAT Problem Solving as a significant predictor of the academic grade, which accounted for 13.6% of the variance, Adjusted $R^2 = .136$; $F(1, 68) = 11.822$, $p < .001$.

Table 8

Backward Stepwise Regression of Canadian Forces Aptitude Test (CFAT) and Critical Reasoning Battery (CRB) Sub-tests on Academic Grade (n = 70)

Predictor	R	R^2	$Adj\ R^2$	ΔR^2	F	p	B	$SE\ B$	β	t	p	r	pr	sr	sr^2
Model 1	0.41	0.17	0.09	0.17	2.17	.06									
CFAT VS							.04	.11	.04	.34	.74	.10	.04	.04	.00
CFAT SA							.08	.12	.08	.64	.53	.22	.08	.07	.00
CFAT PS							.23	.08	.41	2.76	.008	.39	.33	.32	.10
CRBD							.00	.12	.005	.04	.97	.11	.005	.004	.00
CRBN							-.09	.10	-.16	-.96	.34	.18	-.12	-.11	.01
CRBV							.13	.15	.11	.85	.40	.16	.11	.10	.01
Model 2	0.41	0.17	0.11	0.00	2.64	.03									
CFAT VS							.04	.11	.04	.34	.74	.10	.04	.04	.00
CFAT SA							.08	.12	.09	.65	.52	.22	.08	.07	.00
CFAT PS							.23	.08	.41	2.79	.007	.39	.33	.32	.10
CRBN							-.09	.09	-.16	-.99	.32	.18	-.12	-.11	.01
CRBV							.13	.15	.11	.86	.39	.16	.11	.10	.01
Model 3	0.41	0.17	0.12	-0.00	3.32	.02									
CFAT SA							.09	.12	.09	.71	.48	.22	.09	.08	.00
CFAT PS							.23	.08	.42	2.82	.006	.39	.33	.32	.10
CRBN							-.09	.09	-.16	-1.00	.32	.18	-.12	-.11	.01
CRBV							.13	.15	.11	.91	.37	.16	.11	.10	.01
Model 4	0.40	0.16	0.12	-0.00	4.29	.008									
CFAT PS							.24	.08	.44	3.10	.003	.39	.36	.35	.12
CRBN							-.08	.09	-.13	-.89	.38	.18	-.11	-.10	.01
CRBV							.13	.15	.11	.91	.37	.16	.11	.10	.01
Model 5	0.38	0.15	0.13	-0.01	6.06	.004									
CFAT PS							.20	.06	.37	3.19	.001	.39	.36	.36	.13
CRBV							.09	.14	.07	.64	.53	.16	.08	.07	.00
Model 6	0.38	0.15	0.14	-0.00	11.82	.001									
CFAT PS							.21	.06	.39	3.44	.001	.39	.39	.39	.13

Note. CFAT VS = CFAT Verbal Skills; CFAT SA = CFAT Spatial Ability; CFAT PS = CFAT Problem Solving; CRBD = CRB Diagrammatic; CRBN = CRB Numerical; CRBV = CRB Verbal; r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation

Intelligence predictors of flying average. To determine the associations of the intelligence predictors with the flying average outcome variable, a backward stepwise multiple regression analysis was performed using the sub-tests of the Canadian Forces Aptitude Tests and Critical Reasoning Battery (i.e., CFAT Verbal Skills, Spatial Ability, and Problem Solving; Critical Reasoning Battery Diagrammatic, Verbal and Numerical) as the predictors. The composite total score variables for both CFAT and CRB were not included in the analysis as it was thought that the sub-tests would provide more meaningful results. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 9.

The regression with all the variables in the model (Model 1) was not significant, $R = .40$, $R^2 = .16$, $F(6, 63) = 1.95$, $p = .09$. The Adjusted R^2 value of .076 indicated that only 7.6% of the variability in academic grade was predicted by a model with all the CFAT and Critical Reasoning Battery sub-test variables. However, in the initial model, only two of the six variables exhibited a regression coefficient that was significantly different from zero; CFAT Verbal Skills, $\beta = -.28$, $t\,(63) = -2.31$, $p < .05$, and CFAT Spatial Ability, $\beta = .31$, $t\,(63) = 2.36$, $p < .05$. Therefore , lower levels of verbal skills are associated with higher flying averages whereas higher levels of spatial ability are associated with higher flying averages. There were no statistically significant linear relations detected between the mean of flying average and any of the remaining predictors. The two CFAT variables were the only predictors to exhibit a significant relation with the flying average outcome variable, associations which were maintained

Table 9

Backward Stepwise Regression of Canadian Forces Aptitude Test (CFAT) and Critical Reasoning Battery (CRB) Sub-tests on Flying Average (n = 70)

Predictor	*R*	R^2	Adj R^2	ΔR^2	*F*	*p*	*B*	*SE B*	*β*	*t*	*p*	*r*	*pr*	*sr*	sr^2
Model 1	0.40	0.16	0.08	0.16	1.95	.09									
CFAT VS							-2.01	0.87	-.28	-2.31	.02	-.22	-.28	-.27	.07
CFAT SA							2.31	0.98	.31	2.36	.02	.28	.29	.27	.08
CFAT PS							0.30	0.16	.07	0.46	.65	.14	.06	.05	.00
CRBD							0.16	0.98	.02	0.16	.87	.13	.02	.02	.00
CRBN							-0.23	0.75	-.05	-0.31	.76	.09	-.04	-.04	.00
CRBV							0.03	1.19	.00	0.02	.98	.01	.00	.00	.00
Model 2	0.40	0.16	0.09	0.00	2.38	.05									
CFAT VS							-2.01	0.86	-.28	-2.34	.02	-.22	-.28	-.27	.07
CFAT SA							2.31	0.97	.31	2.38	.02	.28	.29	.27	.08
CFAT PS							0.30	0.65	.07	0.46	.65	.14	.06	.05	.00
CRBD							0.16	0.97	.02	0.16	.87	.13	.02	.02	.00
CRBN							-0.23	0.70	-.05	-0.32	.75	.09	-.04	-.04	.00
Model 3	0.39	0.16	0.10	0.00	3.01	.02									
CFAT VS							-2.03	0.84	-.28	-2.42	.02	-.22	-.29	-.28	.08
CFAT SA							2.33	0.96	.32	2.44	.02	.28	.29	.28	.08
CFAT PS							0.31	0.64	.07	0.48	.64	.14	.06	.05	.00
CRBN							-0.19	0.66	-.04	-0.29	.77	.09	-.04	-.03	.00
Model 4	0.39	0.15	0.12	-.001	4.05	.01									
CFAT VS							-2.04	0.83	-.28	-2.45	.02	-.22	-.29	-.28	.08
CFAT SA							2.28	0.93	.31	2.44	.02	.28	.29	.28	.08
CFAT PS							0.21	0.54	.05	0.38	.70	.14	.05	.04	.00
Model 5	0.39	0.15	0.13	-0.00	6.07	.004									
CFAT VS							-2.02	0.83	-.28	-2.45	.02	-.22	-.29	-.28	.08
CFAT SA							2.43	0.85	.33	2.87	.01	.28	.33	.32	.10

Note. CFAT VS = CFAT Verbal Skills; CFAT SA = CFAT Spatial Ability; CFAT PS = CFAT Problem Solving; CRBD = CRB Diagrammatic; CRBN = CRB Numerical; CRBV = CRB Verbal; *r* = zero-order correlation; *pr* = partial correlation; *sr* = semipartial correlation; sr^2 = squared semipartial correlation for all models in the analysis. The final and optimal model (Model 5), included only the CFAT Verbal Skills and Spatial Ability variables as a significant predictors of the flying average, and accounted for 12.8% of the variance, Adjusted $R^2 = .128$; $F(2, 67) = 6.07$, $p < .01$. The CFAT Spatial Ability was positively associated with the flying average, $\beta = .33$, $t\,(69) = 2.87$, $p < .01$, and accounted for 10% of the unique variance in the flying average. Conversely, CFAT Verbal Skills was negatively associated with the flying average, $\beta = -.28$, $t\,(69) = -2.45$, $p < .05$, and accounted for almost 8% of the unique variance in the flying average.

Summary of intelligence predictors. For Hypothesis 2, there were no significant associations between any of the intelligence predictors and the Good/Poor Performance training outcome. However, there were significant associations between intelligence and the academic grade and flying average. CFAT Problem Solving was positively related to the academic grade. The correlation analysis for the intelligence predictors and flying average, which used Kendall's tau b to account for the skew of the flying average, indicated that the only significant correlation was between Non-Spatial Association and the flying average (Kendall's $\tau_b = .18$, $p < .05$). However, the backward regression indicated that CFAT Verbal Skills was negatively related to the flying average, whereas CFAT Spatial Ability was positively related to the flying average.

Specific cognitive ability predictors of flying training performance. Hypothesis 3 proposed that higher levels of work rate (i.e., information processing speed) and digit recall are associated with higher levels of pilot training performance. The descriptive statistics and bivariate correlations are provided at Table 10. The variables Table Reading and Visual Search, which are both measures of work rate, were both positively associated with academic (i.e., ground school) performance. However, none of the specific cognitive ability measures exhibited a significant relation with either of the flying training performance variables.

Table 10

Descriptive Statistics and Correlations for Specific Cognitive Ability Tests (Measured by the Royal Air Force Aircrew Aptitude Test Battery) and Training Performance Outcome Variables

Variables	*n*	*M*	*SD*	1	2	3	4	5	6	7
1. Table Reading[a]	70	65.69	10.19	-						
2. Recall[a]	70	100.74	11.29	.08	-					
3. Vigilance[a]	69	150.07	26.51	.42**	.11	-				
4. Visual Search[a]	70	113.96	11.28	.38**	.23*	.31**	-			
5. Academic Grade[a]	75	96.52	1.99	.29**	.17	.19	.22*	-		
6. Flying_Average[b]	75	67.48	15.39	.12	-.02	-.05	.12	.32**	-	
7. Good/Poor Performance[cd]	75	0.80	0.40	.15	-.13	.17	.04	.37*	1***	-

[a] Pearson correlations (1-tailed)

[b] Kendall's tau correlations (1-tailed)

[c] Biserial correlations (1-tailed)

[d] Good/Poor Performance: 0 = poor; 1 = good

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Specific cognitive ability predictors of good/poor performance. To confirm the non-significant results of the biserial correlations an independent samples t-test was conducted. The means of good and poor training performance were compared for all the RAFAAT specific cognitive ability variables. There were no significant differences in the scores between good and poor performance for any of the RAFAAT predictors (see Table 11). However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 11

Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for Specific Cognitive Ability Measures (RAFAAT) by Good (n = 56) and Poor (n = 14) Training Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
Table Reading	66.23	63.50	-0.90	68	.37	-8.82	3.35
	(10.28)	(9.87)					
Digit Recall	100.25	102.71	0.73	68	.47	-4.29	9.22
	(11.13)	(12.14)					
Vigilance	151.62	144.00	-0.96	67	.34	-23.47	8.23
	(26.87)	(25.07)					
Visual Search	114.11	113.36	-0.22	68	.83	-7.52	6.02
	(11.75)	(9.54)					

Note. Levene's test was not significant for all variables; results presented are based on equal variance assumed. The assumption of homogeneity of variance was not violated.

Specific cognitive ability predictors of academic grade. A backward stepwise multiple regression analysis was performed using the academic grade as the dependent variable and the RAFAAT specific cognitive ability measures (i.e., Table Reading, Digit Recall, Vigilance, and Visual Search) as the independent variables. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 12.

The regression with all the variables in the model (Model 1) was not significant, $R = .32$, $R^2 = .10$, $F(4, 64) = 1.79$, $p = .14$. The Adjusted R^2 value of .04 indicated that only 4% of the variability in academic grade was predicted by a model with all the RAFAAT specific cognitive ability measures in the model. Moreover, in the initial model, none of the four variables exhibited a regression coefficient that was significantly different from zero. Model 2, in which Vigilance was removed from the model, was also not

Table 12

Backward Stepwise Regression of RAFAAT Cognitive Ability Sub-tests on Academic Grade (n = 69)

Predictor	R	R^2	$Adj\ R^2$	ΔR^2	F	p	B	$SE\ B$	β	t	p	r	pr	sr	sr^2
Model 1	0.32	0.10	0.04	0.10	1.79	.14									
Table Reading							0.04	0.03	.21	1.53	.13	.27	.19	.18	.03
Recall							0.02	0.02	.11	0.91	.37	.15	.11	.11	.01
Vigilance							0.00	0.01	.06	0.46	.64	.19	.06	.06	.03
Visual Search							0.02	0.02	.08	0.64	.53	.20	.08	.08	.00
Model 2	0.31	0.10	0.06	-0.00	2.34	.08									
Table Reading							0.05	0.03	.23	1.81	.08	.27	.22	.21	.05
Recall							0.02	0.02	.11	0.95	.35	.15	.12	.11	.01
Visual Search							0.02	0.02	.09	0.73	.47	.20	.09	.09	.00
Model 3	0.30	0.09	0.06	-0.02	3.27	.04									
Table Reading							0.05	0.02	.26	2.34	.03	.27	.27	.26	.07
Recall							0.02	0.02	.13	1.12	.27	.15	.14	.13	.02
Model 4	0.27	0.07	0.06	-0.02	5.27	.02									
Table Reading							0.05	0.02	.27	2.30	.02	.27	.27	.27	.07

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation significant, $R = .31$, $R^2 = .10$, $F(3, 65) = 2.34$, $p = .08$. Only 6% (Adjusted $R^2 = .056$) of the variability in academic grade was accounted for by Model 2, and none of the individual predictors exhibited a regression coefficient that was significantly different from zero. Visual Search was removed in Model 3, and the model with only Table Reading and Recall was significant, $R = .30$, $R^2 = .09$, $F(2, 66) = 3.27$, $p < .05$. Table Reading and Recall accounted for 6% (Adjusted $R^2 = .063$) of the variability in the academic grade. However, only Table Reading exhibited a regression coefficient that was significantly different from zero; $\beta = .26$, $t\,(68) = 2.23$, $p < .05$. The final model, Model 4 with only Table Reading, was significant, $R = .27$, $R^2 = .07$, $F(1, 67) = 5.27$, $p < .05$. For Model 4, 6% (Adjusted $R^2 = .059$) of the variability in academic grade was accounted for by a model with only Table Reading in the equation; $\beta = .27$, $t\,(68) = 2.30$, $p < .05$.

Specific cognitive ability predictors of flying average. A backward stepwise multiple regression analysis was performed using the flying average as the dependent variable and the RAFAAT specific cognitive ability measures (i.e., Table Reading, Digit Recall, Vigilance, and Visual Search) as the independent variables. The results of the backward stepwise regression analysis are presented at Table 13. Although the model fit improved with the removal of predictors at each subsequent step, none predictors of flying average were significant for any of the models. With an Adjusted R^2 value of .003, accounting for less than 1% of the variability in the flying average, Model 4 was the only model that accounted for any variability in the flying average. Therefore, none of the RAFAAT specific cognitive ability measures were significantly associated with the flying average.

Table 13

Backward Stepwise Regression of RAFAAT Cognitive Ability Sub-tests on Flying Average (n = 69)

Predictor	R	R^2	$Adj\ R^2$	ΔR^2	F	p	B	$SE\ B$	β	t	p	r	pr	sr	sr^2
Model 1	0.17	0.03	-0.03	0.03	0.47	.76									
Table							0.17	0.22	.11	0.75	.45	.12	.09	.09	.01
Recall							-0.05	0.18	-.03	-0.25	.80	-.01	-.03	-.03	.00
Vigilance							-0.05	0.08	-.08	-0.54	.59	.00	-.07	-.07	.00
Visual Search							0.18	0.20	.12	0.89	.38	.13	.11	.11	.01
Model 2	0.17	0.03	-0.02	-0.00	0.62	.61									
Table							0.17	0.22	.11	0.77	.44	.12	.10	.09	.01
Vigilance							-0.05	0.08	-.08	-0.57	.57	.00	-.07	-.07	.00
Visual Search							0.17	0.19	.12	0.87	.39	.13	.11	.11	.01
Model 3	0.15	0.02	-0.00	-0.00	0.77	.47									
Table							0.13	0.21	.08	0.62	.54	.12	.08	.08	.01
Visual Search							0.15	0.19	.10	0.78	.44	.13	.10	.10	.01
Model 4	0.13	0.02	0.00	-0.00	1.18	.28									
Visual Search							0.19	0.17	.13	1.09	.28	.13	.13	.13	.02

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation

Summary of specific cognitive ability predictors. Hypothesis 3 predicted that better performance on the specific cognitive ability variables, as measured by the RAFAAT, would be related to better performance on the pilot training outcomes. However, for Hypothesis 3, the only predictor demonstrating significant relations with any of the pilot training performance outcomes was the positive relation between Table Reading (i.e., information processing) and the academic grade. There were no other significant relations between any of the specific cognitive ability variables measured by the RAFAAT and any of the pilot training outcome variables. Therefore, Hypothesis 3 is only partially supported.

Executive functioning predictors of flying training performance. Hypothesis 4 proposed that higher levels of executive functioning (i.e., working memory, attention control, associated learning, and verbal organization), as measured by the ExamCorp battery, are associated with higher levels of pilot training performance. The descriptive statistics and correlation analysis are presented in Table 14. The ExamCorp executive functioning predictor measures Go/No-Go, Response Inhibition, and Sustained Attention were positively associated with the Academic Grade. Additionally, the ExamCorp Non Spatial Association sub-test was positively associated with the Flying Average and Good/Poor Performance, whereas the ExamCorp Random Letter Span sub-test was negatively associated with Flying Average. Based on these results, Hypothesis 3 has been partially supported in that higher levels of Non Spatial Association (associated learning) are associated with a higher flying average. The results also indicate that lower performance on Random Letter Span, representing lower working memory performance, is associated with a higher flying average.

Table 14

Descriptive Statistics and Correlations for Executive Functioning (Measured by the ExamCorp Battery) and Training Performance Outcome Variables

Variables	*n*	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12
1. Non Spatial Association [a]	73	0.07	0.75	-											
2. Spatial Association [a]	73	0.33	0.88	.14	-										
3. Go/No-Go [a]	73	0.31	0.81	-.15	.02	-									
4. Random Letter Span [a]	73	0.20	0.99	.19	.32**	.02	-								
5. Random Object Span [a]	73	0.03	0.92	.26*	.21*	-.09	.29**	-							
6. Recency Judgement [a]	73	0.05	0.96	-.02	.14	-.03	.02	-.06	-						
7. Response Inhibition [a]	73	0.13	0.95	.29**	.14	-.08	.31**	.34**	.03	-					
8. Sustained Attention [a]	73	0.52	0.64	.05	.01	.16	.16	.15	-.09	.26*	-				
9. Word Fluency [a]	73	-0.03	0.89	.32*	.17	.05	.29**	.25*	.26*	.40**	.21*	-			
10. Academic Grade [a]	75	96.52	1.99	.17	.03	-.29**	.00	.07	.12	.26*	.28**	.19	-		
11. Flying Average [b]	75	67.48	15.39	.18*	-.02	-.03	-.13	-.12	.04	-.02	.00	.12	.23**	-	
12. Good/Poor Performance [cd]	75	0.80	0.40	.21*	.04	.03	-.15	-.16	.09	.05	.06	.06	.26*	.76**	-

[a] Pearson correlations (1-tailed)

[b] Kendall's tau correlations (1-tailed)

[c] Biserial correlations (1-tailed)

[d] Good/Poor Performance: 0 = poor; 1 = good

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Executive functioning predictors of good/poor performance. An independent samples t-test was conducted in order to compare the means of good and poor training performance on all the executive functioning variables (see Table 15). For Non-Spatial Association there was a significant difference in the scores for good performance (M = .146, SD = .785) and poor performance (M = -.250, SD = .459); t (63) = -2.48, $p < .05$. Therefore, higher scores on the Non-Spatial Association test were associated with good performance, whereas lower scores on the Non-Spatial Association test were associated with poor performance. For the remaining executive functioning predictors there were no significant differences in the scores between the means of good and poor performance. However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 15
Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for Executive Functioning Measures (ExamCorp) for Good (n = 59) and Poor (n = 14) Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
Non Spatial Association[a]	.146 (.785)	-.250 (.459)	-2.48*	33.68	.02	-0.720	-0.071
Spatial Association[a]	.341 (.761)	.257 (1.299)	-0.23	15.18	.82	-0.852	0.685
Go/No-Go[b]	.325 (.818)	.257 (.822)	-0.28	71	.78	-0.554	0.417
Random Letter Span[b]	.124 (1.015)	.500 (.863)	1.28	71	.20	-0.210	0.962
Random Object Span[b]	-.048 (.949)	.333 (.773)	1.40	71	.17	-0.163	0.926
Recency Judgement[b]	.095 (.946)	-.128 (1.019)	-0.78	71	.44	-0.792	0.345
Response Inhibition[b]	.158 (.945)	.029 (.970)	-0.46	71	.65	-0.692	0.434
Sustained Attention[a]	.535 (.577)	.443 (.892)	-0.37	15.63	.72	-0.624	0.438
Word Fluency[b]	-.002 (.890)	-.136 (.899)	-0.51	71	.61	-0.662	0.394

[a] Levene's test significant; therefore equal variances not assumed. The assumption of homogeneity of variance was violated, and the results may be inaccurate.
[b] Levene's test not significant; therefore equal variances assumed. The assumption of homogeneity of variance was not violated.

Executive functioning predictors of academic grade. To determine the associations of the ExamCorp executive functioning predictors with the academic grade outcome variable, a backward stepwise multiple regression analysis was performed using the executive functioning predictors, specifically non-spatial association, spatial association, go/no-go, random letter span, random object span, recency judgement, response inhibition, sustained attention, and word fluency. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 16.

Table 16

Backward Stepwise Regression of ExamCorp Executive Functioning Measures and Academic Grade (n = 73)

Predictor	R	R^2	Adj R^2	ΔR^2	F	p	B	SE B	β	t	p	r	pr	sr	sr^2
Model 1	0.50	0.25	0.15	0.25	2.35	.02									
Non Spatial Association							0.21	0.33	.08	0.63	.53	.17	.08	.07	.00
Spatial Association							0.05	0.27	.02	0.17	.86	.03	.02	.02	.00
Go/No-Go							-0.78	0.28	-.32	-2.77	.01	-.29	-.33	-.30	.09
Random Letter Span							-0.24	0.25	-.12	-0.94	.35	.00	-.12	-.10	.01
Random Object Span							-0.13	0.27	-.06	-0.50	.62	.07	-.06	-.05	.00
Recency Judgement							0.23	0.25	.11	0.95	.35	.12	.12	.10	.01
Response Inhibition							0.33	0.27	.16	1.21	.23	.26	.15	.13	.02
Sustained Attention							0.96	0.37	.31	2.62	.01	.28	.31	.29	.08
Word Fluency							0.16	0.30	.07	0.52	.61	.19	.07	.06	.00
Model 2	0.50	0.25	0.16	0.00	2.68	.01									
Non Spatial Association							0.21	0.32	.08	0.65	.52	.17	.08	.07	.00
Go/No-Go							-0.78	0.28	-.31	-2.79	.01	-.29	-.33	-.30	.09
Random Letter Span							-0.23	0.24	-.11	-0.94	.35	.00	-.12	-.10	.01
Random Object Span							-0.13	0.26	-.06	-0.49	.63	.07	-.06	-.05	.00
Recency Judgement							0.24	0.24	.11	0.99	.33	.12	.12	.11	.01
Response Inhibition							0.33	0.27	.16	1.22	.23	.26	.15	.13	.02
Sustained Attention							0.96	0.36	.31	2.63	.01	.28	.31	.29	.08
Word Fluency							0.16	0.30	.07	0.52	.60	.19	.07	.06	.00
Model 3	0.50	0.25	0.17	0.00	3.07	.00**									
Non Spatial Association							0.19	0.32	.07	0.60	.55	.17	.07	.06	.00

Go/No-Go							-0.77	0.28	-.31	-2.78	.01	-.29	-.33	-.30	.09
Random Letter Span							-0.25	0.24	-.12	-1.05	.30	.00	-.13	-.11	.01
Recency Judgement							0.25	0.24	.12	1.05	.30	.12	.13	.11	.01
Response Inhibition							0.31	0.27	.15	1.16	.25	.26	.14	.13	.02
Sustained Attention							0.95	0.36	.30	2.63	.01	.28	.31	.28	.08
Word Fluency							0.14	0.29	.06	0.48	.63	.19	.06	.05	.00
Model 4	0.50	0.25	0.18	0.00	3.59	.00**									
Non Spatial Association							0.23	0.31	.09	0.75	.46	.17	.09	.08	.01
Go/No-Go							-0.75	0.27	-.30	-2.76	.01	-.29	-.32	-.30	.09
Random Letter Span							-0.23	0.23	-.11	-0.99	.33	.00	-.12	-.11	.01
Recency Judgement							0.28	0.23	.14	1.26	.21	.12	.15	.13	.02
Response Inhibition							0.34	0.26	.16	1.34	.19	.26	.16	.14	.02
Sustained Attention							0.97	0.36	.31	2.73	.01	.28	.32	.29	.09
Model 5	0.49	0.24	0.18	-0.01	4.22	.00**									
Go/No-Go							-0.78	0.27	-.32	-2.89	.01	-.29	-.33	-.31	.10
Random Letter Span							-0.21	0.23	-.10	-0.91	.37	.00	-.11	-.10	.01
Recency Judgement							0.28	0.23	.13	1.23	.22	.12	.15	.13	.02
Response Inhibition							0.39	0.25	.18	1.57	.12	.26	.19	.17	.03
Sustained Attention							0.97	0.36	.31	2.73	.01	.28	.32	.29	.09
Model 6	0.48	0.23	0.19	-0.01	5.08	.00***									
Go/No-Go							-0.79	0.27	-.32	-2.93	.01	-.29	-.34	-.31	.10
Recency Judgement							0.28	0.23	.13	1.22	.23	.12	.15	.13	.02
Response Inhibition							0.33	0.24	.15	1.37	.18	.26	.16	.15	.02
Sustained Attention							0.94	0.35	.30	2.67	.01	.28	.31	.28	.08
Model 7	0.46	0.21	0.18	-0.02	6.23	.00***									

Go/No-Go							-0.79	0.27	-.32	-2.93	.01	-.29	-.33	-.31	.10
Response Inhibition							0.34	0.24	.16	1.44	.16	.26	.17	.15	.02
Sustained Attention							0.90	0.35	.29	2.55	.01	.28	.29	.27	.08
Model 8	0.44	0.19	0.17	-0.02	8.19	.00***									
Go/No-Go							-0.84	0.27	-.34	-3.12	.00	-.29	-.35	-.34	.11
Sustained Attention							1.04	0.34	.33	3.05	.00	.28	.34	.33	.10

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

All of the models demonstrated a significant association with the academic grade, with the strength of the model fit increasing as the weakest predictors were removed, until the most parsimonious model was reached at Model 8. The regression with all the variables in the model (Model 1) was significant, $R = .50$, $R^2 = .25$, $F(9, 63) = 2.35$, $p < .05$. The Adjusted R^2 value of .145 indicated that 14.5% of the variability in academic grade was predicted by a model with all the ExamCorp executive functioning measures in the model. However, in the initial model, only two of the nine variables exhibited regression coefficients that were significantly different from zero; Go/No-Go, $\beta = -.32$, t $(72) = -2.77$, $p < .01$, and Sustained Attention, $\beta = .31$, $t\,(72) = 2.62$, $p < .05$. Therefore, lower scores on the Go/No-Go measure were associated with higher academic grades, whereas higher scores on Sustained Attention were associated with higher academic grades. There were no statistically significant linear relations detected between the mean of the academic grade and any of the remaining predictors.

The Go/No-Go and Sustained Attention variables were the only predictors to exhibit significant associations with the academic grade outcome variable at any of the models in the analysis, and maintained significance through all models. The final and optimal model (Model 8), which only included Go/No-Go and Sustained Attention, was significant; $R = .44$, $R^2 = .19$, $F(2, 70) = 8.19$, $p < .001$. The Adjusted R^2 value of .167 indicated that almost 17% of the variability in the academic grade was predicted by a model with only the Go/No-Go and Sustained Attention variables in the model. Both variables exhibited regression coefficients that were significantly different from zero; Go/No-Go, $\beta = -.84$, $t\,(72) = -3.12$, $p < .01$, and Sustained Attention, $\beta = .33$, $t\,(72) = 3.05$, $p < .01$. Therefore, lower scores on the Go/No-Go measure were associated with higher academic grades, whereas higher scores on Sustained Attention were associated with higher academic grades. These results somewhat support the hypothesis that higher levels of executive functioning are associated with higher academic grades.

Executive functioning predictors of flying average. To determine the associations of the ExamCorp executive functioning predictors with the flying average outcome variable, a backward stepwise multiple regression analysis was performed using the executive functioning predictors, specifically non-spatial association, spatial association, go/no-go, random letter span, random object span, recency judgement, response inhibition, sustained attention, and word fluency. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 17.

The strength of the model fit increased as the weakest predictors were removed, until the most parsimonious model was reached at Model 8. The regression with all the variables in the model (Model 1) was not significant, $R = .41$, $R^2 = .17$, $F(9, 63) = 1.43$, $p = .19$. The Adjusted R^2 value of .05 indicated that only 5% of the variability in academic grade was predicted by a model with all the ExamCorp executive functioning measures in the model. However, in the initial model, only two of the nine variables exhibited regression coefficients that were significantly different from zero; Non-Spatial Association, $\beta = .29$, $t\,(72) = 2.31$, $p < .05$, and Random Letter Span, $\beta = -.29$, $t\,(72) = -2.24$, $p < .05$. There were no statistically significant linear relations detected between the mean of the academic grade and any of the remaining predictors.

Table 17

Backward Stepwise Regression of ExamCorp Executive Functioning Measures and Flying Average (n = 73)

Predictor	*R*	R^2	Adj R^2	ΔR^2	*F*	*p*	*B*	*SE B*	*β*	*t*	*p*	*r*	*pr*	*sr*	sr^2
Model 1	0.41	0.17	0.05	0.17	1.43	.19									
Non Spatial Association							5.87	2.54	.29	2.31	.02	.23	.28	.27	.07
Spatial Association							0.59	2.11	.04	0.28	.78	-.04	.04	.03	.00
Go/No-Go							0.80	2.19	.04	0.37	.72	.02	.05	.04	.00
Random Letter Span							-4.38	1.95	-.29	-2.24	.03	-.23	-.27	-.26	.07
Random Object Span							-2.94	2.08	-.18	-1.41	.16	-.15	-.18	-.16	.03
Recency Judgement							0.22	1.91	.01	0.12	.91	.03	.01	.01	.00
Response Inhibition							0.84	2.14	.05	0.39	.70	.03	.05	.05	.00
Sustained Attention							1.59	2.86	.07	0.56	.58	.04	.07	.06	.00
Word Fluency							0.93	2.34	.06	0.40	.69	.07	.05	.05	.00
Model 2	0.41	0.17	0.07	-0.00	1.63	.13									
Non Spatial Association							5.84	2.51	.29	2.33	.02	.23	.28	.27	.07
Spatial Association							0.62	2.07	.04	0.30	.77	-.04	.04	.03	.00
Go/No-Go							0.79	2.17	.04	0.36	.72	.02	.05	.04	.00
Random Letter Span							-4.40	1.94	-.29	-2.27	.03	-.23	-.27	-.26	.07
Random Object Span							-2.96	2.05	-.18	-1.45	.15	-.15	-.18	-.17	.03
Response Inhibition							0.83	2.12	.05	0.39	.70	.03	.05	.05	.00
Sustained Attention							1.55	2.82	.07	0.55	.59	.04	.07	.06	.00
Word Fluency							1.01	2.21	.06	0.46	.65	.07	.06	.05	.00
Model 3	0.41	0.17	0.08	-0.00	1.88	.09									
Non Spatial Association							5.87	2.49	.30	2.36	.02	.23	.28	.27	.07

Go/No-Go							0.81	2.15	.04	0.38	.71	.02	.05	.04	.00
Random Letter Span							-4.24	1.86	-.28	-2.29	.03	-.23	-.27	-.26	.07
Random Object Span							-2.90	2.02	-.18	-1.43	.16	-.15	-.18	-.16	.03
Response Inhibition							0.84	2.11	.05	0.40	.69	.03	.05	.05	.00
Sustained Attention							1.49	2.79	.06	0.53	.60	.04	.07	.06	.00
Word Fluency							1.05	2.19	.06	0.48	.63	.07	.06	.05	.00
Model 4	0.41	0.17	0.09	-0.00	2.20	.05									
Non Spatial Association							5.74	2.45	.29	2.35	.02	.23	.28	.26	.07
Random Letter Span							-4.21	1.84	-.28	-2.29	.03	-.23	-.27	-.26	.07
Random Object Span							-2.96	2.00	-.18	-1.48	.15	-.15	-.18	-.17	.03
Response Inhibition							0.76	2.08	.05	0.36	.72	.03	.05	.04	.00
Sustained Attention							1.68	2.73	.07	0.61	.54	.04	.08	.07	.00
Word Fluency							1.13	2.17	.07	0.52	.60	.07	.06	.06	.00
Model 5	0.41	0.17	0.10	-0.00	2.65	.03									
Non Spatial Association							5.87	2.41	.30	2.44	.02	.23	.29	.27	.07
Random Letter Span							-4.11	1.81	-.27	-2.27	.03	-.23	-.27	-.25	.06
Random Object Span							-2.82	1.95	-.18	-1.44	.15	-.15	-.17	-.16	.03
Sustained Attention							1.85	2.67	.08	0.69	.49	.04	.08	.08	.00
Word Fluency							1.32	2.09	.08	0.63	.53	.07	.08	.07	.00
Model 6	0.40	0.16	0.11	-0.01	3.24	.02									
Non Spatial Association							6.26	2.31	.31	2.71	.01	.23	.31	.30	.09
Random Letter Span							-3.90	1.77	-.26	-2.20	.03	-.23	-.26	-.25	.06
Random Object Span							-2.68	1.93	-.17	-1.39	.17	-.15	-.17	-.15	.02
Sustained Attention							2.12	2.62	.09	0.81	.42	.04	.10	.09	.00
Model 7	0.39	0.15	0.12	-0.01	4.12	.01									

Non Spatial Association							6.26	2.31	.31	2.71	.01	.23	.31	.30	.09
Random Letter Span							-3.72	1.75	-.25	-2.13	.04	-.23	-.25	-.24	.06
Random Object Span							-2.51	1.92	-.16	-1.31	.20	-.15	-.16	-.15	.02
Model 8	0.36	0.13	0.11	-0.02	5.26	.00**									
Non Spatial Association							5.59	2.26	.28	2.47	.02	.23	.28	.28	.08
Random Letter Span							-4.31	1.70	-.29	-2.53	.01	-.23	-.29	-.28	.08

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation
* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

The Non-Spatial Association and Random Letter Span variables were the only predictors to exhibit significant associations with the flying average outcome variable for all of the models in the analysis, and maintained significance through all models. The final and optimal model (Model 8), which only Non-Spatial Association and Random Letter Span, was significant; $R = .36$, $R^2 = .13$, $F(2, 70) = 5.26$, $p < .01$. The Adjusted R^2 value of .106 indicated that almost 11% of the variability in the flying average was predicted by a model with only the Non-Spatial Association and Random Letter Span variables in the model. Both variables exhibited regression coefficients that were significantly different from zero; Non-Spatial Association, $\beta = .28$, $t\,(72) = 2.47$, $p < .05$, and Random Letter Span, $\beta = -.29$, $t\,(72) = -2.53$, $p < .05$. Therefore, higher scores on Non-Spatial Association were associated with higher flying average outcomes whereas lower scores on the Random Letter Span measure were associated with higher flying average outcomes. These results partially support the hypothesis, in that higher levels of the executive functioning domain associated learning was associated with a higher flying average.

Summary of the executive functioning predictors. Hypothesis 4 predicted that higher levels of the executive functioning variables would be associated with higher performance on pilot training outcomes. The results indicated that higher levels of non-spatial association (i.e., associated learning) were associated with good performance. Additionally, worse performance on the Go/No-Go measure (i.e., attention control) and better performance on sustained attention were related to higher academic grades. The results of the correlation analysis for the executive functioning predictors and flying average, which used Kendall's tau b to account for the skew of the flying average, a significant correlation was only found between Non-Spatial Association and the flying average (Kendall's $\tau_b = .18$, $p < .05$). However, the backward regression indicated that two executive functioning variables also displayed significant relations with the flying average; non-spatial association (i.e., associated learning) was positively related to the flying average whereas random letter span (i.e., working memory) was negatively related to the flying average. Therefore, hypothesis 4 is partially supported; the negative associations of the significant predictors were unexpected.

Incremental validity of executive functioning measures. Hypothesis 5 proposed that executive functioning, as measured by the ExamCorp battery, accounted for more unique variance in pilot training performance than general cognitive intelligence (as measured with the CFAT), fluid intelligence (as measured by RAFAAT), and specific cognitive abilities (as measured by RAFAAT). The incremental validity analyses for the two continuous outcome variables, academic grade and flying average, were considered separately.

Incremental validity for the academic grade outcome. The investigation of the incremental validity of the executive functioning predictors over the intelligence and specific cognitive ability predictors commenced using the academic grade outcome variable. The results of previous hypothesis testing indicated that CFAT Problem Solving, RAFAAT Table Reading, and two of the ExamCorp executive functioning variables, Go/No-go and Sustained Attention, were significantly associated with the academic grade outcome variable. Therefore, only these four variables were considered for this analysis. The descriptive statistics and bivariate correlations are presented at Table 18.

Table 18

Descriptive Statistics and Correlations[a] *of CFAT Problem Solving, RAFAAT Table Reading, ExamCorp Executive Functioning Variables (i.e., Go/No-Go, Sustained Attention), and Academic Grade*

	n	*M*	*SD*	1	2	3	4	5
1. CFAT Problem Solving	75	25.35	3.64	-				
2. Table Reading	75	65.72	9.89	.37**	-			
3. Go/No-Go	73	0.325	0.80	-.05	-.12	-		
4. Sustained Attention	73	0.526	0.65	.24*	.24*	.12	-	
5. Academic Average	75	96.60	2.04	.39**	.29**	-.30**	.30**	-

Note. CFAT = Canadian Forces Aptitude Test; RAFAAT = Royal Air Force Aircrew Aptitude Test

[a] Pearson correlations (1-tailed)

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

To determine whether the executive functioning variables provided incremental validity over the CFAT and RAFAAT variables in predicting pilot training performance, specifically the academic grade, a hierarchical regression analysis was performed. As suggested by Cohen, Cohen, West and Aiken (2003), the order of entry of the variables into the hierarchical regression was determined *a priori*. CFAT variables were entered as a set into step 1 as these measures are currently used in the Canadian Forces selection system, and the goal of the hypothesis is to determine measures that predict over and above the current measures. Additionally, previous research has recognized general cognitive ability as a significant predictor of general academic and job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998) as well as pilot performance (Martinuusen & Hunter, 2010). The RAFAAT Table Reading variable was added in step 2 as the RAFAAT was also added to the Canadian Forces pilot selection system. The two executive functioning variables, Go/No-Go and Sustained Attention, were added as a set in step 3, to examine the incremental validity of executive functioning over the other cognitive measures in the prediction of the academic grade.

Incremental validity was assessed by examination of the results of a hierarchical regression analysis, which are displayed in Table 19. In the first step of the three-step hierarchical regression the CFAT Problem Solving was regressed onto the academic grade outcome variable. CFAT Problem Solving contributed significantly to the prediction of the academic grade in the first step $R^2 = .15$, $\Delta R^2 = .15$, $F(1, 66) = 11.66$, $p < .001$. In step one, CFAT Problem Solving accounted for almost 14% of the variance in the academic grade (Adjusted R2 = .137). The addition of the RAFAAT Table Reading variable to CFAT Problem Solving in step 2, was also significant, $R^2 = .18$, $\Delta R^2 = .02$, $F(2, 65) = 6.90$, $p < .01$. The R^2 change indicates that an additional 2.5% of the variability in the academic grade was predicted by the combination of the CFAT Problem Solving and RAFAAT Table Reading variables. However, despite the significant association found between the RAFAAT Table Reading and the academic grade variables in Hypothesis 3, the relation became non-significant when added with the CFAT Problem Solving variable in step 2; $\beta = .17$, $t(65) = 1.40$, $p = .17$.

In step 3, the addition of the executive functioning variables, Go/No-Go and Sustained Attention, added a significant amount of incremental validity above the other cognitive measures, $R^2 = .30$, $\Delta R^2 = .12$, $F(4, 63) = 6.67$, $p < .001$. The R^2 change indicates that an additional 12% of the variability in the academic grade was predicted by the combination of CFAT Problem Solving, RAFAAT Table Reading and the two executive functioning variables. There was a significant positive relation between the academic grade outcome variable and both CFAT Problem Solving ($\beta = .28$, $t(65) = 2.44$, $p < .05$) and Sustained Attention ($\beta = .25$, $t(65) = 2.21$, $p < .05$), and a negative relation between the academic grade and the Go/No-Go ($\beta = -.30$, $t(65) = -2.78$, $p < .01$). However, the RAFAAT Table Reading variable remained non-significant, $\beta = .09$, $t(65)$

Table 19

Hierarchical Regression of Executive Functioning Variables (Go/No-Go and Sustained Attention) Over General Cognitive Ability Measures (CFAT Problem Solving), and Specific Cognitive Abilities (RAFAAT Table Reading) on Academic Grade (n = 70)

Predictor	R^2	Adj R^2	ΔR^2	*F (df)*	*p*	*B*	*SE B*	*β*	*t*	*p*	*r*	*pr*	*sr*	sr^2 *incr*
Step 1	0.15	0.14	0.15	11.66	.00***									
CFAT PS				(1, 66)		0.22	0.06	.39	3.42	.00***	.39	.39	.39	.15
Step 2	0.18	0.15	0.03	6.90	.00**									
CFAT PS				(2, 65)		0.18	0.07	.33	2.69	.00**	.39	.32	.30	.09
Table Reading						0.04	0.03	.17	1.40	.17	.29	.17	.16	.02
Step 3	0.30	0.25	0.12	6.67	.00***									
CFAT PS				(4, 63)		0.16	0.07	.28	2.42	.02	.39	.29	.26	.06
Table Reading						0.02	0.02	.09	0.78	.44	.29	.10	.08	.01
Go/No-go						-0.77	0.28	-.30	-2.78	.00**	-.30	-.33	-.29	.08
Sustained Attention						0.77	0.35	.25	2.21	.03	.30	.27	.23	.05

Note. CFAT PS = CFAT Problem Solving; . *r* = zero-order correlation; *pr* = partial correlation; *sr* = semipartial correlation; sr^2 = squared semipartial correlation; *incr* = incremental

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

= 0.78, p = .44. This supports the hypothesis in that the executive functioning variables added a significant increment in validity over the other cognitive measures in the prediction of the academic grade.

Incremental validity for the flying average outcome. Next, the incremental validity of the executive functioning predictors over the intelligence and specific cognitive ability predictors was investigated using the flying average outcome variable. The results of previous hypothesis testing indicated that CFAT Verbal Skills and Spatial Ability, and the ExamCorp executive functioning variables, Non-Spatial Association and Random Letter Span, were significantly associated with the flying average outcome variable. Therefore, only these four variables were considered for this analysis. The descriptive statistics and bivariate correlations are presented at Table 20.

Table 20
Descriptive Statistics and Correlations of CFAT (Verbal Skills and Spatial Ability), ExamCorp Executive Functioning Variables (Non-Spatial Association and Random Letter Span), and Flying Average

	n	M	SD	1	2	3	4	5
1. CFAT Verbal Skills[a]	75	11.23	2.20	-				
2. CFAT Spatial Ability[a]	75	12.17	2.09	.20*	-			
3. Non Spatial Association[a]	73	0.063	0.73	.25*	.26*	-		
4. Random Letter Span[a]	73	0.196	0.99	.03	-.03	.19	-	
5. Flying Average[b]	75	67.48	15.39	-.13	.15*	.18*	-.13	-

Note. CFAT = Canadian Forces Aptitude Test
[a] Pearson correlations (1-tailed)
[b] Kendall's tau correlations (1-tailed)
* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

To determine whether the executive functioning variables provide incremental validity over the CFAT variables in predicting pilot flying performance, a hierarchical regression analysis was performed. As suggested by Cohen, Cohen, West and Aiken (2003), the order of entry of the variables into the hierarchical regression was determined *a priori*. CFAT variables were entered as a set into step 1 as these measures are currently used in the Canadian Forces selection system, and the goal of the hypothesis is to determine measures that predict over and above the current measures. Additionally, previous research has recognized general cognitive ability as a significant predictor of general academic and job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998) as well as pilot performance (Martinuusen & Hunter, 2010). The two executive functioning variables, Non-Spatial Association and Random Letter Span, were added as a set in step 2, to examine the incremental validity of executive functioning in the prediction of the flying average.

The results of the hierarchical regression are displayed in Table 21. In the first step of the two-step hierarchical regression the CFAT variables (Verbal Skills and Spatial Ability) were regressed onto the flying average outcome variable. Both the CFAT measures contributed significantly to the prediction of the flying average in the first step $R^2 = .10$, $\Delta R^2 = .10$, $F(2, 70) = 3.87$, $p < .05$. The addition of the executive functioning measures in step 2 added a significant amount of incremental validity (i.e., a significant increment in R^2) above general cognitive ability, $R^2 = .18$, $\Delta R^2 = .13$, $F(4, 68) = 4.96$, $p < .001$. The R^2 change indicates that an additional 13% of the variability in flying average was predicted by the combination of the CFAT variables (Verbal Skills and Spatial Ability) and executive functioning variables (Non Spatial Association and Random Letter Span). These results support the hypothesis that executive functioning measures represent an increase in predictive power for pilot selection purposes.

There was a significant negative relation between CFAT Verbal Skills and flying average in both step 1 ($\beta = -.24$, $p < .05$) and step 2 ($\beta = -.30$, $p < .01$). In step 1, with

Table 21
Hierarchical Regression of Executive Functioning Variables (Non-Spatial Association and Random Letter Span) Over General Cognitive Ability Measures(CFAT Verbal Skills and Spatial Ability) on Flying Average

Predictor	R^2	Adj R^2	ΔR^2	*F (df)*	*p*	*B*	*SE B*	β	*t*	*p*	*r*	*pr*	*sr*	sr^2 *incr*
Step 1	0.10	0.07	0.10	3.87	.03									
CFAT Verbal Skills				(2, 70)		-1.65	0.78	-.24	-2.12	.04	-.20	-.25	-.24	.06
CFAT Spatial Ability						1.88	0.86	.25	2.19	.03	.20	.25	.26	.07
Step 2	0.23	0.18	0.13	4.97	.001									
CFAT Verbal Skills				(4, 68)		-2.00	0.75	-.30	-2.68	.00	-.19	-.31	-.29	.08
CFAT Spatial Ability						1.34	0.82	.18	1.62	.11	.20	.19	.17	.03
NSA						6.15	2.35	.30	2.61	.01	.22	.30	.28	.08
RLS						-4.17	1.63	-.28	-2.55	.01	-.23	-.30	-.27	.07

Note. CFAT = Canadian Forces Aptitude Test; NSA = Non-Spatial Association; RLS = Random Letter Span; *r* = zero-order correlation; *pr* = partial correlation; *sr* = semipartial correlation; sr^2 = squared semipartial correlation; *incr* = incremental only the CFAT variables in the equation, CFAT Spatial Ability was positively related to flying average (β = .25, p < .05). However, with the addition of the executive functioning variables in step 2, the relation between CFAT Spatial Ability and the flying average drops to non-significance (β = .18, p = .11). Also in step 2, there was also a significant positive relation for Non-Spatial Association and the flying average (β = .30, p < .01), and a significant negative relation between Random Letter Span and the flying average (β = -.28, p < .05).

Summary of the incremental validity of the executive functioning measures. Hypothesis 5 predicted that the executive functioning variables would demonstrate incremental validity over the other cognitive measures. The results of the hierarchical regressions support the hypothesis for both the academic grade and the flying average. Specifically, for the academic grade, both the Go/No-Go and Sustained Attention variables demonstrated incremental validity over the CFAT and Table Reading variables. Similarly, both the Non-Spatial Association and Random Letter Span variables demonstrated incremental validity over the CFAT variables.

Psychomotor ability predictors of flying training performance. Hypothesis 6 proposed that higher levels of psychomotor ability, as measured by two RAFAAT sub-tests (e.g., Sensory Motor Apparatus and Control of Velocity), are associated with higher levels of pilot flying performance. The results of the descriptive statistics and correlation analysis are presented at Table 22. The academic average outcome variable was not included in this analysis as it is not expected that psychomotor ability would be related to academic performance. Of interest is that only one of the RAFAAT measures of psychomotor ability, Sensory Motor Apparatus, demonstrated a significant association with the training flying performance outcomes.

Table 22
Descriptive Statistics and Correlations for Psychomotor Ability (Measured by Sensory Motor Apparatus and Control of Velocity) and Training Performance Outcome Variables

	n	M	SD	1	2	3	4
1. Sensory Motor Apparatus [a]	70	96.58	38.71	-			
2. Control of Velocity [a]	70	104.80	13.19	.37**	-		
3. Flying Average [b]	75	67.47	15.38	.17*	.00	-	
4. Good/Poor Performance [cd]	75	0.800	0.403	.37*	.10	1***	-

[a] Pearson Correlations (1-tailed)
[b] Biserial Correlations (1-tailed)
[c] Biserial Correlations (1-tailed)
[d] Good/Poor Performance: 0 = poor; 1 = good
* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Psychomotor predictors of good/poor performance. An independent samples t-test was conducted in order to compare the means of good and poor training performance on all the psychomotor ability variables (see Table 23). For Sensory Motor Apparatus there was a significant difference in the scores for good performance (M = 101.56, SD = 39.64) and poor performance (M = 76.71, SD = 27.83); t (68) = -2.21, $p < .05$. For the remaining psychomotor ability predictor, Control of Velocity, there was no significant difference in the scores between good and poor performance. However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 23
Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for Psychomotor Ability Measures (RAFAAT) for Good (N = 56) and Poor (N = 14) Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
Sensory Motor Apparatus	101.56 (39.64)	76.71 (27.83)	-2.21	68	.03	-47.30	2.38
Control of Velocity	105.25 (12.98)	103.00 (14.37)	-0.57	68	.57	-10.15	-5.65

Note. Levene's test was not significant for all variables; results presented are based on equal variance assumed. The assumption of homogeneity of variance was not violated.

Psychomotor predictors of flying average. Due to the limited number of variables involved in this analysis, a standard multiple regression analysis was performed between the two psychomotor ability predictors and the flying average outcome variable. The results of the regression analysis are presented at Table 24. The regression with all the psychomotor ability variables in the model was significant, $R = .30$, $R^2 = .09$, $F(2, 67) = 3.32$, $p < .05$. The Adjusted R^2 value, .06, indicates that 6 percent of the variability in flying average is predicted by a model with all the psychomotor ability variables. However, only the Sensory Motor Ability measure was significantly related to the flying average, $\beta = .31$, $t\,(67) = 2.50$, $p < .01$. A statistically significant linear relation was not detected between the mean of flying average on the Control of Velocity measure $\beta = -.04$, $t\,(67) = -.35$, $p = .73$.

Table 24
Standard Regression of Psychomotor Ability Measures (RAFAAT tests Sensory Motor Apparatus and Control of Velocity) on Flying Average (n = 70)

Predictor	*B*	*SE B*	*β*	*t*	*df*	*p*	*r*	*pr*	*sr*	sr^2
Sensory Motor Apparatus	0.13	0.05	.31	2.50	67	.01	.30	.29	.29	.08
Control of Velocity	-0.05	0.15	-.04	-0.35	67	.73	.07	-.04	-.04	.00

R =0 .30
R^2 = 0.09
Adjusted R^2 = 0.06
$F(df)$ = 3.32 (2, 67)
p = .04

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation.

Summary of psychomotor ability predictors. Hypothesis 6 predicted that the psychomotor ability variables, as measured by the RAFAAT, would be significantly related to flying training performance. Of the two psychomotor ability variables, only the Sensory Motor Apparatus, the measure of hand-eye-foot coordination was significant; Sensory Motor Apparatus demonstrated a significant positive relation with both good performance and the flying average. Therefore, Hypothesis 6 was partially supported.

Incremental validity of psychomotor ability measures. Hypothesis 7 builds on Hypothesis 6, predicting that psychomotor ability as measured by RAFAAT will account for unique variance in pilot training performance over and above CAPSS performance. As only Sensory Motor Apparatus demonstrated a significant relation with the training performance variables in the previous hypothesis, the Control of Velocity measure was not included in this analysis. Table 25 provides the descriptive statistics and correlation analysis for this hypothesis. Although CAPSS demonstrated a significant positive association with both the academic grade and flying average, it did not show a significant association with either the Pass/fail Grade or Good/poor Performance. Additionally, despite the significant correlation between CAPSS and flying average, Kendall's $\tau_b = .15$, $p < .05$, the bivariate relation only accounts for 2% of the variance.

Table 25

Descriptive Statistics and Correlations of Sensory Motor Apparatus, CAPPS, and Training Performance Outcome Variables

	n	*M*	*SD*	1	2	3	4	5
1. CAPSS [a]	75	0.83	0.12	-				
2. Sensory Motor Apparatus [a]	70	91.41	38.71	.20*	-			
3. Academic Grade [a]	75	96.52	1.99	.20*	.20*	-		
4. Flying Average [b]	75	67.48	15.39	.15*	.17*	.23**	-	
5. Good/Poor Performance [cd]	75	0.80	0.40	.26	-.37*	.37*	1***	-

Note. CAPSS = Canadian Automated Pilot Selection System

[a] Pearson correlations (1-tailed)

[b] Kendall's tau correlations (1-tailed)

[c] Biserial correlations (1-tailed)

[d] Good/Poor Performance: 0 = poor; 1 = good

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

An independent samples t-test was conducted in order to compare good and poor training performance on the CAPSS variable, as all other variables have been investigated in previous hypotheses (see Table 26). For the CAPSS variable there were no significant differences in the scores between good ($M = .84$, $SD = .13$) and poor ($M = .78$, $SD = .10$) performance; $t(73) = -1.56$ $p = .12$. However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 26

Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for CAPSS Performance for Good (N =60) and Poor (N =15) Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
CAPSS	0.839 (0.128)	0.783 (0.102)	-1.56	73	.12	-0.127	0.015

Note 1. CAPSS = Canadian Automated Pilot Selection System

Note 2. Levene's test was not significant; results presented are based on equal variance assumed. The assumption of homogeneity of variance was not violated.

Incremental validity of predictors for flying average. To test whether the Sensory Motor Apparatus measure provides incremental validity over CAPSS in the prediction of flying training performance, a hierarchical multiple regression analysis was performed. CAPSS was entered in the first step of the two-step hierarchical regression as this measure is currently used in the Canadian Forces selection system, and the goal is to determine measures that predict over and above the current selection measures. As the CAPSS measure is thought to represent a measure of psychomotor ability (Darr, 2010b), the CFAT variables were not included in this analysis. Therefore, the new measure Sensory Motor Apparatus was added in the second step.

In the first step of the two-step hierarchical regression the CAPSS variable was regressed onto the flying average outcome variable (see Table 27). The CAPSS variable contributed to the prediction of the flying average in the first step $R^2 = .06$, $\Delta R^2 = .06$, $F(1, 68) = 4.01$, $p < .05$. The addition of the Sensory Motor Apparatus variable in step 2 added a significant amount of incremental validity (i.e., a significant increment in R^2) above the CAPSS variable, $R^2 = .12$, $\Delta R^2 = .07$, $F(2, 67) = 4.60$, $p < .05$. The addition of Sensory Motor Apparatus resulted in an increment of almost 7% of variability to the prediction of the flying average over and above the CAPSS measure. However, though there was a significant positive relation between CAPSS performance and the flying average in step 1 ($\beta = .24$, $t\,(68) = 2.00$, $p < .05$), that association becomes non-significant with the addition of Sensory Motor Apparatus in the equation in step 2 ($\beta = .18$, $t\,(67) = 1.57$, $p = .12$). The Sensory Motor Apparatus variable was significantly positively related to flying average ($\beta = .26$, $t\,(67) = 2.23$, $p < .05$).

Table 27
Hierarchical Regression of CAPSS and Sensory Motor Apparatus on Flying Average

Predictor	R^2	*Adj* R^2	ΔR^2	*F*	*p*	*B*	*SE B*	β	*t*	*p*	*r*	*pr*	*sr*	sr^2 *incr*
Step 1	0.06	0.04	0.06	4.01	.05									
CAPSS						28.40	14.17	.24	2.00	.05	.24	.24	.24	.06
Step 2	0.12	0.10	0.07	4.06	.01									
CAPSS						22.08	14.07	.18	1.57	.12	.24	.19	.18	.03
SMA						0.11	0.05	.26	2.23	.03	.30	.26	.26	.07

Note. CAPSS = Canadian Automated Pilot Selection System; SMA = Sensory Motor Apparatus; *r* = zero-order correlation; *pr* = partial correlation; *sr* = semipartial correlation; sr^2 = squared semipartial correlation; *incr* = incremental

Summary of incremental validity of psychomotor ability measures. Hypothesis 7 predicted that psychomotor ability, as measured by the RAFAAT Sensory Motor Apparatus, would demonstrate incremental validity over CAPSS variable. The results indicate that this hypothesis was partially supported, in that CAPSS dropped to non-significance with the addition of the SMA variable.

Personality predictors of flying training performance. Hypothesis 8 proposed that higher levels of the big five personality factors (measured by the ExamCorp battery) were associated with higher training performance outcomes. The results of the descriptive statistics and bivariate correlations are presented at Table 28. Of the five personality factors, conscientiousness demonstrated a significant positive association with two training performance outcomes (academic grade and flying average), whereas extraversion demonstrated a significant negative association with the academic grade. Therefore, more conscientious students performed better both academically and on the flying phase (Flying Average) than less conscientious students, and students who had higher levels of extraversion did not perform as well academically as those with lower levels of extraversion.

Table 28
Descriptive Statistics and Correlations of Personality and Training Performance Variables

	n	*M*	*SD*	1	2	3	4	5	6	7	8
1. Agreeableness [a]	73	0.09	0.91	-							
2. Conscientiousness [a]	73	0.54	0.80	.45**	-						
3. Emotional Stability [a]	73	0.59	0.78	.62**	.54**	-					
4. Extraversion [a]	73	0.21	1.01	.36**	.35**	.59**	-				
5. Intellect/Culture [a]	73	0.12	0.84	.56**	.48**	.53**	.57**	-			
6. Academic Grade [a]	75	96.52	1.99	.01	.29**	.04	-.22*	.02	-		
7. Flying Average [b]	75	67.48	15.39	.04	.18*	.08	-.01	.05	.30**	-	
8. Good/Poor Performance [cd]	75	0.80	0.40	.04	.19	-.03	-.10	.05	.37**	1***	-

[a] Pearson correlations (1-tailed)
[b] Kendall's tau correlations (1-tailed)
[c] Biserial correlations (1-tailed)
[d] Good/Poor Performance: 0 = poor; 1 = good
* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Personality predictors of good/poor performance. An independent samples t-test was conducted in order to compare good and poor training performance on all the personality factors (see Table 29). There were no significant differences in the scores between good and poor performance for any of the personality measures. However, the severity of the split of the sample size (i.e., 80% to 20%) for this variable means the results may not be accurate.

Table 29

Independent Samples t-test Results, with Means and Standard Deviations (in Parentheses), for Personality Factors for Good (N = 59) and Poor (N = 14) Performance

	Performance					95% CI	
	Good	Poor	*t*	*df*	*p*	Lower	Upper
Agreeableness	0.105 (0.928)	0.043 (0.865)	-0.23	71	.82	-0.606	0.481
Conscientiousness	0.590 (0.781)	0.321 (0.894)	-1.12	71	.26	-0.744	0.208
Emotional Stability	0.578 (0.803)	0.621 (0.710)	0.19	71	.85	-0.423	0.510
Extraversion	0.178 (1.02)	0.343 (1.01)	0.54	71	.59	-0.439	0.769
Intellect/Culture	0.137 (0.833)	0.057 (0.876)	-0.32	71	.75	-0.589	0.418

Note. Levene's test was not significant for all variables; results presented are based on equal variance assumed

Personality predictors of academic grade. A backward stepwise regression analysis was performed using the personality predictors to determine the associations with the academic grade outcome variable. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003). The results of the backward stepwise regression analysis are presented at Table 30.

All of the models involving the personality predictors demonstrated a significant

Table 30

Backward Stepwise Regression of Personality Variables and Academic Grade (n = 73)

Predictor	R	R^2	$Adj\ R^2$	ΔR^2	F	p	B	$SE\ B$	β	t	p	r	pr	sr	sr^2
Model 1	0.22	0.16	0.22	3.74	3.74	.005									
Agreeableness							-0.31	0.33	-.14	-0.94	.35	.01	-.11	-.10	.01
Conscientiousness							0.99	0.34	.39	2.94	.01	.29	.34	.32	.10
Emotional Stability							0.34	0.43	.13	0.79	.43	.04	.10	.09	.01
Extraversion							-0.88	0.29	-.44	-3.02	.00	-.22	-.35	-.33	.10
Intellect/Culture							0.22	0.37	.09	0.60	.55	.02	.07	.07	.00
Model 2	0.21	0.17	0.00	0.36	4.62	.002									
Agreeableness							-0.24	0.31	-.11	-0.78	.44	.01	-.09	-.08	.01
Conscientiousness							1.03	0.33	.41	3.17	.00	.29	.36	.34	.12
Emotional Stability							0.33	0.43	.13	0.78	.44	.04	.09	.08	.01
Extraversion							-0.81	0.27	-.41	-3.05	.00	-.22	-.35	-.33	.11
Model 3	0.21	0.17	-0.01	0.60	6.00	.001									
Conscientiousness							0.99	0.32	.39	3.09	.00	.29	.35	.33	.11
Emotional Stability							0.18	0.38	.07	0.48	.63	.04	.06	.05	.00
Extraversion							-0.80	0.26	-.41	-3.05	.00	-.22	-.34	-.33	.11
Model 4	0.20	0.18	0.00	0.23	8.98	.00***									
Conscientiousness							1.05	0.29	.42	3.68	0.00	.29	.40	.39	.15
Extraversion							-0.74	0.23	-.37	-3.27	0.00	-.22	-.36	-.35	.12

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

association with the academic grade. The regression with all the personality variables in the model (Model 1) was significant, $R = .47$, $R^2 = .22$, $F(5, 67) = 3.74$, $p < .01$. The Adjusted R^2 value of .16 indicated that 16% of the variability in academic grade was predicted by a model with all the personality variables in the model. However, in the initial model, only two of the five variables exhibited regression coefficients that were significantly different from zero; Conscientiousness, $\beta = -.39$, $t\,(70) = 2.94$, $p < .01$, and Extraversion, $\beta = -.44$, $t\,(70) = -3.02$, $p < .01$. Therefore, higher scores on Conscientiousness were associated with higher academic grades, whereas lower scores on Extraversion were associated with higher academic grades. There were no statistically significant linear relations detected between the mean of the academic grade and any of the remaining predictors.

The Conscientiousness and Extraversion variables were the only predictors to exhibit significant associations with the academic grade average outcome variable in the analysis, and these variables maintained significance through all models. The most parsimonious model included both Conscientiousness and Extraversion, and was significant, $R = .45$, $R^2 = .20$, $F(2, 70) = 8.98$, $p < .001$. The Adjusted R^2 value of .18 indicated that 18% of the variability in the academic grade was predicted by a model that only included the Conscientiousness and Extraversion variables. Both variables exhibited regression coefficients that were significantly different from zero; Conscientiousness, $\beta = .42$, $t\,(70) = 3.68$, $p < .001$, and Extraversion, $\beta = -.37$, $t\,(70) = -3.27$, $p < .01$. Therefore, higher scores on Conscientiousness were associated with higher academic grades, whereas lower scores on Extraversion were associated with higher academic grades. Therefore, based on these results hypothesis 8 is partially supported.

Personality predictors of flying average. A backward stepwise regression analysis was also performed using the personality predictors to determine the associations with the flying average outcome variable. The backward method was chosen to identify the variables that provide the most parsimonious model given a modest number of variables in the initial model (Cohen, Cohen, West, & Aiken, 2003).

The results of the backward stepwise regression analysis involving the personality and flying average variables are presented at Table 31. Model 1, the regression with all the personality variables in the model, was not significant, $R = .32$, $R^2 = .10$, $F(5, 67) = 1.56$, $p = .18$ The Adjusted R^2 value of .04 indicated that less than 4% of the variability in flying average was predicted by a model including all the personality variables. However, in the initial model (and all subsequent models), only Conscientiousness exhibited a regression coefficient that was significantly different from zero, $\beta = .35$, $t\,(70) = 2.46$, $p < .05$. Therefore, higher scores on Conscientiousness were associated with higher flying averages. There were no statistically significant linear relations detected between the mean of the flying average and any of the remaining personality predictors. The Conscientiousness variable was the only predictor to exhibit significant associations with the flying average outcome variable throughout the analysis. The most parsimonious model included only Conscientiousness in the regression equation; Model 5 was significant, $R = .28$, $R^2 = .08$, $F(1, 71) = 6.18$, $p < .05$. The Adjusted R^2 value of .07 indicated that almost 7% of the variability in the flying average was predicted by a model that only included the Conscientiousness variable. Therefore, hypothesis 8 was partially supported, in that Conscientiousness was the only personality variable that was significantly associated with the flying average outcome variable.

Table 31

Backward Stepwise Regression of Personality Variables and Flying Average (n = 73)

Predictor	R	R^2	Adj R^2	ΔR^2	F	p	B	SE B	β	t	p	r	pr	sr	sr^2
Model 1	0.32	0.10	0.04	0.10	1.56	.18									
Agreeableness							-0.64	2.63	-.04	-0.24	.81	.07	-.03	-.03	.00
Conscientiousness							6.53	2.65	.35	2.46	.02	.28	.29	.28	.08
Emotional Stability							0.78	3.40	.04	0.23	.82	.10	.03	.03	.00
Extraversion							-2.07	2.31	-.14	-0.90	.37	-.03	-.11	-.10	.01
Intellect/Culture							-0.89	2.92	-.05	-0.30	.76	.04	-.04	-.04	.00
Model 2	0.32	0.10	0.05	-0.00	1.96	.11									
Agreeableness							-0.38	2.35	-.02	-0.16	.87	.07	-.02	-.02	.00
Conscientiousness							6.71	2.52	.36	2.67	.01	.28	.31	.31	.09
Extraversion							-1.85	2.07	-.13	-0.89	.38	-.03	-.11	-.10	.01
Intellect/Culture							-0.90	2.90	-.05	-0.31	.76	.04	-.04	-.04	.00
Model 3	0.32	0.10	0.06	-0.00	2.64	.06									
Conscientiousness							6.61	2.42	.36	2.73	.00	.28	.31	.31	.10
Extraversion							-1.86	2.06	-.13	-0.90	.37	-.03	-.11	-.10	.01
Intellect/Culture							-1.08	2.66	-.06	-0.41	.68	.04	-.05	-.05	.00
Model 4	0.32	0.10	0.08	-0.00	3.93	.02									
Conscientiousness							6.25	2.25	.34	2.79	.00	.28	.32	.32	.10
Extraversion							-2.27	1.78	-.15	-1.27	.21	-.03	-.15	-.14	.02
Model 5	0.28	0.08	0.07	-0.02	6.18	.02									
Conscientiousness							5.24	2.11	.28	2.49	.02	.28	.28	.28	.08

Note. r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation

Summary of personality predictors. Hypothesis 8 predicted that higher levels of both conscientiousness and extroversion, and lower levels of neuroticism, would be associated with higher pilot training performance outcomes. This hypothesis was partially supported in that higher levels of conscientiousness were associated with both higher academic grades and flying averages; however, lower levels of extroversion were associated with higher academic grades.

Incremental validity of personality measures. The purpose of Hypothesis 9 was to investigate whether personality factors (i.e., openness, conscientiousness, extraversion, agreeableness, and emotional stability) account for unique variance in pilot training performance over and above cognitive ability, executive functions, psychomotor ability, and the work sample (CAPSS). The variables used to test this hypothesis were those showing significant associations with flying average as identified in previous hypothesis. Conscientiousness was the only personality factor included in this analysis as it was the only personality factor that was significantly associated with flying average performance, as identified in hypothesis 8. The descriptive statistics and bivariate correlations are provided in Table 32.

Table 32

Descriptive Statistics and Correlations of CAPSS, Psychomotor, Cognitive, Executive Functioning, Personality and Training Performance Variables

	n	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. CAPSS [a]	75	0.83	0.12	-									
2. Sensory Motor Apparatus [a]	70	91.41	38.71	-.18	-								
3. CFAT Verbal Skills [a]	75	11.23	2.20	-.14	.04	-							
4. CFAT Spatial Ability [a]	75	12.17	2.09	.07	-.04	.165	-						
5. Non Spatial Association [a]	73	0.07	0.75	-.13	-.19	.25*	.26*	-					
6. Random Letter Span [a]	73	0.20	0.99	-.24*	.001	.03	-.03	.19	-				
7. Conscientiousness [a]	73	0.54	0.80	.01	-.05	.06	.11	-.06	-.11	-			
8. Academic Grade [a]	75	96.52	1.99	.22*	-.20*	.12	.22*	.17	-.003	.29**	-		
9. Flying Average [b]	75	67.48	15.39	.20*	-.30**	-.20*	.27**	.23*	-.23*	.28**	.30**	-	
10. Good/Poor Performance [cd]	75	0.80	0.40	.26	-.37	-.23	.19*	.30	-.21*	.19	.37*	1***	-

Note. CAPSS = Canadian Automated Pilot Selection System; CFAT = Canadian Forces Aptitude Test

[a] Pearson correlations (1-tailed)

[b] Kendall's tau correlations (1-tailed)

[c] Biserial correlations (1-tailed)

[d] Good/Poor Performance: 0 = poor; 1 = good

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

Two separate hierarchical regression analysis were performed for each of the outcome variables, academic grade and flying average, to investigate the incremental validity of the personality variables over and above the other intelligence, cognitive, executive functioning, and psychomotor ability (when appropriate) measures.

Incremental validity of personality for academic grade. The first analysis was performed using the academic average outcome variable. The predictors included in this analysis were those that demonstrated significant relations with the academic grade outcome in previous hypothesis, including CFAT Problem Solving, and the two executive functioning variables, Go/No-Go and Sustained Attention. Although hypothesis 3 revealed a significant relation between the RAFAAT Table Reading variable and the academic grade, it was not included in the analysis as an earlier hierarchical regression analysis (hypothesis 5) revealed that RAFAAT Table Reading becomes non-significant when additional predictors are added to the model. The CAPSS variable was also included due to the significant bivariate correlation with the academic grade variable identified in Table 30. In this case, CAPSS is believed to represent more than a psychomotor ability measure; a previous study (Darr, 2010b) suggested that the variance in CAPSS not explained by psychomotor ability may actually represent a learning ability measure.

The order of entry of the variables into the hierarchical regression was determined *a priori*, as suggested by Cohen, Cohen, West and Aiken (2003). The CAPSS variable and the CFAT Problem Solving variables were entered as a set into step 1 as these measures are currently used in the Canadian Forces selection system, and the goal of the hypothesis is to determine measures that predict over and above the current measures.

Additionally, previous research has recognized general cognitive ability as a significant predictor of general academic and job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998) as well as pilot performance (Martinuusen & Hunter, 2010). The two executive functioning variables, Non-Spatial Association and Random Letter Span were added as a set in step 2, as it is believed that executive functioning ability would be more predictive than personality in such a cognitively demanding occupation. Finally, based on the results of hypothesis 8, Conscientiousness and Extraversion were added in step 3 to examine the incremental validity of personality in the prediction of the academic average. The results of the hierarchical regression analysis are displayed in Table 33.

In the first step of the three-step hierarchical regression the CAPSS and CFAT Problem Solving variables were regressed onto the academic grade outcome variable. The variables entered as a set in step 1 contributed significantly to the prediction of the academic grade in the first step $R^2 = .23$, $\Delta R^2 = .23$, $F(2, 70) = 10.18, p < .001$. There was a significant relation between the academic grade outcome variable and both predictors in step 1; CPASS $\beta = .26$, $t\,(70) = 2.43, p < .05$, and CFAT Problem Solving $\beta = .42$, $t\,(70) = 3.99, p < .001$.

The addition of the executive functioning (Go/No-Go and Sustained Attention) at step 2 accounted for a significant increment in the prediction of the academic grade; $R^2 = .34$, $\Delta R^2 = .11$, $F(4, 68) = 8.57, p < .001$. Therefore the addition of the executive functioning variables to the model with CAPSS and CFAT Problem Solving resulted in an increment to the validity of the prediction of the academic grade of 11%. Additionally, there was a significant relation between the academic grade outcome variable and three of the predictors in step 2; CPASS $\beta = .22$, $t\,(70) = 2.17, p < .05$,

Table 33

Hierarchical Regression of CAPSS, CFAT Problem Solving, ExamCorp Executive Functions (Go/No-Go and Sustained Attention), and Personality Variables (Conscientiousness and Extraversion) on Academic Grade

Predictor	R^2	Adj R^2	ΔR^2	F	p	B	$SE\ B$	β	t	p	r	pr	sr	sr^2 incr
Step 1	0.23	0.20	0.23	10.18	.00***									
CAPSS						4.10	1.69	.26*	2.43	.02	.22	.28	.26	.07
CFAT PS						0.24	0.06	.42	3.99	.002	.40	.43	.42	.18
Step 2	0.34	0.30	0.11	8.57	.00***									
CAPSS						3.53	1.63	.22	2.17	.03	.22	.25	.22	.05
CFAT PS						0.20	0.06	.36	3.45	.001	.40	.39	.34	.12
Go/No-Go						-0.76	0.25	-.31	-3.04	.003	-.29	-.35	-.30	.09
Sustained Attention						0.64	0.33	.20	1.93	.06	.28	.23	.19	.04
Step 3	0.43	0.38	0.10	8.45	.00***									
CAPSS						2.98	1.55	.19	1.92	.06	.22	.23	.18	.03
CFAT PS						0.19	0.06	.33	3.33	.001	.40	.38	.31	.10
Go/No-Go						-0.63	0.24	-.25	-2.65	.01	-.29	-.31	-.25	.06
Sustained Attention						0.48	0.32	.15	1.52	.13	.28	.18	.14	.02
Conscientiousness						0.81	0.26	.32	3.16	.003	.29	.36	.29	.09
Extraversion						-0.48	0.21	-.24	-2.33	.02	-.22	-.28	-.22	.05

Note. CAPSS = Canadian Automated Pilot Selection System; CFAT PS = CFAT Problem Solving; r = zero-order correlation; pr = partial correlation; sr = semipartial correlation; sr^2 = squared semipartial correlation; *incr* = incremental

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

CFAT Problem Solving $\beta = .36$, $t\,(70) = 3.45$, $p < .001$, and Go/No-Go $\beta = -.31$, $t\,(70) = -3.04$, $p < .001$.

The final model with all the predictor variables in the equation was significant, $R^2 = .43$, $\Delta R^2 = .10$, $F(6, 66) = 8.45$, $p < .001$. This represents increment to the prediction of the academic grade of 9.9% over the model with the CAPSS, CFAT Problem Solving and executive functioning variables in the equation, and an increment of almost 21% over the current model that only includes CAPSS and CFAT Problem Solving. Together these variables account for 38% of the variability in the academic grade outcome. However, two of the six variables exhibited non-significant associations with the academic grade variable; the Sustained Attention variable maintained a non-significant association from step 2, $\beta = .15$, $t\,(70) = 1.52$, $p = .13$, and the CAPSS variable was reduced to non-significance with the addition of the two personality variables, $\beta = .19$, $t\,(70) = 1.92$, $p < .06$. Therefore, hypothesis 9 is supported in that the personality variables, conscientiousness and extraversion added incremental validity to the prediction of the academic grade over CAPSS, CFAT Problem Solving, and the executive functioning (Go/No-Go and Sustained Attention) variables.

Incremental validity of personality for flying average. The order of entry of the variables into the hierarchical regression was determined *a priori*, as suggested by Cohen, Cohen, West and Aiken (2003). The CAPSS measure and two CFAT variables (Verbal Skills and Spatial Ability) were entered as a set into step 1 as these measures are currently used in the Canadian Forces selection system, and the goal of the hypothesis is to determine measures that predict over and above the current measures. Additionally, previous research has recognized general cognitive ability as a significant predictor of general academic and job performance (Hunter & Hunter, 1984; Schmidt & Hunter, 1998) as well as pilot performance (Martinuusen & Hunter, 2010). The psychomotor measure, Sensory Motor Apparatus, was added in step 2, as previous research has indicated that one of the best predictors of pilot training performance is psychomotor ability (Martinussen & Hunter, 2010). The two executive functioning variables, Non-Spatial Association and Random Letter Span, were added as a set in step 3, as it is believed that executive functioning will be more predictive than personality in such a cognitively demanding occupation. Finally, Conscientiousness was added in step 4 to examine the incremental validity of personality in the prediction of the flying average. The results of the hierarchical regression analysis are displayed in Table 34.

In the first step of the four-step hierarchical regression the CAPSS and CFAT variables (Verbal Skills and Spatial Ability) were regressed onto the flying average outcome variable. The variables entered as a set in step 1 contributed significantly to the prediction of the flying average in the first step $R^2 = .13$, $\Delta R^2 = .13$, $F(3, 64) = 3.19$, $p < .05$. There was a non-significant relation between the CAPSS variable and the flying average, when the CFAT variables were in the model; $\beta = .12$, $t\,(67) = 1.06$, $p = .30$; this non-significant association is maintained throughout the analysis. However, CFAT Verbal Skills was negatively related to the flying average, $\beta = .\text{-}.25$, $t\,(67) = -2.09$, $p < .05$, whereas CFAT Spatial Ability was positively related to the flying average, $\beta = .25$, $t\,(67) = 2.11$, $p < .05$.

The psychomotor ability measure, Sensory Motor Apparatus, added in step 2, accounted for a significant increment in R^2; $R^2 = .20$, $\Delta R^2 = .07$, $F(4, 63) = 3.81$, $p < .01$. Therefore, the Sensory Motor Apparatus variable accounted for an additional 6% of the

Table 34

Hierarchical Regression of CAPSS, CFAT (Verbal Skills and Spatial Ability), Psychomotor Ability, Executive Functions (Non Spatial Association and Random Letter Span), and Conscientiousness on Flying Average

Predictor	R^2	*Adj* R^2	ΔR^2	*F*	*p*	*B*	*SE B*	β	*t*	*p*	*r*	*pr*	*sr*	sr^2 *incr*
Step 1	0.14	0.10	0.14	3.58	.02									
CAPSS						19.85	13.56	.17	1.46	.30	.18	.13	.12	.02
CFAT Verbal Skills						-1.69	.83	-.24	-2.03	.04	-.22	-.25	-.24	.06
CFAT Spatial Ability						1.85	.89	.25	2.09	.04	.21	.26	.25	.06
Step 2	0.20	0.15	0.06	4.01	.00**									
CAPSS						14.10	13.45	.12	1.05	.51	.18	.08	.08	.01
CFAT Verbal Skills						-1.65	.811	-.24	-2.03	.04	-.22	-.26	-.24	.06
CFAT Spatial Ability						1.79	.87	.24	2.07	.04	.21	.26	.24	.06
SMA						.10	.05	.25	2.17	.03	.30	.27	.26	.07
Step 3	0.32	0.26	0.12	4.90	.00***									
CAPSS						9.44	13.15	.08	0.72	.68	.18	.05	.04	.00
CFAT Verbal Skills						-2.18	.78	-.31	-2.80	.01	-.22	-.35	-.30	.09
CFAT Spatial Ability						1.21	.83	.16	1.45	.15	.21	.18	.15	.02
SMA						.08	.04	.20	1.83	.06	.30	.24	.20	.04
NSA						6.16	2.44	.30	2.53	.01	.22	.31	.27	.07
RLS						-4.67	1.74	-.30	-2.68	.01	-.26	-.34	-.30	.09
Step 4	0.38	0.31	0.05	5.23	.00***									
CAPSS						10.03	12.73	.09	0.79	.61	.18	.07	.05	.00
CFAT Verbal Skills						-2.22	.75	-.32	-2.95	.00	-.22	-.36	-.31	.09
CFAT Spatial Ability						1.02	.81	.14	1.26	.22	.21	.16	.13	.02
SMA						.07	.04	.19	1.74	.08	.30	.23	.18	.03

NSA	6.46	2.36	.31	2.73	.01	.22	.34	.28	.08
RLS	-4.11	1.70	-.27	-2.41	.01	-.26	-.31	-.26	.07
Conscientiousness	4.46	1.96	.24	2.27	.03	.28	.28	.23	.05

Note. CFAT = Canadian Forces Aptitude Test; SMA = Sensory Motor Apparatus; NSA = Non-Spatial Association; RLS = Random Letter Span; *r* = zero-order correlation; *pr* = partial correlation; *sr* = semipartial correlation; sr^2 = squared semipartial correlation; *incr* = incremental

* $p < 0.05$, ** $p < 0.01$, *** $p < .001$

variability in the flying average performance over the CAPSS and CFAT variables. There was little change to the CFAT variables, as they maintained their significance and direction of association with the flying average; CFAT Verbal Skills, β = .-.25, t (67) = -2.09, $p < .05$ and CFAT Spatial Ability, $\beta = .24$, t (67) = 2.09, $p < .05$. There was a significant positive relation between Sensory Motor Apparatus and the flying average, $\beta = .26$, t (67) = 2.25, $p < .05$.

In step 3, the addition of the set of executive functioning measures also resulted in a significant increment in R^2; $R^2 = .32$, $\Delta R^2 = .13$, $F(6, 61) = 4.89$, $p < .001$. Therefore, the addition of the two executive functioning variables accounted for an additional 12% of the variability in the prediction of flying average over the CAPSS, CFAT, and psychomotor ability variables. Again, CAPSS remained non-significant. Although the association between the CFAT Verbal Skills and flying average variables remained significant, $\beta = -.32$, t (67) = -2.87, $p < .01$, the associations between the CFAT Spatial Ability and flying average dropped to non-significance, $\beta = .16$, t (67) = 1.45, $p = .15$, as did the association between the psychomotor ability measure and flying average, $\beta = .21$, t (67) = 1.89, $p = .06$. However, there was a significant positive association between the Non-Spatial Association variable and flying average, $\beta = .31$, t (67) = 2.56, $p < .05$, and a significant negative association between Random Letter Span variable and flying average, $\beta = -.31$, t (67) = -2.80, $p < .01$.

Finally, the addition of the personality measure, conscientiousness, in step 4 also contributed significantly to the prediction of the flying average over the CAPSS, general cognitive ability, psychomotor ability, and the executive functioning variables, $R^2 = .38$, $\Delta R^2 = .05$, $F(7, 60) = 5.20$, $p < .001$. The model with all the variables in the equation in step 4 accounts for 30% of the variability in the flying average score, and the addition of the personality variable accounts for an additional 5% of the variability in the flying average. Again, the associations of CAPSS (β = .06, t (67) = .52, p = .61), CFAT Spatial Ability (β = .14 t (67) = 1.25, p = .22), and psychomotor ability (β = .19, t (67) = 1.78, p = .08) with the flying average remained non-significant. The CFAT Verbal Skills and both executive functioning variables maintained their significance and the direction of their associations; CFAT Verbal Skills, β = -.33, t (67) = -3.00, p < .01, Non-Spatial Association, β = .32, t (67) = 2.75, p < .01, and Random Letter Span, β = .24, t (67) = 2.27, p < .05. Therefore, hypothesis 9, regarding the incremental validity of the personality variable over the other predictors was partially supported.

Summary of incremental validity of personality measures. Hypothesis 9 predicted that personality factors, namely conscientiousness and extroversion, would demonstrate incremental validity over the cognitive and psychomotor ability measures. However, the results indicate that this hypothesis was partially supported, as the addition of the personality variables led to a subsequent drop to non-significance for a number of the other predictors, including CAPSS, CFAT Spatial Ability, and SMA (psychomotor ability).

Discussion

In the Canadian Forces, the current pilot selection model includes the Canadian Forces Aptitude Test (CFAT; a measure of general mental ability), the Canadian Automated Pilot Selection System (CAPSS) simulator, and various sub-tests of the Royal Air Force Aircrew Aptitude Test (RAFAAT) battery. The overall purpose of the present research was to look beyond the currently used tests to investigate additional predictors of Primary Flight Training performance that could be used to improve pilot selection, with particular relevance to the Canadian Forces. At the time of this study, the RAFAAT measures were not included in the current Canadian Force's pilot selection model. Although the RAFAAT has not yet been validated with a Canadian Forces pilot training sample, a preliminary study (Darr, 2010b) resulted in the inclusion of a psychomotor ability measure in the Canadian Forces pilot selection model. Therefore, in addition to the current Canadian Forces pilot selection measures, this study also represented a validation of the RAFAAT battery and an examination of the effectiveness of a number of additional measures. A specific focus of this study was to explore the predictive power of executive functions when added to the current testing and selection approaches in the Canadian Forces, with the addition of the computer-based ExamCorp battery. Additionally, this current study represents a first in the Canadian Forces with regard to examining the relations of personality factors, specifically the Big Five, as predictors of pilot training performance.

The significant findings of this study offer a proof of concept for including executive functions in pilot selection testing, which should be further validated in a longitudinal study. The discussion here addresses the findings beyond all the separate hypotheses that were outlined to guide the analyses, but rather in a high-level, integrative fashion. The discussion starts with a brief overview of the results. The overview is followed by a more detailed discussion of the results, specified by domain (previous flying experience, work sample, cognitive abilities, executive functioning, personality, and psychomotor ability). Within each domain the significant findings will be discussed for each training outcome variable (i.e., good/poor performance, academic grade, and flying average). Additionally, the discussion for each domain will also include the findings that were counter to expectations. Following the discussion of the results, a tentative model is proposed based on the significant findings and the implications of this model for selection theory are discussed. Subsequent sections include the practical implications of this research, an acknowledgement of the limitations, and suggestions for future research. Finally, the overall significance and contribution of the study are presented.

Key Findings

Overall, the different measures displayed varying significant relations with the pilot training outcomes of the Primary Flight Training course. Significant relations were found between previous flying experience in and the current CFAT and CAPSS selection measures. Many of the ExamCorp measures of both the executive functioning and personality domains also demonstrated significant relations with the pilot training performance outcomes. With the exception of one measure of psychomotor ability and information processing measure, the validation of the RAFAAT indicated that this computer-based battery was not as successful in the prediction of pilot training performance as the ExamCorp battery. Given the small sample size and the restricted range, statistical artifacts, and marginally skewed data, the results should be interpreted cautiously, as statistical artifacts and small sample sizes have a tendency to bias the size of the correlations downward, thus creating inaccurate results (Hunter & Burke, 1995). Notwithstanding these factors, the next sections provide detailed discussions of the main results, commencing with previous flying experience, and followed by the cognitive measures, personality, and finally, psychomotor ability.

Previous Flying Experience

The results of this study reflect the strong positive relation between previous flying experience and Primary Flight Training performance. Indeed, of all the predictors, previous flying experience demonstrated the strongest relations with the academic grade and flying average, and was also positively related to CAPSS performance. These results are aligned with numerous studies that have found that previous flying experience is a good predictor of both Primary Flight Training performance and CAPSS performance (Carretta & Ree, 1994; Darr, 2009, 2010b; Deitcher & Johnston, 2004; Johnston & Catano, 2010; Kaplan, 1965; Martinussen, 1996; Ree, Carretta, & Teachout, 1995; Woycheshin, 2002). Johnston and Catano (2010) explained that previous flying experience represents accumulated job knowledge that influences performance on CAPSS and both academic and flying performance.

Work-Sample Measure

The work-sample used in the Canadian Forces is the Canadian Automated Pilot Selection System (CAPSS), which simulates a small, single-engine light aircraft in the Instrument Flying Rules (IFR) environment. The current study found that CAPSS performance predicted both academic and flying performance in Primary Flight Training. This result is aligned with the findings of previous research that have reported CAPSS to be a good predictor of Primary Flight Training (Darr, 2009, 2010b; Deitcher & Johnston, 2004, Johnston & Catano, 2010; Spinner, 1991; Woycheshin, 1999, 2001, 2001). As a work sample, it requires the candidate to perform the tasks that are specific to the job (Hunter & Burke, 1995); a number of military and civil aviation selection systems have included some form of computer-based simulator assessment (Suarez, Barborek, Nikore, & Hunter, 1994).

General Intelligence and Specific Cognitive Abilities

The present study proposed that the measures from the Cognitive Reasoning Battery (CRB) of the Royal Air Force Aircrew Aptitude Test (RAFAAT) would predict pilot training performance in Primary Flight Training over and above the Canadian Forces Aptitude Test (CFAT). Of the CFAT variables, Problem Solving predicted academic performance. Additionally, the correlation analysis for the significant relations between the predictors and flying average, used Kendall's τ_b to account for the skewed flying average in a small sample, and found only Problem Solving was associated with the flying average. Conversely, the backward regression indicated that CFAT Verbal Skills and Spatial Ability were significant predictors of flying performance. However, due to the skew associated with the flying average, it is thought that the results of the parametric regression analysis are not as accurate as the non-parametric correlation analysis which accounted for the skew in the flying average variable by applying the Kendall's τ_b correlation. Therefore, only CFAT Problem Solving was used in the proposed model in this study.

Contrary to the hypothesis, none of the three measures included in the Critical Reasoning Battery (verbal, numerical, or spatial reasoning) proved to be significant predictors of pilot training performance. A number of factors could account for this non-significant result, the most likely being that all participants were previously selected based on meeting a minimum CFAT cut-off, which has restricted the range of similar measures. Additionally, the CFAT and CRB sub-tests were moderately to highly correlated, suggesting that they were indeed measuring similar competencies.

The RAFAAT battery also includes a number of sub-tests that measure more specific cognitive abilities, such as information processing (work rate), attention (vigilance), and short-term memory. Of the four sub-tests (table reading, digit recall, vigilance, and visual search), only Table Reading demonstrated a significant positive relation with academic performance. The United Sates Air Force includes a Table Reading sub-test in their pilot selection model as a measure of perceptual speed (Drasgow, Nye, Carretta, & Ree, 2010), and a similar test in the Royal Air Force is known as an information-processing measure (Southcote, 2004). Regardless of the differences in terminology of the domain, the significant result reflects the importance of working quickly and accurately, which is necessary for achieving academic success. Aside from the Table Reading variable, none of the RAFAAT specific cognitive ability measures demonstrated significant relations with either of the pilot training performance measures. The lack of significant findings for these measures could be attributed to the small sample size.

Executive Functions

Executive functioning refers to higher level cognitive processes, mainly centered in the prefrontal cortex, and includes working memory, attention control (inhibition), associated learning, and word fluency (Peterson, 2003). It was expected that the executive function measures would be associated with better flight performance because of the complexity of flying tasks, particularly those associated with monitoring instruments and the environment, and identifying and responding to potential inconsistencies/emergencies. A number of pilot selection models have included some measures of specific cognitive abilities such as attention and memory (Skinner & Ree, 1987; see Bailey, M., 1999). However, the ExamCorp battery offers a different method of assessing working memory from other current measures, and it includes additional measures, such as associated learning, that have not previously been considered in pilot selection. The ExamCorp battery is a commercially-available off-the-shelf battery of computer-based measures of executive functioning. It includes three measures of working memory, two measures of attention control and inhibition, three measures of associated learning, and one measure of word fluency. The ExamCorp battery investigates psychological domains that have not yet been investigated in the Canadian Force's pilot selection model. These results are first discussed in relation to the associations with the academic grade outcome variable and followed by the flying average.

After taking into account the current pilot selection measures, the relations between executive functioning and pilot training performance in Primary Flight Training was investigated. Consequently, the results identified three executive functioning measures as being positively related to academic performance. Specifically, one associated learning measure (Go/No-go), and two measures of attention control (Sustained Attention and Response Inhibition) predicted academic performance. The Go/No-go represents the ability to learn the inhibitory task rule, and the score reflects the total number of trials required to learn the inhibition rule. A higher number of learning trials is indicative of poorer performance. The results from this study indicated that having more difficulty learning the inhibition rule involved with the Go/No-go task was negatively related to the academic grade. This result supports the hypothesis, as it was expected that individuals who are quicker to learn associated rules (i.e., fewer number of trials) would achieve higher training performance outcomes than individuals who were slower to learn such associated rules. The ability to learn rules quickly is necessary in the complex aviation environment, in which changes in an instrument, system, or weather can affect aircraft operations, and potentially result in aircraft emergencies if the correct assumptions and actions are not arrived at quickly.

The Response Inhibition task, a measure of attention control, represents the ability to inhibit a response upon the appearance of the target. The results supported the hypothesis in that response inhibition was positively related to the academic average. It was expected that higher levels of academic performance entails the ability to inhibit responses (i.e., be less impulsive) to outside stimuli that would detract from studies. Sustained Attention, another measure of attention control, represents the ratio of correct responses to the number of target appearances within a string of distracting objects. Sustained Attention, or vigilance, was also positively related to academic performance. The results supported the hypothesis that sustained attention would be positively related to pilot training performance. It also suggests that it is likely the self-regulatory type of sustained attention (i.e., persistence) that is associated with the attainment of better academic grades, which is based on theory regarding the types of inattention that are associated with the different forms of ADHD (Preston, Heaton, McCann, Watson, and Selke, 2009).

Although one of the executive functioning measures, Non-Spatial Association, predicted flying performance, the negative relation was not expected. The Non-Spatial

Association variable was a measure of associated learning, and reflects the total number of consecutive trials required to successfully respond to stimuli; therefore, higher scores on this measure reflect a higher number of trials required to learn the rule, and thus lower levels of associated learning ability. Contrary to the hypothesis, the positive relation between Non-Spatial Association (error score) and flying performance indicates that poorer abilities to make conditional associations are associated with better flying performance. However, there is no logical explanation for this relation, and it is contrary to previous research (Darr, 2010b; Johnston & Catano, 2010) that indicates that learning ability reflected as a measure of *g* and is positively associated with pilot training performance. This unexpected result may be the effect of the interaction between the skewed Non-Spatial Association variable and the skewed flying average variable. Therefore, the unexpected results could be due to the influence of statistical biases and not necessarily reflect real relations.

Overall, only three of the executive functioning measures (Go/No-go, Sustained Attention, and Response Inhibition) investigated in this study predicted pilot training performance, and only academic performance, not flying performance, was predicted by the executive functioning measures. However, it was expected that the executive functioning measures would be more relevant as predictors of flying performance, due to the complex, cognitively-loaded demands of the flying environment. Future research should investigate whether the higher levels of executive functioning are less important during the early stages of flight training when procedures are fairly simple, but more important later on when pilots are flying more complex aircraft, entailing higher workloads, multi-tasking, dealing with complex situations, and decision making.

Notwithstanding, this study demonstrated that the executive functioning measures provided incremental validity over the current measures used in pilot selection in the Canadian Forces (i.e., CAPSS, CFAT, and Sensory Motor Apparatus [SMA]), and their inclusion, even in the selection for Primary Flight Training, would add to the predictive power of the selection battery.

Personality

Personality factors, measured by the ExamCorp personality sub-tests of the Big 5 personality factors, also played a role in the prediction of pilot training performance in Primary Flight Training. In this study, conscientiousness was positively associated with both the academic grade and the flying average, and extroversion was negatively associated with the academic grade. The results for conscientiousness are in line with previous research that has found positive links between conscientiousness and training and job performance for both general (Barrick & Mount, 1991; Hogan & Ones, 1997; Mount & Barrick, 1995, Salgado, 1997; Tett, Jackson, & Rothstein, 1991) and aviator-specific occupations (Carretta, Rodgers, & Hansen, 1996; Martinussen, 1996; Siem & Murray, 1994). As the Canadian Forces has recently added the Trait-Self Descriptive Personality Inventory (TSD-PI; Christal, 1994) Inventory to its selection model, future research should be conducted to determine whether the findings from the current study would be replicated with the TSD-PI.

Contrary to expectations, the current study found a negative relation between extraversion and academic performance. This study contradicts previous studies, including a meta-analysis of 26 studies by Campbell, Casteneda, and Pulos (2009), that have generally found a small positive relation between extraversion and pilot training success (see meta-analysis in). Similarly, Barrick, Mount, and Judge (2001) found a positive relation between extroversion and training performance in studies involving non-military samples. The negative relation between extraversion and academic performance could reflect the fact that in the demanding, cognitively-loaded pilot training environment, in which practice and dedication to studies are critical, less extraverted pilot students spend more time on their studies and less time socializing with classmates, and thus achieve better academic results. However, a recent study involving medical school students suggests that personality characteristics, particularly openness and extraversion, were more relevant in later medical training in applied settings when effective patient interaction became the important criterion, than in early training when academic grades were used as the criterion (Lievens, Ones, & Dilchert, 2009). Based on the contradictory results of the extraversion measure with previous research, it is suggested that only the measure of conscientiousness should be considered in the Canadian Force's pilot selection model.

Other studies suggest that personality has different relations with both pilot training and job performance. In their meta-analysis, Campbell and colleagues (Campbell et al., 2009) found anxiety, a facet of neuroticism, was significantly negatively related to pilot training performance. As the personality measure used in the current study did not permit facet-level investigation, future research should involve a measure that is able to assess the more detailed facet-level personality characteristics. Additionally, although high levels of conscientiousness are certainly advantageous to persevere and be successful throughout the progressively more demanding pilot training process, perhaps different personality characteristics become more relevant during later training and actual job performance when the environment becomes more cognitively complex, and effective crew interaction and leadership responsibilities become more involved in daily tasks. Lievens and colleagues (2004) found personality characteristics to be increasingly relevant when candidates have been preselected based on high-end cognitive abilities, which is certainly the case with pilot candidates. Another study (Studer-Luethi, Jaeggi, Buschkuehl, & Perrig, 2011) found a link between personality, namely neuroticism and conscientiousness, and performance in cognitive training, specifically working memory tasks. Such a result shows the relevance of personality characteristics in the more advanced levels of pilot training, which are more cognitively complex. Therefore, perhaps the validation of the personality characteristics should incorporate advanced pilot training and pilot job performance (effective/successful pilots), not initial pilot training performance, as the criterion.

Psychomotor Ability

The results of this study found that a multi-limb coordination sub-test, Sensory Motor Apparatus (SMA), of the RAFAAT was a strong predictor of flying performance in Primary Flight Training. The multi-limb coordination measure required the candidate to use a joystick and foot pedals to control the position of a dot within a target crosshair on the computer screen. The results for this study supports the hypothesis and are aligned with previous research that indicates that psychomotor ability is one of the best predictors of pilot training performance (Martinusson, 1996), and is a domain that many countries have included in their pilot selection models since the end of the First World War (Martinussen & Hunter, 2010). Furthermore, the results indicate that the computer-based measure of psychomotor ability accounted for more variance in the flying performance than the current CAPSS measure. Good psychomotor abilities and multi-limb coordination are required to coordinate flying tasks such as the operation of the yoke (stick), which controls aircraft direction and altitude, and the rudder pedals, which control yaw. To effect a smooth turn, a pilot must simultaneously coordinate these two systems (yoke and rudder pedals). Therefore, this psychomotor ability measure shows promise for inclusion in a future selection model that may not include CAPSS, specifically when one considers the lower financial costs and ease of administration and maintenance associated with a computer based measure.

A Possible Selection Decision Model

The current study proposes that an alternative decision model to the one currently employed in the Canadian Forces should be investigated further. The current decision model based on a multiple hurdle model in which applicants are required to meet minimum cut-off scores for each predictor before moving on to the next predictor (Catano, Cronshaw, Wiesner, Hackett, & Methot, 1997). Applicants must first meet the minimum cut-off score on the CFAT, before progressing to the structured interview, and performance on the structured interview is assessed before candidates are sent to the Aircrew Selection Centre for the pilot-specific selection tests. However, at the Aircrew Selection Centre the decision model is then based on a multiple cut-off model, where applicants are rejected if their scores do not meet the cut-off score for all the predictors (CAPSS performance, psychomotor ability, etc). Candidates who meet the cut-off on all predictors are then considered in a top-down selection process, in which the scores on all predictors are weighted equally, candidates are ranked according to the combination of the scores, and selection decision is based on the top performers (Catano et al., 1997).

The current method of calculating the candidate's score does not allow much variability, as scores range from 0 to 4 (half points are possible), with a score greater than 2 being considered for selection. However, there may be numerous candidates within the same score ranking, and currently the most impartial method to make selection decisions within each ranking rests on something as inconsequential as the earliest date of application. This lack of variability prompted the proposal of an alternate decision making model for pilot selection in the Canadian Forces.

The current study proposes that future research should investigate a multiple regression (weighted regression) model to improve pilot selection decision making. The multiple regression model presented in this study would require further investigation in recognition of the limitations of this study, specifically with respect to the small sample size, which makes generalizability a challenge. The proposed decision-making model would weight a candidate's score on all predictors and sum them to produce a total score that represents the predicted training performance score (Catano et al., 1997). Catano and colleagues (1997) contend that linear model approaches generally outperform other decision making approaches. Indeed, Hunter and Burke (1995) suggested using the statistical approach (i.e., regression model) to create a pilot composite score that combines the pilot selection data, over all other decision making approaches. The United States Air Force has used a regression model since 1993, and studies have demonstrated that the Pilot Candidate Selection Method is a significant predictor of pilot training performance (Carretta, 2011).

There were a number of factors that were considered in the construction of the proposed model here. First, the sample size was small relative to the number of predictors, which does not meet a key regression assumption, meaning the results are likely inaccurate (Hunter & Burke, 1995; Tabachnick & Fidell, 2007). As a result, this model is presented as a tentative model to be evaluated in future research with a larger sample size. Second, the regression model would only be applied to candidates who meet the minimum cut-off for CAPSS and CFAT predictor variables; future research is required to establish cut-offs for the other predictor variables. Third, the regression model was created by considering the appropriate outcome measures and the associated predictors. The creation of the weighted regression model required a decision as to which training performance measure would permit the most accurate selection decision. Although the majority of the new measures of executive functioning only demonstrated significant relations with academic performance, few pilot candidates fail the ground school portion of the training. As performance on the flying phase of training has the greatest impact on the candidate's overall success on Primary Flight Training, the outcome variable of greatest interest is the flying average. Accordingly, the predictors that were included in the regression equation were those that this study identified as having significant relations with flying performance. Such predictors include previous flying experience, CAPPS performance, psychomotor ability (Sensory Motor Apparatus), CFAT Problem Solving, and Conscientiousness. The Non-Spatial Association variable was not included in the model as few selection models would consider a lower score in associated learning as ideal for selection. Based on these factors, future research should investigate the weightings of the proposed model, which is presented as follows:

$$\begin{matrix}\text{Flying}\\ \text{Performance}\end{matrix} = .25 \begin{matrix}\text{Psychomotor}\\ \text{Ability}\end{matrix} + .23\ \text{Conscientiousness} + .13 \begin{matrix}\text{Previous}\\ \text{Flying}\\ \text{Experience}\end{matrix} + .11 \begin{matrix}\text{CFAT}\\ \text{Problem}\\ \text{Solving}\end{matrix} + .08 \begin{matrix}\text{CAPSS}\\ \text{Performance}\end{matrix}$$

This model reflects that pilot candidates with higher levels of psychomotor ability, conscientiousness, previous flying experience, problem solving ability, and CAPSS performance were significant predictors of the flying average. The model using these predictors of flying experience was significant, $R^2 = .20$, $F(5, 62) = 3.20$, $p < .05$. The composite weighted score for each candidate is calculated by multiplying the candidate's actual score on each predictor above by the weighted score and then summing the results of those weighted scores. This calculation results in a regression equation which would then be used as the basis to rank candidates in the top-down selection process. Future research with a larger sample should compare the results of the current top-down selection model to the results of a top-down selection model involving a weighted regression score, such as the one proposed here.

Although the proposed model presents an interesting conceptualization of the best weighted combination of predictors based on the results of the current study, there are limitations to such a model that should be acknowledged. As this model used flying performance as the outcome, none of the predictors of academic performance were included. Therefore, none of the executive functioning or information processing measures were included in the model, which was a goal of the study. However, the choice of outcome measure for the model was based on the importance of the flying performance as the greatest variance was found in this measure. As suggested earlier in the discussion, perhaps the lack of significant findings between the complex executive functioning measures and flying performance are due to the emphasis of psychomotor ability in Primary Flight Training, and executive functioning becomes more important in flying performance in subsequent, more advanced flight training. To this end, future research should investigate this hypothesis. Notwithstanding the current Canadian Forces top-down selection model, which includes all significant predictors (i.e., for both academic and flying performance) of both academic and flying performance, a weighted regression model would permit more precise selection scores to better differentiate successful pilot candidates.

Practical and Theoretical Contributions

This current study has a number of practical and theoretical implications. First, although the executive functioning variables were not included in the proposed selection model, this study demonstrated the relevance of executive functioning measures, such as the associated learning and response inhibition measures, as predictors of pilot training performance in the Canadian Forces. However, the emergence of the significance of the executive functioning variables with academic performance and not with flying performance, suggests that higher-order cognitive processes may not be as important as psychomotor ability in Primary Flight Training. Therefore, future research is required to indentify whether executive functioning is more relevant in advanced pilot training courses. The investigation of the computer-based Royal Air Force Aircrew Aptitude Test battery identified a psychomotor ability measure that has the practical implication of adding to the Canadian Force's pilot selection model. The psychomotor ability measure also offers an alternative to a costly, high fidelity simulator.

The proposed model also has both practical and theoretical implications. A practical implication is the support for an alternative to the equally-weighted top-down pilot selection model currently used by the Canadian Forces, which has never been implemented in the Canadian Forces. The weighted model also has practical implications by allowing for greater differentiation between candidates and thus permit better selection decisions. The theoretical implication of this study is the support for the weighted approach to pilot selection decision-making models, which represents a theoretical explanation of the interactions of the various factors and how they affect behaviour and performance.

Furthermore, this study may have practical implications regarding methods to address the inadequate pilot training criteria. To capture the complex cognitive aspects of Primary Flight Training, it is suggested that the flight training assessments (Progress cards) include behavioural indicators of such constructs as mental workload or situation awareness. The development of such behavioural indicators of complex cognitive processes would require additional research involving cognitive task analysis and multiple working groups with Subject Matter Experts. Therefore, a practical implication is that more salient links between the training and selection systems in the Canadian Forces could be established to achieve these goals.

Limitations and Future Research

There were several limitations to the current study, mostly stemming from the inadequate criterion variables, small sample size, restricted range, and methodological issues. A key limitation to this study, and many studies involving pilot selection, involves the outcome variables, including absence of an assessment of cognitive performance and the inadequacy of the dichotomized outcome. Flight training assessments generally measure a student's performance on a number of aviation tasks, yet none of the assessment items attempt to measure the more cognitively-loaded aspects of flying, such as perceived work-load or situation awareness. As the challenge for many pilot selection researchers has been to develop accurate measures of pilot cognitive abilities, it would be beneficial to align these cognitively-loaded selection measures with more accurate measures of the cognitive aspects of actual flying performance. Although work-load and situation awareness are sometimes ambiguous constructs, the development of operational definitions and assessment using behavioural indicators of such complex cognitive processes should be considered. Such an endeavor would require flight instructors to undergo specific training to familiarize themselves with the scoring methodology to ensure standardization. Yet the time and effort required to develop and implement such behavioural-based assessments of cognitive performance may yield much richer data from which more robust measures of cognitive performance may be validated.

This process could be facilitated by a more cognitively-focused pilot job analysis in the Canadian Forces, to identify the tasks and KSAOs relating to situation awareness, as well as more specific cognitive demands experienced by pilots. Indeed, such research could build on Cass and Parush's (2011) study which included a cognitive task analysis for Canadian Forces helicopter pilots. Consequently, it is recommended that a combination job analysis method (CJAM; Levine, 1983) should be performed, in which the results of the functional job analysis conducted by Darr (2010a) are augmented with a more thorough Cognitive Task Analysis (Wei & Salvendy, 2004). Moreover, future research should investigate situation awareness in fixed-wing pilots using CogScreen – AE, which Cass and Parush (2011) found predicted situation awareness constructs (recall, retention, working memory, attribute identification, and motor tracking) for rotary wing pilots on a simulated helicopter deck landing task. Measures such as CogScreen – AE could then be used to link selection to the new cognitively-loaded training assessments proposed above.

Dichotomized training variables are another limitation to selection research. According to Hunter and Burke (1995), "what will be measured and utilized as the criteria is essential" (p.12). A number of researchers (Carretta & Ree, 2006; Hunter & Burke, 1995; Hunter & Schmidt, 1990) have acknowledged the shortcomings of dichotomized training outcomes such as Pass/Fail, which are prevalent in pilot research. The dichotomization of the training outcome results in a reduced maximum correlation size which is decreased even further due to uneven group proportions (i.e., between Pass/Fail; Cohen, 1983; Hunter & Burke, 1995). In addition, the variability in the pilot training criteria is lost when Pass/Fail is used to describe the training outcome, as the variable does not elaborate on the strength of the success or failure (Hunter & Burke, 1995). More detailed investigation of the specific factors relating to the training failures may identify additional domains or competencies that have previously been overlooked, and ultimately may lead to the identification and development of a more descriptive overall training assessment.

A small sample size also limited the results in this study. The sample size for this study was 75, which was reduced to 68 for some of the analyses due to missing data. As a result, many of the non-significant findings may be due to a lack of statistical power. This is particularly relevant for the RAFAAT sub-tests, which were limited to the reduced sample size of 68. Conversely, small samples also have the potential to produce statistical artifacts that artificially inflate correlations (Burke, Hobson, & Linsky, 1997). Additionally, this data set included a number of variables that were skewed to various degrees, and smaller samples are not robust to the violations of normality which larger sample sizes afford (Tabachnick & Fidell, 2007). Consequently, the results may not be an accurate representation of the associations in the population, and these relations should be re-examined in follow-on research with a larger sample size.

An additional concern associated with the small sample size is the effect of uneven group splits for dichotomous training outcomes, such as Pass/Fail. For this study the Pass/Fail training variables had such a severe group split that it was considered unsuitable for further analyses. The unfortunate implication of applied research, particularly in pilot selection, is small sample sizes. Regardless of the limitations due to the small sample size for this study, some significant results were found that provide insight into the cognitive factors that influence pilot training performance. However, to compensate for the small sample size, future research should involve concurrent validity studies with trained, experienced pilots, as well as longitudinal studies, which follow the pilot candidates past the training stage and into their actual flying careers.

Range restriction was another limitation that was present in this study. All candidates had previously been selected based on performance on the CFAT and CAPSS performance, which likely reduced the correlations between predictor and outcome variables, thus creating inaccurate results (Hunter & Burke, 1995). The effects of range restriction are even more pronounced as training attrition at each successive course reduces the number of candidates at further stages (Hunter & Burke, 1995). Although there are calculations available to correct for the restriction of range, in this case it is not possible as the initial validation of the CAPSS was also used on a restricted sample of candidates who had already been selected based on the pilot selection system at the time.

Alternatively, to mitigate the effects of range restriction, Hunter and Burke (1995) suggest developing selection measures that assess advanced skills using simulators.

Methodological issues concerning the timing of the administration of the RAFAAT and ExamCorp batteries were another limitation to this study. First, all participants had already been pre-selected for pilot training, and the RAFAAT and ExamCorp test batteries were not completed in a high-stakes situation. Consequently, the results may not adequately reflect performance in a selection situation, where participants are motivated to perform better as they know selection is based on better performance on the tests. Additionally, administration of the computer-based testing occurred at different times during the training cycle (i.e., prior to course start, after ground school, and during flying phase) for the various courses involved in the study. Different stressors and individual concerns associated with the different administration timings could have influenced the results. Early in the study an attempt was made to mitigate the influence of the timing of the administration, and an additional day was added to the training syllabus. Notwithstanding, the most salient factor of the administration is that the testing was not performed in a high-stakes environment, and therefore participants may not have given maximum effort.

Conclusion

This current study contributes to the understanding of the important predictors in pilot selection in the Canadian Forces. In general, the findings of the associations of the predictors with academic and flying performance in the present study validate findings from previous research in the context of the Canadian Forces. Previous flying experience (PFE), the Canadian Automated Pilot Selection System (CAPSS) simulator, the Canadian

Forces Aptitude Test (CFAT) Problem Solving subtest, psychomotor ability (RAFAAT), and conscientiousness predicted flying performance on Primary Flight Training. With the exception of psychomotor ability, the same predictors (PFE, CAPSS, CFAT Problem Solving and conscientiousness), as well as an information processing measure, attention control, and two associated learning measures predicted academic performance on Primary Flight Training.

Considering the fiscal pressures of the current era, when departments are forced to continue to perform with fewer financial resources, it is imperative that significant predictors of pilot training and job performance be employed in order to reduce training attrition. The right combination of computer-based tests would allow the assessment of the cognitive domain, which has previously been untapped in the Canadian Force's pilot selection model. Computer-based testing of psychomotor and cognitive abilities have the added benefit of generally being less expensive to purchase and both easier and less costly to maintain than simulator-based work sample measures. The lack of significant results between the executive functioning measures and flying performance suggested that future hypothesis should test whether the same is true for flying performance on more advanced pilot training.

Although the results of this study provide theoretical and practical implications that have the potential to improve the current pilot selection model, there is still unaccounted-for variance in the model. As some of this unaccounted-for variance may be a result of statistical artifacts, follow-on research with a larger sample size is recommended. Additionally, longitudinal and concurrent validity studies of these measures are suggested in order to determine the associations with subsequent pilot training courses and job performance. Finally, these studies could investigate whether there are differences in pilots' test results across various airframes.

References

Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence, 10*(2), 101-139.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*(1), 17-42.

Anderson, N., Born, M., & Cunningham-Snell, N. (2002). Recruitment and selection: Applicant perspectives and outcomes. In N. Anderson, D. S. Ones, H. K. Sinangil, and C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology*. Thousand Oaks, CA: Sage Publications Inc.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A* (1), 5-28.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory. *Trends in Cognitive Science*, 4, 414-423.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bowser (Ed.), *The Psychology of Learning and Motivation, Vol. 8* (pp. 47-89). New York: Academic Press.

Bailey, M. (1999). *Evolution of aptitude testing in the RAF*. Directorate of Recruiting and Selection (Royal Air Force). Paper presented at the RTO HFM Workshop on Officer Selection held in Monterey, CA (RTO MO -55).

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

Barrick, M. R., & Mount, M. K. (2001).

Barrick, M. R., & Mount, M. K. (2003). Impact of meta-analysis methods on understanding personality-performance relations. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 197-222). Mahwah, NJ: Lawrence Erlbaum.

Barrick, M. R., & Mount, M. K. (2005). Yes, personality matter: Moving on to more important matters. *Human Performance, 18*, 359-372.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance in the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9-30.

BaseOps (2012). *Test of Basic Aviation Skills (TBAS)*. Retrieved from http://www.baseops.net/militarypilot/tbas.html

Beier, M. E., & Ackerman, P. L. (2007). Personnel selection and testing. In Francis Durso (Ed.) *Handbook of applied cognitive psychology*, (2nd ed.) (pp. 604-627). John Wiley & Sons, Ltd: West Sussex, England.

Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioral Science, 12*, 8-17.

Boswell, R., & Kuschnereit, F. (2009, April). *Canadian Forces Aptitude Test Manual* (CR 2009-002). Ottawa, ON: Director General Military Personnel Research and Analysis.

Bracket, M. A & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29*(9), 1147-1158. doi: 10.1177/0146167203254596

Burke, E., Hobson, C., & Linsky, C. (1997). Large sample validations of three general predictors of pilot training success. *The International Journal of Aviation Psychology, 7*(3), 225-234.

Burke, E., Kokorial, A., Lescreve, F., Martin, C. J., Van Raay, P., & Weber, W. (1995). Computer-based assessment: A NATO survey. *International Journal of Selection and Assessment, 3*(2), 75-83.

Byrdorf, P. (1993). Pilot candidate selection in the Royal Danish Air Force and EURO NATO aircrew selection working group test battery: A research program. In R. S. Jensen, & D. Neumeister (Eds.), *Proceedings of the 7th International Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

Campbell, J. S., Castaneda, M., & Pulos, S. (2009). Meta-analysis of personality assessments as predictors of military aviation training success. *The International Journal of Aviation Psychology, 20*(1), 92-109.

Carretta, T. R. (1989). USAF pilot selection and classification systems. *Aviation, Space, and Environmental Medicine, 60*(1), 46-49.

Carretta, T. R. (1992). Short-term retest reliability of an experimental U.S. Air Force pilot candidate selection test battery. *The International Journal of Aviation Psychology, 2*(3), 161-173.

Carretta, T.R. (2011). Pilot Candidate Selection Method: Still an effective predictor of US Air Force pilot training performance. *Aviation Psychology and Applied Human Factors, 1*(1), 3-8.

Carretta, T. R., Rodgers, M. N., & Hansen, I. (1996). *The identification of ability requirements and selection instruments for fighter pilot training* (Tech Report 2). Euro-NATO Aircrew Human Factor Working Group.

Carretta, T. R., Perry, & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology, 6*(1), 21-41.

Carretta, T. R., & Ree, M. J. (1994). Pilot candidate selection method (PCSM): Sources of validity. *International Journal of Aviation Psychology, 4*, 103-117.

Carretta, T. R., & Ree, M. J. (1995). Air Force Officer Qualifying Test validity for predicting pilot training performance. *Journal of Business Psychology, 9*(4), 379-388.

Carretta, T. R., & Ree, M. J. (1997). Expanding the nexus of cognitive and psychomotor abilities. *International Journal of Selection and Assessment, 5*(3), 149-158

Carretta, T. R., & Ree, M. J. (2000). General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment, 8*(4), 227-236

Carretta, T. R., & Ree, M. J. (2003). Pilot selection methods. In B. H. Kantowitz (Series Ed.), P.S. Tang, & M. A. Vidulich (Eds.), *Human factors in transportation: Principles and practices of aviation psychology* (pp. 357-396). Mahwah, NJ: Lawrence Erlbaum Associates.

Carretta, T. R., & Ree, M. J. (2006). The role of measurement error in familiar statistics. *Organizational Research Methods, 9*(1), 99-112.

Carroll, J. B. (Ed.). (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Carter, J., Boswell, R., Kushnereit, F., & Ebel-Lam, A. (in press). *Canadian Forces Aptitude Test Manual: 2012 Update* (Technical Report). Ottawa, ON: Director General Military Personnel Research and Analysis.

Cascio, W. F. (1991). *Applied psychology in personnel management* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Cass, E. & Parush, A. (in press). Can situation awareness be predicted? Investigating relationships between CogScreen – AE and pilot situation awareness. *International Journal of Aviation Psychology*.

Catano, V. M., Cronshaw, S. F., Wiesner, W. H., Hackett, R. D., & Methot, L. L. (Eds.). (1997). Selection II: Testing. In M. Belcourt (Series Ed.), *Recruitment and selection in Canada*. Toronto, ON: Nelson.

Catano, V. M., Wiesner, W. H., Hackett, R. D., & Methot, L. L. (Eds.). (2005). Selection II: Testing. *Recruitment and selection in Canada* (3rd ed.). Toronto, ON: Nelson.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal Social Psychology, 38*, 476-506.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1-22.

Cattell, R. B. & Horn, J. L. (1987). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement, 15*, 139-164.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Causse, M., Dehais, F., & Pastor, J. (2011). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *The International Journal of Aviation Psychology, 21*(3), 217-234.

Christal, R. (1994). *The air force self descriptive inventory*. Final R&D Status Report: United States Air Force.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249–253.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (Eds.). (2003). In *Applied multiple regression/correlation analysis for the behavioral sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Costa, P.T. (1996). Work and personality: Use of the NEO PI-R in industrial/organizational psychology. *Applied Psychology: An international Review, 45*(3), 225-241.

Costa, P. T. & McRae, R. R. (1988). From catalogue to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology, 55*, 258-265.

Costa, P. T. & McRae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, Florida: Psychological Assessment Resources, Inc.

Damos, D. L. (2007). *Foundations of military pilot selection systems: World War 1 (Technical Report 1210)*. Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

Darr, W. (2009). *A psychometric examination of the Canadian Automated Pilot Selection System (CAPSS),* (TM 2009-024). Ottawa, ON: Director General Military Personnel Research & Analysis.

Darr, W. (2010a). *Job analysis: Air Force pilots (jet, rotary wing, and multi-engine streams* (DGMPRA TN 2010-013). Ottawa, ON: Director General Military Personnel Research & Analysis.

Darr, W. (2010b). *The Royal Air Force Aircrew Aptitude Test (RAFAAT): Preliminary evidence for validity* (DGMPRA TN 2010-015). Ottawa, ON: Director General Military Personnel Research & Analysis.

Deitcher, J. J., & Johnston, P. J. (2004). *Validation of the Canadian Forces pilot selection system* (Air Personnel Research Report 04/2). Ottawa, ON: Chief of the Air Staff.

Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology, 41*, 417-440.

Dockeray, F. C., & Isaacs, S. (1921). Psychological research in aviation in Italy, France, England, and the American Expeditionary Forces. *Comparative Psychology, 1*, 115-148.

Donohue, J. J. (2005). Validation of the 2004 CFAT percentile Norms (Technical Note 2005-01). Ottawa, ON: Director Human Resources Research and Evaluation.

Drasgow, F., Nye, C. D., Carretta, T. R., & Ree, M. J. (2010). Factor structure of the Air Force Officer Qualifying Test Form S: Analysis and comparison with previous forms. *Military Psychology, 22*, 68-85, DOI: 10.1080/08995600903249255.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721-733). Cambridge, MA: MIT Press.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situational awareness. *International Journal of Aviation Psychology, 4*, 241-264.

English, A. D. (Ed.). (1996). "A good airman is worth a battalion of infantry": Aviator recruiting and selection, 1914-45. In, *The cream of the crop: Canadian aircrew, 1939-1945* (pp. 17-40). Quebec: McGill Queen's University Press

ExamCorp (2012). *ExamCorp*. Retrieved January 2, 2012, from http://www.examcorp.com/v5/about/who.shtml

Eysenk, H. J. (1991). Dimensions of personality: 16, 5, or 3? Criteria for a taxonomy paradigm. *Personality and Individual Differences, 12*, 773-790.

Eysenk, H. J., & Eysenk, S. B. G. (1964). *Manual of the Eysenk Personality Inventory*. London, UK: University of London Press.

Fatolitis, P. G., Jentsch, F. G., Hancock, P. A., Kennedy, R. S., & Bowers, C. (2010). *Initial validation of novel performance-based measures: Mental rotation and psychomotor ability* (NAMRL Monograph 10-52). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Fitts, P., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks Coleman.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from difference sources. *Journal of Abnormal and Social Psychology, 44*, 329-344.

Framingham, J. (2011). Minnesota Multiphasic Personality Inventory (MMPI). *Psych Central*. Retrieved on January 22, 2012, from http://psychcentral.com/lib/2011/minnesota-multiphasic-personality-inventory-mmpi/

Fuster, J. M. (1997). *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe* (3rd ed.). Philadelphia, PA: Lippincott-Raven.

Ghez, C. & Krakauer, J. (2000). The organization of movement. In E.R. Kandel, J.H. Schwartz & T.M. Jessell (Eds.), *Principles of Neural Science* (4th ed.). (pp. 653-673). New York: McGraw-Hill.

Gnan, M., Flynn, C. F., & King, R. E. (1995). *Psychological pilot selection in the U.S. Air Force, the Luftwaffe, and the German aerospace research establishment* (Technical Report 1995-0003). Brooks AFB, TX: Armstrong Laboratory, Aerospace Medicine Directorate.

Goffin, Rothstein, & Johnson (1996). Personality testing and the assessment centre: Incremental validity for managerial selection. *Journal of Applied Psychology, 81*, 746-756.

Griffin, G. R., & Koonce, J. M. (1996). Review of psychomotor skills in pilot selection research of the U.S. military services. *The International Journal of Aviation Psychology, 6*(2), 125-147.

Hardy, D. J., Satz, P., D'Elia, L., & Uchiyama, C. L. (2007). Age-related group and individual differences in aircraft pilot cognition. *International Journal of Aviation Psychology, 17*(1), 77-90.

Haynes, S. N. & Lench, H. C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment, 15*(4), 456-466.

Higgins, D. M. (2009). *The roles of psychometric intelligence and prefrontal cognitive ability in the prediction of academic and job performance* (Doctoral dissertation). Retrieved from Dr. Jordan Peterson, December 2011.

Higgins, D.M., Peterson, J.B., Lee, A. and Pihl, R.O. (2007). Prefrontal cognitive ability, intelligence, Big Five personality and the prediction of advanced academic and workplace performance. *Journal of Personality and Social Psychology, 93*, 298-319.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *Personality – current theory & research: Nebraska symposium on motivation*. Lincoln, NE: University of Nebraska Press.

Hogan, R. (1986). *Manual for the Hogan Personality Inventory*. Minneapolis, MN: National Computer System.

Hogan, J. & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. Johnston, & S. Briggs (Eds.), *Handbook of Personality Psychology*. San Diego, CA: Academic.

Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology, 27*, 437-485.

Horn, J. L. & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253-270.

Hough, L. M. (1992). The "Big Five" personality variables - construct confusion: Description versus prediction. *Human Performance, 5*, 139-155.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*, 209-244.

Hough, L. M. (1998). Personality at work: Issues and evidence. In M. D. Hakel (Ed.), *Beyond multiple choice: Evaluating alternatives to traditional testing for selection* (pp. 131-166). Mahwah, NJ: Lawrence Erlbaum Associates.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.

Hunter, D.R., & Burke, E. F. (1994). Predicting aircraft pilot training success: A meta-analysis of published research. *The International Journal of Aviation Psychology, 4*(4), 297-313.

Hunter, D.R., & Burke, E.F. (1995). *Handbook of pilot selection.* Brookfield, VT: Avebury Aviation.

Hunter, J.E. & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96*(1), 72-98.

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and Job Performance: Economic and Social Implications. *Psychology, Public Policy, and Law, 2*(3/4), 447-472.

Hussey, A. K. (2004). *Air Force flight screening: Evolutionary changes, 1917 – 2003.* San Antonio, TX: Office of History and Research, Randolph Air Force Base.

Ibel, R. H., & Cotton, C. A. (1994). *Validation of experimental scales for Non-Commissioned Member applicant selection and classification* (Technical Note 17/94). Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd ed.). Port Huron, MI: Research Psychologists.

Johnston, P. J., & Catano, V. M. (2010). *Investigating the validity of previous flying experience, both actual and simulated, in predicting initial and advanced military pilot training performance.* Paper presented at the 52nd Conference of the International Military Testing Association.

Kaplan, H. (1965). *Prediction of success in Army aviation training* (Technical Report 1142). Washington, DC: U.S. Army Personnel Research Office.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General, 122*, 411–428.

Kragh, V. (1960). The Defense Mechanism Test: A new method for diagnosis and personnel selection. *Journal of Applied Psychology, 44*, 303 – 309.

Levine, E. L. (1983). *Everything you always wanted to know about job analysis.* Tampa, FL: Mariner

Levine, E. L., Spector, P. E., Menon, S., & Narayanan, L. (1996). Validity generalization for cognitive, psychomotor, and perceptual tests for craft jobs in the utility industry. *Human Performance, 9*(1), 1-22.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology, 94*(6), 1514-1535.

Long, G. E. & Varney, N. C. (1975). *Automated pilot aptitude measurement system* (Tech. Rep. No. AFHRL-TR-75-58). Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division.

Martinussen, M. (1996). Psychological measures as predictors of pilot performance: A meta-analysis. *The International Journal of Aviation Psychology, 6*, 1-20.

Martinussen, M. & Hunter, D. R. (2010). *Aviation psychology and human factors*. Boca Raton, FL: CRC Press.

McCrae, R. R. & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*, 710-721.

McGrew, K. S. (2005). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.)(pp. 136–181). New York: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1-10

Miller, E. K., & Wallis, J. D. (2009). Executive function and higher-order cognition: Definition and neural substrates. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (Vol. 4) . (pp. 99-104). Oxford: Academic Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howeter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Morrow, D. G., & Rodvold, M. (1998). Issues in air traffic control communication. In M. Smolensky & E. Stein (Eds.), *Human factors in air traffic control* (pp. 421-456). New York: Academic.

Morrow, D. G., Menard, W. E., Ridolfo, H. E., Stine-Morrow, E. A. L., Teller, T., & Bryant, D. (2003). Expertise, cognitive ability, and age effects on pilot communication. *The International Journal of Aviation Psychology, 13*(4), 345-371.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resources Management, 13*, 153-200.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165.

NATO Training Group (2008). Data manual 2008. *Human Performance in Military Aviation Working Group (HPMAWG).*

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.

Norman, W. T. (1967). *2, 800 personality trait descriptors: Normative operating characteristics for a university population* (Research Report 08310-1-T). Ann Arbor, MI: University of Michigan.

Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1-18). New York: Plenum.

North, R. A., & Griffin, G. R. (1977). *Aviation selection 1919-1977* (Special Report No. 77-2). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Okros, A. C., Spinner, B., & James, J. A. (1991). *The Canadian Automated Pilot Selection System* (Research Report 91-1). Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

O-Net OnLine (2010a). *53-2012.00 Commercial pilot.* Retrieved January 5, 2012 from http://www.onetonline.org/link/summary/53-2012.00

O-Net OnLine (2010b). *53-2011.00 Airline pilots, co-pilots, and flight engineers.* Retrieved January 5, 2012 from http://www.onetonline.org/link/summary/53-2011.00

Pelchat, D. W. (1999). *Analysis of the Canadian Automated Pilot Selection System (CAPSS): Findings from the first two years of operation* (Report 99-12). Ottawa, ON: Director Human Resources Research and Evaluation.

Peterson, J. B. (2003). *The prefrontal cortex, personality, performance, and economic value: Summary and research appendices.* (Technical Report). Toronto, ON: University of Toronto.

Peterson, J. B. (2010). ExamCorp Assessment Process. *ExamCorp White Paper.*

Ree, M. J. & Carretta, T. R. (1996). Central role of g in military pilot selection. *The International Journal of Aviation Psychology, 6*(2), 111-123

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology, 80*, 721-730.

Ree, K. J. & Earles, J. A. (1991). Predicting training success: Not much more than *g*. *Personnel Psychology, 44*(2), 321-332.

Ree, M. J. & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86-89.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79*(4), 518-524.

Robins, R.W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality correlates of self-esteem. *Journal of Research in Personality, 35*, 463-482.

Rothstein, M. G. & Jelly, R. B. (2003). The challenge of aggregating studies of personality. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 223-262). Mahwah, NJ: Lawrence Erlbaum.

Rothstein, M. G. & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resources Management Review, 16*, 155-180.

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Rummans, T. A., Kaufer, D. I., LaFrance, W. C., & Coffey, C. E. (2002). Executive control function: A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences, 14*(4), 377-405.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82*, 30-43.

Seamster, T. L., Redding, R. E., & Kaempf, G. L. (1997). *Applied cognitive task analysis in aviation.* Brookfield, VT: Ashgate.

Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274.

Shallice, T., & Burgess, P.W., (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 11 4*, 727-741.

Siem, F. M. (1992). Predictive validity of an automated personality inventory for Air Force pilot selection. *The International Journal of Aviation Psychology, 2*(4), 261-270

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analyses of form O* (AFHRL Tech. Rep. No. 86-68). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Skoff, B. (2004). Executive functions in developmental disabilities. *Insights on Learning Disabilities, 15*(2), 4-10.

Skomorovsky, A. & Donohue, J. (2010). *Predicting training performance for military pilots using Canadian Automated Pilot Selection System and Canadian Forces Aptitude Test* (DGMPRA TM 2010-006). Ottawa, ON: Directorate of Military Personnel Research and Analysis.

Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychology Measurement, 27*, 967-984

Society for Industrial and Organizational Psychology, Inc. (2003). *Principles for Validation and Use of Personnel Selection Procedures, 4th ed.*, College Park, MD: Author. Retrieved from http://www.siop.org/_Principles/principles.pdf

Society for Industrial and Organizational Psychology, Inc. (2012). *Employment testing overview*. Retrieved from http://www.siop.org/Workplace/employment%20testing/overview.aspx

Southcote, A. (2004). *Officer and aircrew selection centre aptitude test manuals* (Psychologist Report 04/04). RAF College Cranwell, UK: Directorate of Recruiting and Selection (Royal Air Force).

Southcote, A. (2007). Assessing the feasibility of a tri-service selection aptitude test battery. In P. Murphy, G. Fogarty, & J. Swann (Eds.), *Select Proceedings of the 49th Annual Conference of the International Military Testing Association, Gold Coast, Australia, October 8-12, 2007*. Retrieved from http://www.internationalmta.org/Documents/2007/Proceedings2007.pdf

Spangler, S. (2009, March 24). UAV pilot shortage and military intelligence. *Jetwhine, 3*. Retrieved September 25, 2011 from http://jetwhine.com/2009/03/uav-pilot-shortage-military-intelligence/

Spearman, C. (1904). "General Intelligence", objectively determined and measured. *The American Journal of Psychology, 15*(2), 201-292.

Spinner, B. (1988). *Using the Canadian Automated Pilot Selection System to predict performance in primary flying training: Straight and level flight* (Technical Note 15/88). Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Spinner, B. (1991). Predicting success in primary flying school from the Canadian Automated Pilot Selection System: Derivation and cross-validation. *The International Journal of Aviation Psychology, 1*(2), 163-180.

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*(1), 138-161.

Street, D. R., Jr., Helton, K. T., & Dolgin, D. L. (1992). *The unique contribution of selected personality tests to the prediction of success in naval pilot training* (Report No. NAMRL-1374). Pensacola, FL: Naval Aerospace Medical Research Laboratory

Stevens, J. (2002). Chapter 3. In *Author, Applied multivariate statistics for the behavioural sciences* (4th ed.). (pp. 80-172). Hillsdale, NJ: Erlbaum.

Stuss, D. T., Eskes, G. A., & Foster, J. K. (1994). Experimental neuropsychological studies of frontal lobe functions. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9). (pp. 149–185). Amsterdam: Elsevier Science.

Suarez, J., Barborek, S., Nikore, V., & Hunter, D. R. (1994). *Current trends in pilot hiring and selection.* (AAM-240-94-1). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education Inc.

Taylor, J. L., Noda, A. M., O'Hara, R., Mumenthaler, M. S., Kraemer, H. C., & Yesavage, J. A. (2000). *Longitudinal study of older aircraft pilots.* Paper presented at the Cognitive Aging Conference 2000, Atlanta, GA.

Tett, R. P., Jackson, D. N., & Rothstein, M. G. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*, 703-742.

Tett, R. P., Jackson, D. N., Rothstein, M. G., & Reddon, J. R. (1994). Meta-analysis of personality-job performance relations: A reply to Ones, Mount, Barrick, and Hunter (1994). *Personnel Psychology, 47*(1), 157-172.

Thorndike, E. L. (1919). Scientific Personnel Work in the Army. *Science, XLIX* (1255), 53-61.

Tsang, P.S., & Vidulich, M.A. (Eds.). (2008). *Principles and practices of aviation psychology* (pp. 357-396). Mahwah, NJ: Lawrence Erlbaum.

Tupes, E. C. & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (ASD-TR-61-97). Lackland Air Force Base, TX: Aeronautical Systems

Division, Personnel Laboratory.

United States General Accounting Office. (1999). *Military personnel: Actions needed to better define pilot requirements and promote retention* (Report No. GAO/NSIAD-99-211). Report to the Chairman and Ranking Minority Member, Subcommittee on Military Personnel, Committee on Armed Services, House of Representatives. Washington, DC: Author. Retrieved September 25, 2011 from www.goa.gov/archive/1999/ns99211.pdf

Wei, J. & Salvendy, G. (2004). The cognitive task analysis methods for job and task design: Review and reappraisal. *Behavior and Information Technology, 23*(4), 273-299.

Werrel, K. (2007). Aces and -86s: The fight for air superiority during the Korean War. In J. Neufeld & G. M. Watson Jr. (Eds.), *Coalition air warfare in the Korean War, 1950-1953: Proceedings, Air Force Historical Foundation Symposium, Andrews AFB, Maryland, May 7-8, 2002*. Retrieved from http://www.afhso.af.mil/shared/media/document/AFD-100924-016.pdf

Wheeler, J. L., & Ree, M. J. (1997). The role of general and specific psychomotor tracking ability in validity. *International Journal of Selection and Assessment, 5*(2), 128-136.

Wickens, C. (2007). Aviation. In F. T. Durso (Ed.), *Handbook of applied cognition* (2nd ed.). (pp. 361-389). West Sussex, England: John Wiley & Sons, Ltd.

Wickens, C. D., & Carswell, C. M. (2007). Human information processing. In G. Salvendy (ed.), *The Handbook of Human Factors and Ergonomics* (3rd ed.). New York: John Wiley & Sons.

Wiggins, J. S. (1968). Personality structure. *Annual Review of Psychology, 19*, 293-350.

Woycheshin, D. (1999, November). *CAPSS: The Canadian Automated Pilot Selection System.* Paper presented at the meeting of the North Atlantic Treaty Organization, Research and Technology Agency held in Monterey, CA.

Woycheshin, D. (2001). *Analysis of Canadian Automated Pilot Selection System (CAPSS) results: February 1998 to March 2000* (Air Personnel Research Report 01/1). Ottawa, ON: Chief of Air Staff.

Woycheshin, D. (2002). Validation of the Canadian Automated Pilot Selection System (CAPSS) Against Primary Flying Training Results. *Canadian Journal of Behavioural Science, 34*(2), 84-91.

Yerkes, R. M. (1919). Report of the psychological committee of the National Research Council. *Psychological Review, 26*, 83-149.

Zelazo, P. D., Carter, A., Reznick, J. S., and Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*, 198–226.

Appendices

Appendix A: Combination of Pilot Tasks and KSAOs From O-Net (O-Net Online, 2010a, 2010b) and Canadian Forces Job Analyses (Darr, 2010), Including the Associated Cognitive, Psychomotor, and Personality Domains.

	Cognitive Domain				Psycho-motor Ability	Personality				
	Gc	Gf	Specific Cognitive Abilities	Executive Functions		Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Tasks										
complete check lists accurately, while responding appropriately to interruptions			x	x						
use instrumentation to guide flights when visibility is poor [a]	x	x		x						
respond to and report in-flight emergencies and malfunctions [a]		x		x						x
work as part of a flight team with other crew members, especially during take-offs and landings [a]						x	x	x	x	
contact control towers for takeoff clearances, arrival instructions, and other information, using radio equipment [a]				x						

respond to tower instructions (repeat instructions) and take appropriate action			x	x						
steer aircraft along planned routes, using autopilot and flight management computers [a]	x			x						
monitor gauges, warning devices, and control panels to very aircraft performance and to regulate engine speed [a]			x	x						
monitor airspace and gauges to ensure accurate situational awareness		x		x						
start engines, operate controls, and pilot aircraft to transport passengers, mail, or freight, adhering to flight plans, regulations, and procedures [a]	x			x						
inspect aircraft for defects and malfunctions, according to pre-flight checklists [a]		x		x						
check passenger and cargo distributions and fuel amounts to ensure that weight and balance specifications are met [a]	x	x		x						
monitor engine operation, fuel consumption, and functioning of aircraft systems during flight [a]	x	x		x						
Knowledge										
transportation - aircraft operations and flight [a]	x	x								
public safety and security [a]										
mathematics [a]	x	x								
English language [a]	x	x								
geography - describing features of land, sea and air masses, including their physical characteristics, locations, interrelationships [a]		x								
physics [a]	x	x								

knowledge of high school math [b]	x	x								
Skills										
operation & control - controlling operations of equipment or systems [a]	x	x			x					
operation monitoring - watching gauges, dials, or other indicators to make sure a machine is working properly [a]		x	x	x	x					
critical thinking - using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions, or approaches to problems [a]		x		x						
judgement & decision making - considering the relative costs and benefits of potential actions to choose the most appropriate one [a b]		x		x						
complex problem solving - identifying complex problems and reviewing related information to develop and evaluate options and implement solutions [a b]		x		x						
perform basic mental calculations under pressure [b]	x	x								x
active listening - giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times [a]		x				x			x	x
coordination - adjusting actions in relation to others' actions [a]				x					x	x
reading comprehension - understanding written sentences and paragraphs in work related documents [a b]	x	x		x						
speaking - talking to others to convey information effectively [a]		x		x						

monitoring - monitoring/assessing performance of yourself, other individuals, or organizations to make improvements or take corrective actions [a]		x		x						
situation awareness (monitoring) - scanning gauges / instruments and outside environment in order to identify potential issues	x	x		x	x					
Abilities										
control precision - the ability to quickly and repeatedly adjust the controls of a machine or a vehicle to exact positions [a]		x			x					
problem sensitivity - the ability to tell when something is wrong or is likely to go wrong (involves recognizing there is a problem) [a]		x		x		x				
ability to 1) anticipate and 2) prioritize activities so that they happen within a spec time period [b]				x						
ability to focus on task/handle complexity under pressure [b]				x					x	x
ability to manage time and deal with time pressures [b]				x			x		x	x
response orientation - the ability to choose quickly between 2 or more movements in response to 2 or more different signals (lights, sounds, pictures). it includes the speed with which the correct response is started with the hand, foot, or other body part [a]				x	x					
depth perception - the ability to judge which of several objects is closer or farther away from you, or to judge the distance between you and an object [a]										

spatial orientation - the ability to know your location in relation to the environment or to know where other objects are in relation to you [a]			x	x	x					
far vision - the ability to see details at a distance										
Multi-limb coordination - the ability to coordinate 2 or more limbs (e.g., 2 arms, 2 legs or 1 arm and 1 leg) while sitting, standing, or lying down [a b]					x					
hand-eye-foot coordination [b]					x					
divided attention [b]				x						
attend to multiple stimuli [b]				x						
rate control - the ability to time your movements or the movement of a piece of equipment in anticipation of changes in the speed and/or direction of a moving abject or scene [a]				x	x					
reaction time - the ability to quickly respond (with finger, hand, or foot) to a signal (sound, light, picture) when it appears [a]					x					
oral comprehension - the ability to listen to and understand information and ideas presented through spoken words and sentences [a]	x	x		x						
oral comprehension & memory - the ability to comprehend and oral instructions given by the ATC, accurately repeat those instructions, and to accurately perform actions as required [b]				x						
maintain calm during stressful situations [b]				x						x
mental stability [b]										x
handle baggage interference [b]		x		x						x

Work Activities										
running, manoeuvring, navigating, or driving aircraft [a]	x	x	x	x	x		x			
making decisions & solving problems - analysing information and evaluating results to chose the best solution and solve problems [a]	x	x	x	x		x	x			
getting information - observing, receiving, and otherwise obtaining information from all relevant sources [a]		x		x		x	x	x	x	x
monitoring and reviewing information from materiel, events, or the environment to detect or assess problems [a]										
analyzing data or information - identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts [a]		x		x			x			
Work Styles										
attention to detail [a]				x			x			
dependability [a]							x			
self control - maintaining composure, keeping emotions in check, controlling anger, and avoiding aggressive behaviour, even in very difficult situations [a]				x						x
stress tolerance - accepting criticism and dealing calmly and effectively with high stress situations [a]										x
leadership - willingness to lead, take charge, and offer opinions and direction [a]								x		
cooperation - being pleasant with others on the job and displaying a good-natured, cooperative attitude [a]							x		x	x

analytical thinking - analyzing information and using logic to address work-related issues and problems [a]		x		x						
integrity - being honest and ethical [a]										
adaptability/flexibility - open to change (positive or negative) and to considerable variety in the workplace [a]		x		x		x				
persistence - persistence in the face of obstacles [a]							x			
independence [a]							x			
confidence [b]										x

Note:
[a] O-Net Online Job Analysis (2010a, 2010b)
[b] Canadian Forces Job Analysis (Darr, 2010)

Appendix B: Figures of CAPSS Simulator and RAFAAT Apparatus Set-up and Sub-tests

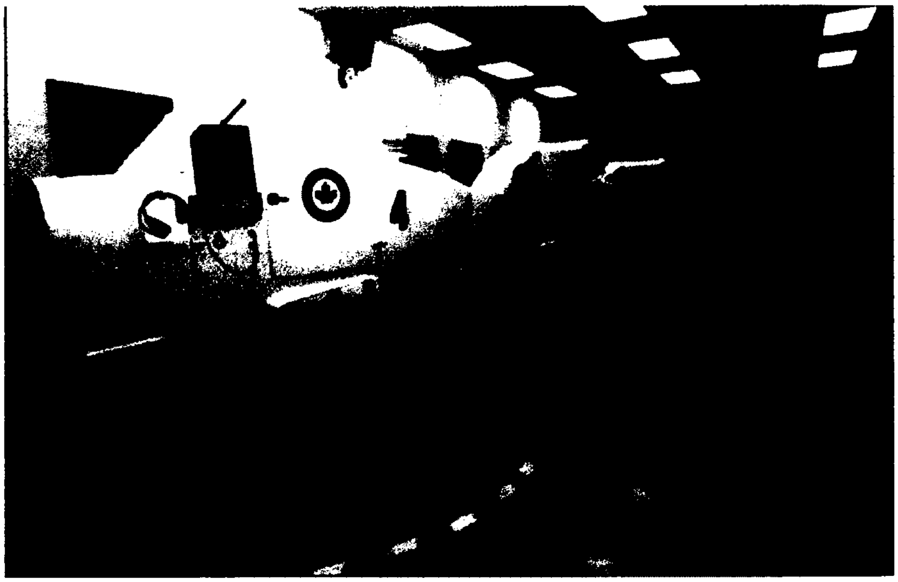


Figure 1. The Canadian Automated Pilot Selection System (CAPSS) at the Aircrew Selection Centre in Trenton, Ontario.

Figure 2. The Royal Air Force Aircrew Aptitude Test (RAFAAT) battery apparatus, as set up at 3 Canadian Forces Flying Training School (3 CFFTS), in Portage La Prairie, Manitoba.

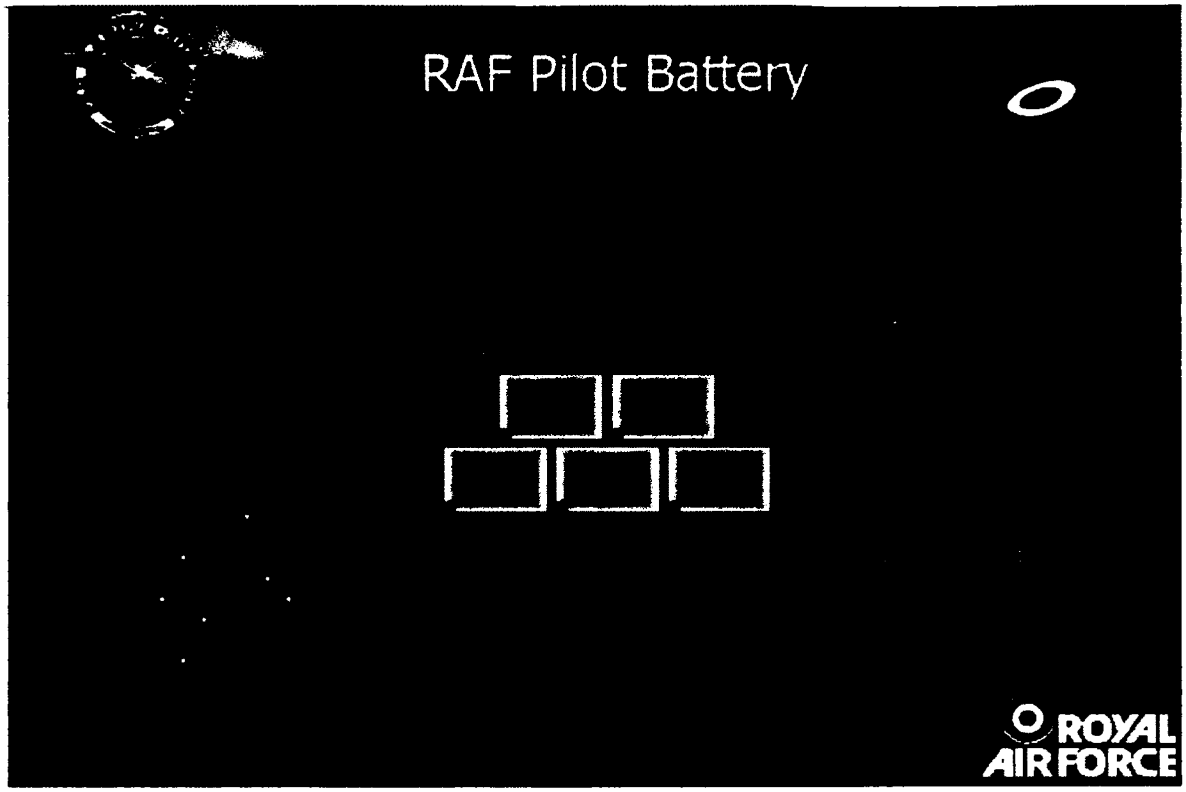


Figure 3. A small sample of the screen displays for five RAFAAT sub-tests. This display includes sensory motor apparatus (top left), control of velocity test (top right), vigilance (bottom left) and digit recall (bottom right). The instrument comprehension test (middle) was not included in the Canadian Forces battery due to overlap with the work measure, the Canadian Automated Pilot Selection System (CAPSS) simulator.

Appendix C: Participant Recruiting Brief

SLIDE 1

Good morning everyone. I'm Lieutenant Commander Dawn Herniman. I'm a Personnel Selection Officer and I'm here to conduct research that the Canadian Forces and the Air Force want to investigate in order to improve the pilot selection process.

SLIDE 2: WHY IS IT IMPORTANT TO PARTICIPATE?

It costs the Canadian Forces approximately $3 million to train a pilot up to wings qualification. Because of this significant training cost, and the enormous financial impacts attrition causes, we want to make sure the Canadian Forces has the best Pilot selection system possible. Ensuring a better selection system is not only a cost saving measure, but it improves pilot training through-put by helping to streamline the training. **Your participation in this study is voluntary. However, your involvement is critical to help the Canadian Forces improve the Pilot selection system. We can't do it without you.**

Now I'm going to explain a little bit about the study.

SLIDE 3: AIM

Title "Investigating Predictors of Primary Flight Training Performance: Potential Impacts for CF Pilot Selection"

I'm investigating two domains that research indicates are essential to Pilot performance, and that the Air Force wants to explore further.

SLIDE 4: EXAMCORP

The ExamCorp battery measures both the executive functions and personality domains.

Cognitive executive functions. Components of executive functions include attentional control, working memory, task prioritization, decision-making, and action/response initiation and inhibition. These types of executive function tasks are critical to pilot aviate, navigate, communicate, and systems monitoring and management tasks. The second domain is **Personality**. Previous research has indicated that there is a link between personality characteristics and pilot performance.

SLIDE 5: OTHER MEASURES

I also want to see how the ExamCorp battery stacks up against current selection tests, including the Canadian Forces Aptitude Test (CFAT), Canadian Automated Pilot Selection System (CAPSS), and the Royal Air Forces Aircrew Aptitude Tests (RAFAAT). The performance measures that will be used to judge whether these measures can predict pilot performance will be your PFT training data. If you choose to participate, I will require your authorization to access to your course report and individual flight assessments, as well as your archived selection tests.

SLIDE 6: RAFAAT

If you did not complete the RAFAAT battery during Aircrew Selection, I also ask that you complete this test. However, I understand that your time is precious, and the main focus is on completing the ExamCorp.

SLIDE 7: OTHER FACTORS/INFLUENCES

In addition to executive functions and personality, I'm going to investigate how PFT performance, ExamCorp performance, and RAFAAT results are influenced by other factors such as education background, previous flying experience, video game experience, and training-related anxiety.

SLIDE 8: INFORMED CONSENT

Every study conducted that involves an experimental design requires Informed Consent. Informed Consent includes the disclosure to you of the purpose of the study, an explanation of how I will ensure your confidentiality, and emphasis on the fact that your participation is voluntary. This information is provided in greater detail in the Informed Consent Form, which I will ask you to sign if you choose to participate.

SLIDE 9: CONFIDENTIALITY

I want to emphasize, again, that your participation in this study is voluntary and that the test results will not be provided to your chain of command or training staff, and it will not affect you career in any way. I will not inform your chain of command of who does or does not participate, but the total number of participants over the 7 PFT classes involved in this study will be reported in the final analysis. I urge you not to discuss the project, and especially the test measures, so as not to compromise the process and results. Furthermore, please do not identify who is or is not participating, in order to ensure as much confidentiality to yourselves as possible.

SLIDE 10: WITHDRAWING

You can withdraw at any time and the data that you have completed to that point will be destroyed.

SLIDE 11: SCHEDULE

Testing will be conducted in the conference room, Tuesday, Wednesday, and Thursday. In order to be flexible to your flying schedule, <u>ExamCorp</u> sessions can commence at any time from 08:00-11:00 and 13:00-16:00. However, for those of you who haven't completed the RAFAAT and agree to do so, <u>RAFAAT</u> sessions will be conducted twice daily, at 08:15 and 12:45.

SLIDE 12: QUESTIONS? My contact information.

SLIDE 13: PARTICIPATION

Again, I want to emphasize two things:

1. there will be no impact on your career or training outcome whether you decide to participate in this study or not, and
2. your participation is voluntary.

If you wish to participate, I have some information you can complete now. However, if you are not interested in participating in this study, you may leave the room now.

I will now take the name, initials, and service number for each of you and give you a participant identification number. After that I will ask you to read and sign the Informed Consent Form.

We can also take 10 minutes to complete the Demographics Questionnaire, so that only the computer-based testing is left to complete.

Appendix D: Participant Informed Consent

Participant Informed Consent

The purpose of an informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. The informed consent has to provide sufficient information such that you have the opportunity to determine whether or not you wish to participate in the study.

Study Title: Investigating Predictors of Primary Flight Training: Potential Impacts for CF Pilot Selection

Purpose: The purpose of this study is to improve Pilot selection for the Canadian Forces. To do so, this study will assess the appropriateness of a potential Pilot selection test called ExamCorp. As such, the study will assess the relationships between ExamCorp and the other selection measures, including the Canadian Forces Aptitude Test (CFAT), the Canadian Automated Pilot Selection System (CAPSS), and the Royal Air Force Aircrew Aptitude Test (RAFAAT), and performance on Primary Flight Training. Additionally, a number of other demographic factors will be investigated in order to determine the relationships with performance on selection tests and Primary Flight Training.

Task Requirements: You will first be asked to fill out a demographic questionnaire regarding your background. Next, you will be asked to complete the ExamCorp test, which is a computer-based test that takes approximately 90 minutes to complete. If you haven't yet completed the RAFAAT as part of your selection testing, an additional 3.5 hours of testing is required to complete the test. In order to conduct proper analyses we will also match your ExamCorp and RAFAAT tests with your archived selection tests, which include the Canadian Forces Aptitude Test (CFAT), Canadian Automated Pilot Selection System (CAPSS). As the hypotheses involve improving Pilot selection, the intent of this study is to assess the ability of these test measures to predict performance on Primary Flight Training (PFT). Therefore, your consent is also required to release the results of your PFT course performance (i.e. PFT course report).

Potential Risk and Discomfort: There are no physical or emotional risks and discomforts to participating in this study.

Research Coordination Number

This study has received clearance by the Carleton University Psychology Research Ethics Board (Reference # 11-101). This study has also been coordinated through the DGMPRA Social Science Research Review Board (Approval # 997/11-F), in accordance with CANFORGEN 198/08.

Anonymity/Confidentiality: Every effort will be taken to ensure your confidentiality and safeguard your anonymity. As we must match your results with training and other section results (i.e., CFAT, RAFAAT, CAPSS) you will be provided with an identification number that will be used to complete all materials for this study. Consequently, your name, initials and service number will be recorded on a master participant list that matches your personal information with a participant identification number. Aside from the Participant Informed Consent form, you will use only your participant identification number to identify yourself on all other forms and the ExamCorp battery, thus ensuring your confidentiality is protected. It should be noted that your ExamCorp test results will be downloaded from the laptop to a memory stick at the end of each test session and forwarded electronically to the principal researcher's account on the ExamCorp server for analysis. However, as you will be identified only by your participant identification number, your confidentiality will not be compromised. Your RAFAAT test results will also be downloaded from the laptop to a memory stick at the end of the test session, and the results will go into the CF RAFAAT test archives. Once all data has been matched (i.e. tests have been completed or retrieved from archives and course training performance is provided), your personal information, including your name, initials, and service number, will be removed from the dataset. While there is the remote possibility that someone might be able to deduce your identity based on some combination of the demographic questions, all individual level data will be kept strictly confidential. The final analysis and report will only involve aggregate data, which will not include any identifying information. During all phases of this study, all discs, memory sticks, forms, and data that identify your personal information will be safeguarded by me, in my capacity as the principal investigator, in accordance with the Canadian Privacy Act. Furthermore, upon completion of this study, all relevant data will be archived and secured in a regulation Protected B file cabinet at Director General Military Personnel Research and Analysis (DGMPRA) in accordance with governmental archival policies and the Canadian Privacy Act. However, only the current research team will have access to the data.

Research Personnel:
The following people are involved in this research project and may be contacted at any time if you have questions or concerns:
LCdr Dawn Herniman, email: dhernima@connect.carleton.ca
Dr. Avi Parush, 613-520-2600, ext. 6026, email: avi_parush@carleton.ca

Ethical Concerns:
Should you have any ethical concerns about this study, please contact:
Dr. Monique Sénéchal (Chair, Carleton University Ethics Committee for Psychological Research, 613-520-2600, ext. 1155, email: monique_senechal@carleton.ca); or Dr. Anne Bowker, (Chair, Department of Psychology, Carleton University, 613-520-2600, ext. 8218, email: psychchair@carleton.ca).
The Canadian Forces contact for ethical concerns is Colin Kemp, email: colin.kemp@forces.gc.ca

ATIP Considerations
I am aware that under the Access to Information Act, Canadian citizens are entitled to obtain copies of research reports and research data (including the database pertaining to this project) held in Federal government files. Similarly, under the Privacy Act, Canadian citizens are entitled to copies of all information concerning them that is held in Federal government files including research databases. Prior to releasing requested information, the Directorate of Access to Information and Privacy (DAIP) screens the data to ensure that individual identities are not disclosed.

Right to Withdraw: Your participation in this study is entirely voluntary. At any point during the study, you have the right to refrain from answering certain questions or to withdraw from the study without penalty whatsoever. Should you choose to withdraw from this study, all information collected on you will be destroyed.

Acceptance
I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I

am also aware that the data gathered in this study are confidential and anonymous with respect to my personal identity. My signature indicates that I agree to participate in this study.

Participant's Name: ______________________________

Participant's Signature ______________________________ Date: ______________

Researcher's Name: ______________________________

Researcher's Signature: ______________________________ Date: ______________

Appendix E: Demographic Questionnaire

Demographic Questionnaire

Participant ID# __________

In order to provide information required for proper analysis and to assess possible selection related issues, please respond to the questions below by circling the correct response or filling in the blanks as appropriate. Completion of any question is considered voluntary.

1. What is your gender?
   a. Female
   b. Male

2. How old are you?
   a. 20 – 25
   b. 26 – 30
   c. 31 – 35
   d. 36 – 40
   e. 41 and older

3. What is the highest level of education you have completed?
   a. High school
   b. Community college or part of under graduate degree
   c. Undergraduate degree (i.e. Bachelor's)
   d. Graduate degree

4. Have you completed a Private Pilot's License?
   a. Yes – approximately how many hours have you logged? ___________
   b. No

5. On what Aircraft type do you have flying experience? (Indicate all that apply, and approximate hours logged, including PFT)
   a. Fixed wing – single engine: hours logged _____

b. Fixed wing – twin engine: hours logged _____

c. Fixed wing – jet: hours logged _____

d. Rotary wing – hours logged _____

6. How long has it been since you successfully completed your Private Pilot's License?

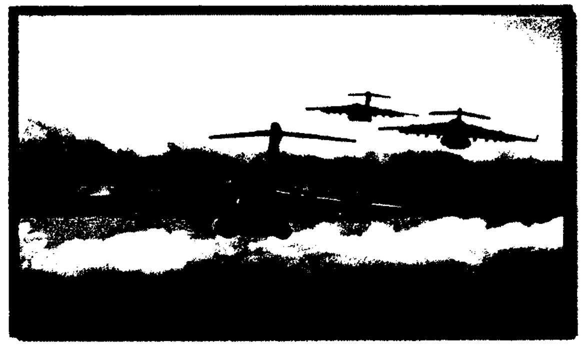
a. 6 – 12 months

b. 1 – 2 years

c. 3 – 5 years

d. 6 – 10 years

e. more than 10 years



7. Indicate how often you played each video game genre in the past 2 years:

Video Game Genre	I never play this video game genre	Less than 1 hour per week	1–3 hours per week	3–5 hours per week	5–7 hours per week	More than 7 hours per week
First Person Shooter (i.e. Medal of Honor, Halo, Battlefield)						
Role Playing Games (i.e. Mass Effect, Dragon Age, Elder Scrolls)						
Puzzle games (i.e. Tetris, Zuma, Bejeweled)						
Strategy games (i.e. Civilization, Shattered Galaxy, Dreamlords)						
Flight Simulator (i.e. Microsoft Flight Simulator, X-Plane, ProFlightSimulator)						

Appendix F: Participant Pre-Brief

(Verbatim Instructions)

Thank you for participating in the "*Investigating predictors of Primary Flight Training success: Potential impacts for CF Pilot selection*" study. Your involvement will help to improve the Canadian Forces pilot selection model. Your participation is critical for the administration of the demographics questionnaire, and the administration of the ExamCorp.

Demographics Questionnaire

We will start with the demographics questionnaire, which should take approximately 10 minutes to complete. When you have finished the questionnaire, please turn the page over so that I know you are done. I will then come by and pick them up. Once I have all the completed questionnaires we will take a short 5 minute break before starting the ExamCorp test.

Do you have any questions?

ExamCorp administration

Now you will complete the ExamCorp test. As a reminder, it will take approximately 90 minutes to complete this test. This time includes time for reading instructions and one short break. There are no tricks to the tests. Everything you need to know will be explained in the instructions on the screen. In each test you will see examples and be given an opportunity to practice where it is appropriate. It is perfectly normal for some people to finish these tests before others; it does not mean that one person's results will be better than another.

Apparatus. During the tests you will be using the mouse in front of you. Place the mouse wherever you like, making sure you can reach it comfortably. Please do not touch any other buttons or switches on the computer.

Nothing else will be required to complete these tests. Place any personal belongings, on the floor under your seat. If you are carrying a mobile phone or pager please switch it off now and for the duration of the testing.

You will notice some cabling lying on the floor connecting the test stations. Please do not trip or walk on the cables.

Breaks. You will have a 5 minute break at a pre-determined interval. You may leave the room during the break, but please be sure to return within the 5 minutes to ensure there is no delay to the testing. As the test is self-paced, please be very quiet as others may be working.

If you have any problem during the testing, raise your hand and I will come to help you. Please be as honest and accurate as possible. Some of the cognitive tests won't move on until you have achieved all the correct answers. Also, for the personality questions, please answer the questions honestly.

When you have completed the test, raise your hand. I will then give you the Participant Debrief Form to complete. If you do not have to complete the RAFAAT exams, your participation is considered complete. If you have not completed the RAFAAT test I ask you to return to complete that testing as well.

Do you have any questions?

Thank you for your participation!

Using the keyboard, enter your Participant ID Number instead of your name, as well as all your relevant information, to begin the testing.

RAFAAT Administration

The Royal Air Force Aircrew Aptitude Test (RAFAAT) is a battery of computer-based tests that are now being administered at the Canadian Forces Aircrew Selection Centre. There are a total of 11 tests and it will take you approximately 3.5 hours to complete them. This time includes time for reading instructions and two short breaks. There are no tricks to the tests. Everything you need to know will be explained in the instructions on the screen. In each test you will see examples and be given an opportunity to practice where it is appropriate. It is perfectly normal for some people to finish these tests before others; it does not mean that one person's results will be better than another.

Apparatus. During the tests you will be using the small keypad in front of you; it has all the functions you will require. Place the small keypad wherever you like, making sure you can reach it comfortably. Please do not touch the keyboard or any other buttons or switches on the computer. You also have a joystick in front of you that will be required for one test; when you are asked to use it, do not be afraid to move it across the full range horizontally and vertically. Finally, you will require a set of tables and some scrap paper for two other tests: they will be provided at that time. You will need to have a pen or pencil handy for that test.

Nothing else will be required to complete these tests. Place any personal belongings, on the floor under your seat. If you are carrying a mobile phone or pager please switch it off now and for the duration of the testing.

You will notice some cabling lying on the floor connecting the test stations. Please do not trip or walk on the cables.

Breaks. You will have two ten minute breaks at pre-determined intervals. Take that time to have a stretch and go to the washroom; however, please be very quiet as others will be working, and please be back on time so that we may stay on schedule.

If you have any problem during the testing, raise your hand and I will come to help you.

Finally, please give your best effort on these tests. The best way to do well is to relax and try to enjoy this testing session.

Are there any questions?

Press ANY NUMBER KEY on the keypad to begin the testing.

Appendix G:Participant Debrief

Participant Debrief

Thank you for participating in this important Pilot selection study. The necessary data has been collected from you as a subject matter expert. Although this concludes your active involvement in data collection, your involvement in this study permits us to access your archived selection test results (i.e. CFAT, CAPSS, and RAFAAT) as well as access to you final PFT course report.

What are we trying to learn in this research?

As indicated by the title on your informed consent, the purpose of this study is to assess the ExamCorp test as a potential pilot selection test. We also want to determine what combination of selection tests predict performance on Primary Flight Training, and could possibly replace the CAPSS simulator in the future.

Why is this important to scientists or the general public?

Pilot training in the Canadian Forces is extremely costly, with average training expenses totalling approximately $3 Million to train a pilot to wings qualification. Therefore, pilot selection testing is an expensive and time consuming endeavour. Militaries around the world strive to select the best pilot applicants possible, with the lowest costs to the organization. Given the demanding requirements, dynamic environment, and high training costs of the pilot occupation, a carefully designed and tested pilot selection system is imperative.

What are our hypotheses and predictions?

The purpose of this study is to assess the ExamCorp test as a potential pilot selection test, and investigate additional demographic factors that may predict performance on PFT and selection tests.

Should you choose to remain involved in this study, it is with the understanding that your data will be protected and not shared with anyone outside of the study including your chain of command. Although the results of your course performance will be known by your training staff, they will not be informed of your performance on the ExamCorp test or on any of the tests related to this study.

Where can I learn more?

There has been much research on cognitive and personality predictors of pilot performance. Here are a few examples:

Wickens, D. (2007). Aviation. In F. T. Durso (Ed.), Handbook of Applied Cognition: Second Edition, (pp. 361-389). Chichester, West Sussex: John Wiley & Sons, Ltd.

What if I have questions later?

If you have any remaining concerns, questions, or comments about the experiment, please feel free to contact:

Principal Investigator:

LCdr Dawn Herniman Tel. , dhernima@connect.carleton.ca

Carleton University Faculty Sponsor:

Dr. Avi Parush, Tel. 520.2600, ext. 6026, avi_parush@carleton.ca

Director Military Personnel Research and Analysis 2 Office of Primary Interest:

Colin Kemp, Tel. colin.kemp@forces.gc.ca

Director Air Personnel Strategy research staff officer:

Capt Edith Knight, Tel. edith.knight@forces.gc.ca

Ethical concerns:

Chair of Carleton University Ethics Committee for Psychological Research:

Dr. Monique Sénéchal, Tel. 613-520-2600, ext. 1155, monique_senechal@carleton.ca

Any other concerns:

Chair of the Department of Psychology, Carleton University:

Dr. Anne Bowker, Tel. 613-520-2600, ext. 8218, psychchair@carleton.ca

Research Coordination Number

This study has received clearance by the Carleton University Psychology Research Ethics Board (Reference # 11-101). This study has also been coordinated through the DGMPRA Social Science Research Review Board (Approval # 997/11-F), in accordance with CANFORGEN 198/08.

Thank you for participating in this research!