Validation of the Naval Officer Assessment Board

by

Lynne Carmen Marie Poirier

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Applied Psychology.

September, 2010, Halifax, Nova Scotia

Approved :	Dr. Victor Catano Supervisor
Approved:	Dr. E. Kevin Kelloway Committee Member
Approved:	LCdr F.A.J. Boyes Committee Member
Approved:	Dr. Damian O'Keefe External Examiner
Date:	September 8, 2010

Your file Votre référence
ISBN: 978-0-494-72011-0
Our file Notre référence
ISBN: 978-0-494-72011-0

Table of Contents

List of Appendices

List of Tables

List of Figures

Acknowledgements

It takes the support of a small army to complete a Masters thesis. Thankfully, I was blessed with a top-notch crew. I would first like to thank my husband, Martin, who uprooted his work and life to move to Halifax; who did a disproportionate share of the cooking and laundry over the past two years; who never stopped believing in me; and who provided timely reminders to 'just keep swimming'.

Next, I offer my thanks to my thesis advisor, Dr. Victor Catano, for sharing his wisdom, for giving me opportunities to learn, and for always leaving his door open to chat without notice. I would also like to thank my thesis committee members, Dr. Kevin Kelloway and Lieutenant-Commander Fraser Boyes, for their sound and practical advice.

This thesis would not have been possible without the efforts of several individuals who purged the archives for the data I needed. They include: Lieutenant-Commander Fraser Boyes, Major Cindy Suurd, Lieutenant-Navy Darren Hill, the Selection and Assessment section under the Director General Military Personnel Research and Analysis, and the standards section at the Naval Officer Training Centre.

Finally, I thank my family and closest friends: Mom, Dad, Soo, Jen, and Nicole. Thank you for your never-ending support and love.

Abstract

Validation of the Naval Officer Assessment Board

By Lynne Carmen Marie Poirier

Abstract: Using a construct-oriented approach, the current study examined the predictive validity of the Canadian Forces Recruiting Centre measures and the Naval Officer Assessment Board measures, which constitute a multistage selection system designed to select Maritime Surface and Subsurface officers for the Canadian Forces. The results showed that cognitive ability at stage one and information-processing at stage two were valid predictors of Phase III training performance. Only information-processing was a valid predictor of Phase IV training performance. The Naval Officer Assessment Board measures at stage two of selection showed incremental predictive validity over the Canadian Forces Recruiting Centre measures at stage one of selection for both Phase III and Phase IV training. Implications and recommendations for Maritime Surface and Subsurface officer selection are discussed.

September 8, 2010

Validation of the Naval Officer Assessment Board

Distinguishing between constructs and methods in predictive validation studies is critical to obtaining meaningful research results (Arthur & Villado, 2008). Constructs are used to predict job performance; they are the crux of predictive validity. Methods refer to the processes or techniques used to collect predictor information. The *Principles for the Validation and Use of Personnel Selection Procedures* (Society for Industrial and Organizational Psychology [SIOP], 2003) specifically call for the separation of constructs and methods to avoid confounds in the interpretation of validation study results.

The purpose of the current study is to examine the criterion-related validity of a multistage selection process using a construct-oriented approach. More specifically, the present study will re-examine the contribution of constructs within the Canadian Forces Recruiting Centre selection stage and within the Naval Officer Assessment Board selection stage in predicting performance on Maritime Surface and Subsurface officer training phases. In addition, the study aims to determine whether constructs measured in the second selection stage (i.e. during the Naval Officer Assessment Board) predict success on Maritime Surface and Subsurface training phases above constructs measured in the first selection stage (i.e. at Canadian Forces Recruiting Centres).

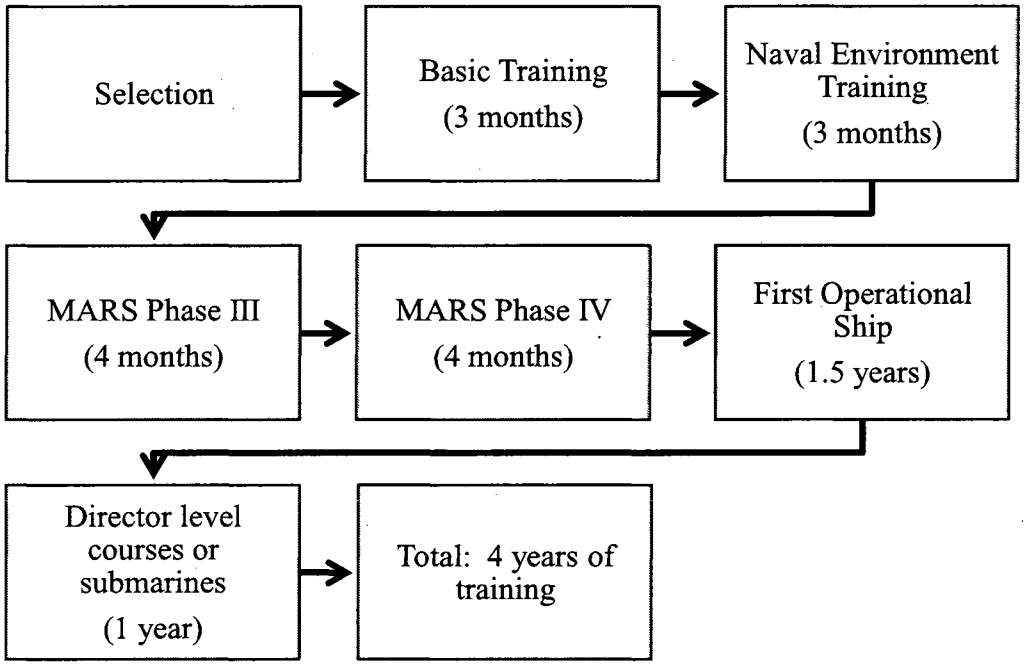
Maritime Surface and Subsurface officers are responsible for the command and control of all maritime operations, both on board ship and in shore-based positions. On board ships, Maritime Surface and Subsurface officers are in charge of various systems (e.g., navigation, bridge management, warfare, aircraft control, and information

management). After several years of experience and training, some Maritime Surface and Subsurface officers command their own ship.

Given the risks involved in the job of a Maritime Surface and Subsurface officer, training these officers involves a significant investment of resources for the CF. Figure 1 shows the training sequence for junior Maritime Surface and Subsurface officers. After selection, Maritime Surface and Subsurface officers complete basic officer training and naval environmental training, otherwise known as Phase I and Phase II. Next, they complete the Maritime Surface and Subsurface specific training known as Phase III and Phase IV. Phase III and Phase IV are critical; students learn the relevant theory and put it into practice in simulations and at sea. The failure rate for first attempt at Phase III can be as high as 40%. By the end of Phase IV, Maritime Surface and Subsurface officers have already spent approximately one and a half years in training. To become qualified Maritime Surface and Subsurface officers, they spend roughly another one and a half years on ship, then complete director level courses. In total, it takes approximately four years to produce a fully qualified Maritime Surface and Subsurface officer. In some cases, it takes longer because students might have to repeat training (i.e., a whole phase or part of a phase) up to two times following a course failure. Therefore, it is important to select those candidates that are most likely to succeed in training.

Figure 1

Training Sequence of Junior Maritime Surface and Subsurface Officers



Selection of Maritime Surface and Subsurface Officers in the Canadian Forces

Over the past five years, the Canadian Forces has recruited an average of 100 Maritime Surface and Subsurface officers per year into its Regular Force cadre. Every year, approximately 200 applicants are screened at a Canadian Forces Recruiting Centre for employment into the Maritime Surface and Subsurface officer occupation, under three external entry streams: a) Direct Entry Officer, b) Continuing Education Officer Training Plan, and c) Regular Officer Training Plan. The current study focuses on candidates who applied under the Direct Entry Officer stream and the Continuing Education Officer Training Plan because they were selected using a multistage system (i.e., stage one at the recruiting centre, stage two at the Naval Officer Assessment Board). The Direct Entry Officer stream is for applicants who possess a university degree whereas the Continuing

Education Officer Training Plan applies to applicants who do not possess a university degree—they will complete their degree on a part-time basis after enrolment. Candidates enrolled under the Regular Officer Training Plan undergo assessment solely at the recruiting centre (i.e., they do not undergo a second stage of selection at the Naval Officer Assessment Board); therefore they are excluded from this study.

All Direct Entry Officer and Continuing Education Officer Training Plan candidates must first complete a screening process at a Canadian Forces Recruiting Centre, which includes: a) psychological testing, b) medical screening, and c) a security check. Psychological testing components include a cognitive ability test and a structured selection interview. For officer applicants, the minimum cut-off score on the cognitive ability test is the 25th percentile (in comparison with all officer applicants); norms are established for both Anglophone and Francophone officer candidates. Applicants who possess a university degree are not required to meet the minimum cut-off score on the cognitive ability test; they must nonetheless complete the test. There is currently no minimum cut-off score for the structured interview component.

For Direct Entry Officer and Continuing Education Officer Training Plan candidates who meet the minimum cut-off score on the cognitive ability test, or who possess a university degree, their test and interview scores are combined to form a composite, referred to as their military potential score. More specifically, the military potential score includes the following areas, which are each worth 30 points for a total out of 90: aptitude (cognitive ability up to 15 pts, education up to 15 points), personality (30 points), and person-environment fit (30 points). The interview determines the latter

60 points. All eligible candidates attend the next available Naval Officer Assessment Board; there is no short list due to the small number of applicants nation-wide. Every year, approximately 75 Direct Entry Officer and Continuing Education Entry Plan candidates attend the boards, which are held three to four times per year in Esquimalt, British Columbia. The maximum capacity per board is 60 candidates.

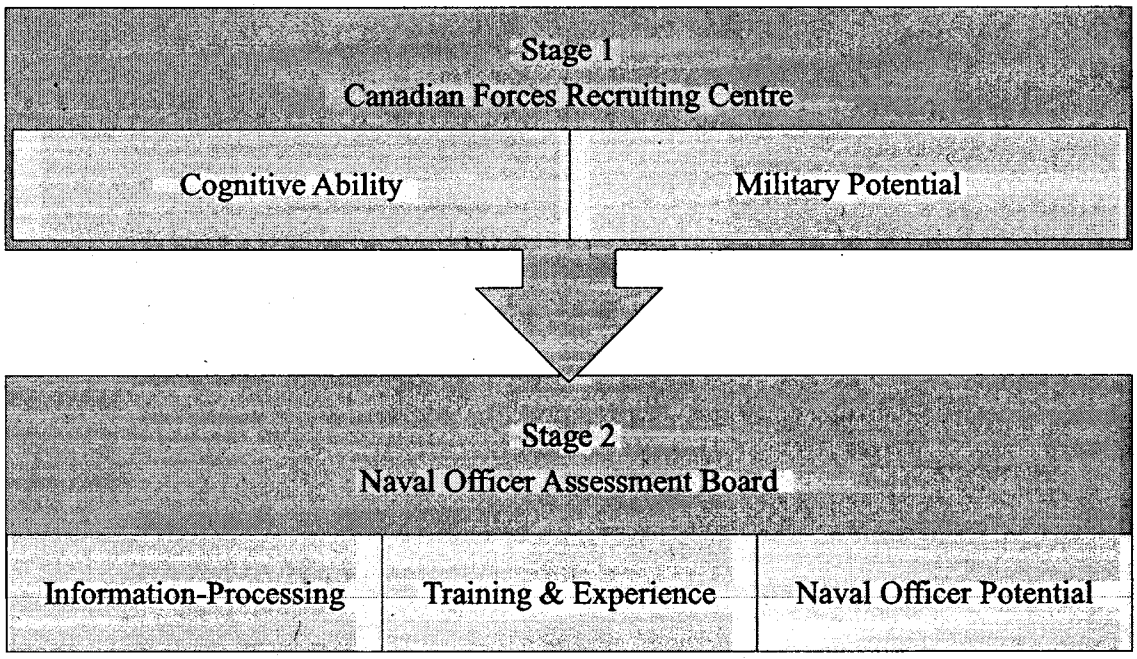
The Naval Officer Assessment Board serves two purposes: a) candidate orientation, and b) candidate assessment. The candidate orientation phase consists of a realistic job preview conducted throughout the five-day Naval Officer Assessment Board. The realistic job preview provides a multi-faceted orientation to the Navy, to the Maritime Surface and Subsurface Occupation, and to the military in general, through the use of audiovisual presentations, a day sail, and tours around the naval training and employment facilities (Boyes, 2008). The realistic job preview helps to manage the expectations of candidates so that they have a realistic view of the organization and the occupation; briefings and presentations also provide information on the positive aspects of the job of a naval officer and the Navy. Results of a meta-analysis by Phillips (1998) indicate that realistic job previews are related to organizational outcomes such as higher performance and lower levels of attrition.

The candidate assessment phase of the Naval Officer Assessment Board requires each candidate to undergo three assessment components: a) a file review, b) an information-processing test, and c) a structured interview (Boyes, 2008). Upon completion of the Naval Officer Assessment Board, candidates are selected for enrolment

in the CF based on their Naval Officer Assessment Board score. The multistage selection process is represented in Figure 2.

Figure 2

Multistage Selection Process for Maritime Surface and Subsurface Officer Candidates[a]



Note. [a]Applies only to the recruitment of external candidates under the Direct Entry Officer stream and the Continuing Education Officer Training Plan.

Despite regular use of Naval Officer Assessment Board scores in selection decisions, a criterion validation study of the Naval Officer Assessment Board process, in its current form, has never been undertaken. Previous criterion validation studies of Maritime Surface and Subsurface selection measures (Bradley, 1990; Hain, 2003; Okros, Johnston, & Rodgers, 1988; Rodgers & Johnston, 1985) are characterized by a confound between predictor constructs (e.g., cognitive ability) and methods (e.g., file review). Using a construct-oriented approach, this study aims to assess the predictive validity of

the Maritime Surface and Subsurface officer selection measures included in both selection stages (i.e. at the recruiting centre and at the Naval Officer Assessment Board).

History of the Naval Officer Assessment Board and Validation Research

In 1976, the Canadian Forces established the Naval Officer Interview Board to assess the suitability of candidates for naval officer occupations and to provide them with exposure to the naval environment prior to enrolment. The Naval Officer Interview Board was created to reduce the attrition rates during Maritime Surface and Subsurface officer occupation training phases, after reports of a loss of 23% of Maritime Surface and Subsurface officers during the Basic Officer Training Phase and another 40-50% loss during Maritime Surface and Subsurface officer occupation training phases (Rodgers, 1984). In 1984, the board's processes were reviewed following reports that the Naval Officer Interview Board was having no impact on Maritime Surface and Subsurface officer training attrition rates (Rodgers, 1984). The board was renamed the Naval Officer Selection Board. The goal of the Naval Officer Selection Board's revised process was to improve performance rates of naval officers on both the basic officer training course and subsequent occupation training, as well as to reduce attrition during naval officer occupation training, through better selection practices and a realistic job preview. It was renamed the Naval Officer Assessment Board in 1989.

Since its inception, the Naval Officer Assessment Board has undergone periodic reviews and modifications. The 1976 version—the Naval Officer Interview Board—included a 45-minute unstructured board interview, file review, tour of the dockyard, films and motivational material. The overall score was based solely on the board

interview results. Attempts to standardize the board interview began in 1982. Although the interview remained unstructured, ratings were based on five specific dimensions: a) academic/technical, b) motivation, c) interpersonal relationships, d) communication, and e) performance before the board. Study results indicated that the average validity coefficient between the Naval Officer Interview Board scores and performance on the leadership practical during the basic officer training course was .15 (Rodgers & Johnston, 1985).

In 1984, the process was re-vamped into an assessment centre style process—renamed the Naval Officer Selection Board—that included seven weighted exercises that formed a composite score: a) a leadership task (15%), b) a conducting officer assessment (15%), c) an interview board (25%), d) a file review (25%), e) an in-basket exercise (10%), and f) two leaderless group discussions (5% each). Statistically significant correlations between individual components of the Naval Officer Selection Board and performance on the leadership practical on the basic officer training course ranged from .16 (for the leadership task) to .34 (for the file review); the total Naval Officer Selection Board score had a .34 correlation with the leadership practical results (Okros et al., 1988). These results represent correlations for both Maritime Surface and Subsurface officers and maritime engineering officers; separate validity coefficients were not provided for each occupation.

Subsequent revisions to the Naval Officer Selection Board led to a reduction in the number of independent measures to five: a) the interview, b) the file review, c) the conducting officer's assessment, d) the leadership stand performance, and e) one

leaderless group discussion. At the same time, the Naval Officer Selection Board was renamed the Naval Officer Assessment Board. Results of validation research on the Naval Officer Assessment Board indicated that both the file review and the leadership stand reached statistically significant corrected correlations with Maritime Surface and Subsurface phase III training results (.23 and .24 respectively); the weighted composite Naval Officer Assessment Board score had a statistically significant corrected correlation of .26 with Maritime Surface and Subsurface phase III training results (Bradley, 1990). Only the file review and the composite achieved statistically significant corrected correlations with Maritime Surface and Subsurface phase IV training results (.32 and .26 respectively; Bradley, 1990). Concurrent validation studies also tested the predictive validity of two experimental measures: the Passage Planning Test (a modification of the US Flight Planning Test designed as a complex cognitive-perceptual test to assess pilot candidates; Okros, 1988a) and the Problem Sensitivity Test (an adventure game concept developed specifically for the CF; Okros & Lynn, 1989). The Passage Planning Test achieved statistically significant correlations with Maritime Surface and Subsurface Phase III (.21) and Maritime Surface and Subsurface Phase IV (.30) training results (Bradley, 1990). In contrast, the Problem Sensitivity Test did not show any significant correlations with either training phases. As a result, the Passage Planning Test was retained as an experimental measure, and then later incorporated as a permanent measure, known as the Maritime Officer Selection Test, within the Naval Officer Assessment Board process.

Further revisions to the Naval Officer Assessment Board resulted in a reduction in the number of measures to four: a) a file review, b) the Maritime Officer Selection Test, c) a structured interview, and d) an essay (Scholtz, 2003). The essay was later removed after it failed to show any significant predictive validity (Boyes, 2008). An unpublished draft validation report of the Naval Officer Assessment Board, which included the three remaining measures, indicated that the Maritime Officer Selection Test was the best predictor of academic performance on both Maritime Surface and Subsurface Phase III and Phase IV training (Hain, 2003). Results by Hain (2003) also showed that the file review and the overall Naval Officer Assessment Board score were good predictors of performance on Maritime Surface and Subsurface Phase III training, but not Maritime Surface and Subsurface Phase IV training. The structured interview failed to predict performance in either phase of training (Hain, 2003). Results of Hain's study must be treated with caution given the small sample size ($n = 24$).

In light of questions regarding the overall usefulness of the Naval Officer Assessment Board, a recent call has been made to re-evaluate the predictive validity of both the Canadian Forces recruiting assessment and the Naval Officer Assessment Board measures against Maritime Surface and Subsurface officer training outcomes (Boyes, 2008). One issue of concern is the final pass/fail training results and first attempt pass/fail training results for both Phase III and Phase IV training. If students fail on their first attempt, they are often afforded the opportunity to retake the course. With a reported 40% failure rate on first attempts (Boyes, 2008), subsequent training attempts represent a significant investment by the Navy. The current study will examine the predictive

validity of Maritime Surface and Subsurface selection measures against both first and final pass/fail results on Phase III and Phase IV training.

To ensure a construct approach in the current study, each part of the multistage selection process is broken down into its construct elements. The entire process includes five predictor constructs: a) cognitive ability; b) military potential; c) information-processing; d) training and experience; and e) naval officer potential. The constructs discussed below are presented within their respective selection components (i.e., recruiting centre and Naval Officer Assessment Board) for ease of understanding. Although the focus of this study is on constructs, references to methods are included to provide context and construct-related information where relevant.

Recruiting Centre Component – Cognitive Ability and Military Potential

Cognitive ability. Research indicates that cognitive ability—also known as general mental ability, intelligence or simply *g*—remains one of the best predictors of job and training performance (Schmidt & Hunter, 1998). Meta-analytic results estimate that general mental ability has a mean corrected validity of between .51 and .55 with overall job performance (Schmidt & Hunter, 1998; Schmidt, Shaffer, & Oh, 2008). General mental ability is also a significant predictor of training performance across numerous jobs and occupations. Results from meta-analyses indicate that the mean corrected correlation between general mental ability and training performance ranges from .54 to .67 (Hunter & Hunter, 1984; Levine, Spector, Menon, Narayanan, & Cannon-Bowers, 1996; Schmidt & Hunter, 1998). In a military context, Salgado (1995) found an uncorrected correlation of .38 between cognitive ability and training performance in pilot trainees. Similarly,

Olea and Ree (1994) found uncorrected correlations of .18 and ..31 between cognitive ability and training performance in pilot and air navigator trainees respectively (.31 and .46 corrected).

Cognitive ability and the Canadian Forces Aptitude Test. At the Canadian Forces Recruiting Centre, candidates complete the Canadian Forces Aptitude Test—a cognitive ability test. The Canadian Forces Aptitude Test is a 60-item speeded cognitive ability test composed of three subscales: a) verbal skills, b) spatial ability, and c) problem solving. When comparing the Canadian Forces Aptitude Test to other measures of general cognitive ability, researchers found evidence of convergent validity with the Wonderlic Personnel Test, the Raven's Standard Progressive Matrices, and the Kaufmann Broad Intelligence Test – Version 2 (Albert, 1998; Leahy, 2008; Vanderpool, 2003a).

Correlations between the total score on the Canadian Forces Aptitude Test and training performance ranged from .20 to .45 across various non-commissioned member occupations (Boswell & Kuschnereit, 2009). For Maritime Surface and Subsurface officers, a significant correlation of .56 was found between the Canadian Forces Aptitude Test and academic scores on Phase III training; the Canadian Forces Aptitude Test did not predict overall pass/fail results on Phase III or Phase IV training (Hain, 2003). Using a sample size larger than the one used by Hain (2003) should lead to the discovery of a significant correlation between the cognitive ability as measured by the Canadian Forces Aptitude Test and pass/fail results in Maritime Surface and Subsurface training phases.

Hypothesis 1a. Cognitive ability as measured by the Canadian Forces Aptitude Test will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 1b. Cognitive ability as measured by the Canadian Forces Aptitude Test will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Specific abilities. While tests of general mental ability predict job performance, a number of organizations use subscales within cognitive ability tests to tailor hiring decisions in environments where specific abilities are deemed predictive of training or job performance (Grubb, Whetzel, & McDaniel, 2004). There are studies on measures of verbal, spatial, and problem-solving abilities in both civilian and military contexts. Kanfer, Wolf, Kantrowitz, and Ackerman (2010) conducted a study of predictors of academic and job performance using a group of university students enrolled in a science and engineering cooperative school-work program. The study included several measures of verbal ability, numerical ability, and spatial ability. Researchers found significant correlations between academic performance—as measured by GPA—and verbal ability ($r = .25, p < .05$), as well as numerical ability ($r = .21, p < .05$); there was no significant correlation between academic performance and spatial ability. None of the specific abilities was related to procedural performance, which was defined as the practical application of mathematical, science, and engineering knowledge to job tasks (Kanfer et al, 2010). In a study of the Air Force Officer Qualification Test, Carretta (2009) found uncorrected correlations between training performance and verbal skills ($r = ..29, p < .01$;

.38 corrected), quantitative skills ($r = ..27, p < .01$; .37 corrected), and spatial ability ($r = ..14, p < .01$; .23 corrected). A study of the US Armed Services Vocational Aptitude Battery subtests showed mean validities of .40 for the word knowledge subtest, .49 for arithmetic reasoning, and .47 for mathematics knowledge against measures of job performance across various occupations; the test did not assess spatial ability (Brown, Le, & Schmidt, 2006).

Specific abilities and the Canadian Forces Aptitude Test. Specific abilities measured by the Canadian Forces Aptitude Test include verbal skills, spatial ability, and problem-solving (i.e., a mathematical component). For some occupations, applicants must meet a cut-off score in one or two subtests because specific subtest scores have been linked to performance in training. Researchers found the following correlations between the Canadian Forces Aptitude Test subtests and training performance: .20 to .39 for verbal skills, .22 for spatial ability, and .22 to .25 for problem-solving (Girard, 2004; Hodgson, 2005; Scholtz, 2004). Based on the previous research findings, it is expected that specific abilities will predict performance in training.

Hypothesis 2a. The verbal skills, spatial ability, and problem-solving subscales of the Canadian Forces Aptitude Test will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 2b. The verbal skills, spatial ability, and problem-solving subscales of the Canadian Forces Aptitude Test will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Military potential. As indicated previously, the cognitive ability score is combined with measures of personality, person-environment fit, and education to form the military potential score. For the purposes of this study, military potential refers only to the personality, person-environment fit, and education components assessed at the recruiting centre (i.e., cognitive ability is excluded and analyzed separately). Ideally, these three constructs would be examined separately. However, scores by constructs were not provided for the current study. The three constructs are discussed separately below, and then combined to formulate hypotheses.

Personality. The study of personality examines several traits to explain individual differences in behaviour. There is consensus among researchers that personality traits can be classified into five broad dimensions (Digman & Inouye, 1986). The factors in the five-factor model of personality—also known as the Big Five—are commonly labelled openness to experience, conscientiousness, extraversion, agreeableness, and emotional stability (Barrick & Mount, 1991).

Several studies show that personality traits based on the five-factor model are related to job performance, but overall correlations remain low because the facets within the five factors vary in their criterion-related validities (Hough & Oswald, 2000). For example, meta-analytic results showed the following mean uncorrected predictive validities between personality factors and job performance: .06 for emotional stability, .06 for extraversion, .06 for agreeableness, .03 for openness to experience, and .12 for conscientiousness (Barrick, Mount & Judge, 2001). Validities were slightly higher for training performance and extraversion (.13), agreeableness (.07), openness to experience

(.14) and conscientiousness (.13); validities for emotional stability were slightly lower (.05). Validity coefficients were higher in studies that included a job analysis to select personality measures (e.g., Tett, Jackson, & Rothstein, 1991). Barrick, Mount, and Judge (2001) found that two factors—conscientiousness and emotional stability—predicted across all jobs and performance criteria.

In relation to training performance, Marcus, Goffin, Johnston, and Rothstein (2007) found that personality was more strongly correlated with typical performance whereas cognitive ability was more strongly correlated with maximum performance on the job, suggesting that cognitive ability would be a better predictor of training outcomes than personality. Research studies among US military members revealed that personality traits were correlated with training outcomes, but were not significant predictors after controlling for cognitive ability (Dean, Conte, & Blankenhorn, 2006; Driskell, Hogan, Salas, & Hoskin, 1994). Nonetheless, results suggest that personality may play a role in predicting attitudinal and motivational factors that, in turn, predict training outcomes (Driskell et al., 1994). Therefore, although research indicates that personality measures do not predict performance better than cognitive ability, the use of personality in conjunction with cognitive ability may improve selection outcomes (Catano, Wiesner, Hackett, & Methot, 2005). For instance, a study of air traffic controller trainees revealed that personality and cognitive ability jointly predicted skill acquisition and job performance (Oakes, Ferris, Martocchio, Buckley, & Broach, 2001).

Methods of assessing personality in selection include personality inventories, interviews, and behavioural observations. In the current study, personality is assessed

using a highly structured interview based on three factors of the five-factor model (i.e., conscientiousness, openness, and emotional stability). Meta-analytic research results indicate that the factors relevant to this study individually predict job performance when used in interviews. Uncorrected correlations between job performance and the factors assessed in highly structured interviews were .20 for conscientiousness,.16 for openness to experience, and .31 for emotional stability (Huffcutt, Conway, Roth, & Stone, 2001).

Person-Environment Fit. Person-environment fit—also labelled person-organization fit—relates to the similarities in values, needs, and characteristics between an individual and an organization. The assessment of person-environment fit in personnel selection is predicated on belief that the better the fit between the individual and the organization, the more likely the individual will perform well on the job (Arthur, Bell, Villado, & Doverspike, 2006). Based on a limited number of research studies, meta-analytic results showed that measures of organization fit in highly structured interviews had a mean uncorrected predictive validity of .32 with job performance (.58 corrected for range restriction; Huffcut et al., 2001). Subsequent meta-analytic findings using a much larger sample of studies found lower mean validities between person-organization fit and various measures of job performance. Arthur, Bell, et al. (2006) reported a corrected mean validity of .12 between person-organization fit and overall job performance.

Assessments of person-environment fit might also include elements of person-job fit, which pertain to a candidate's knowledge of the target occupation and related interests or skills. Meta-analytic research by Huffcutt et al. (2001) indicates that measures of job

knowledge and skills, as assessed in selection interviews, have a mean predictive validity of .23 (.42 corrected for range restriction) with job performance. In highly structured interviews, the mean validity is .18 (.33 corrected for range restriction) between job knowledge and skills, and job performance. Measures of occupation related interests in selection interviews have a mean validity of .13 (.24 corrected for range restriction) with job performance. In highly structured interviews, the mean validity is .14 (.26 corrected for range restriction) between occupational interests and job performance (Huffcutt et al., 2001).

Education. Education—also labelled academic background—refers to the level of formal education of a job applicant (e.g., completed university degree or one year of college), the relevance of an applicant's major field of study (e.g., science or arts), and measures of academic achievement such as grade point average (GPA). In a meta-analysis on the relationship between education and job performance, Ng and Feldman (2009) found that education level was related to objective measures of performance (corrected $r = .24$) and supervisor ratings of job performance (corrected $r = .09$). Education level was negatively related to performance in training programs (corrected $r = -.03$; Ng & Feldman, 2009). Meta-analytic results also suggest that years of education is significantly related to job performance ($r = .10$; Schmidt & Hunter, 1998).

In addition to level and years of education, academic achievement is commonly measured using grade point average (GPA). A meta-analysis by Roth, BeVier, Switzer, and Schippmann (1996) found that undergraduate GPA was a significant predictor of job performance across different types of organizations; the mean validity was .16 (.36

corrected). Validities were higher when the time between graduation and job performance assessments was shorter. For instance, the mean validity between GPA and job performance was .23 (.49 corrected) when the time lapse between graduation and job performance assessment was one year. Mean validities lowered to .15 (.33 corrected) when the time lapse was two to five years; mean validities lowered to .05 (.12 corrected) when the time lapse was six years or more (Roth et al., 1996).

Personality, person-environment fit, and education in the Canadian Forces structured recruitment interview. Within the Canadian Forces recruitment process, measures of personality and person-environment fit are assessed using a structured selection interview, while education is assessed using a standard rating form. For the purposes of this study, scores on the three constructs are combined to form the military potential score. For the personality construct, six interview questions cover the following traits and behaviours: a) work ethic, b) organizational citizenship, c) achievement striving, d) dependability, e) openness to experience, and f) stress management. According to the Personnel Psychology Directive 102 (2006), the first four questions (work ethic, organizational citizenship, achievement striving, and dependability) map onto the conscientiousness factor of the five-factor model. Openness to experience maps onto the openness to experience factor, while stress management maps onto the emotional stability factor of the five-factor model (Personnel Psychology Directive 102, 2006). The structured recruitment interview was not designed to assess agreeableness and extraversion (Skomorovsky, 2009).

Research on the convergent validity of the personality portion of the structured recruitment interview revealed that the interview lacked convergent validity with three well-established measures of personality (i.e., the Trait-Self Descriptive Personality Inventory [TSD-PI], the Neuroticism-Extraversion-Openness to Experience Personality Inventory – Revised [NEO-PI-R], and the Honesty-Emotionality-Extraversion-Agreeableness-Conscientiousness-Openness to Experience Personality Inventory [HEXACO-PI]; Skomorovsky, 2009), suggesting that the personality portion of the structured recruitment interview is not a valid measure of personality because it lacks construct validity. Moreover, the personality portion of the structured recruitment interview failed to predict training performance on basic recruit training (Skomorovsky, 2009) and on occupation training in the logistics family of occupations (Girard, 2009).

The person-environment fit portion of the structured recruitment interview is divided into four areas: a) target occupation related skills, b) knowledge of target occupation, c) congruency between an applicant's interests and the target occupation, and d) realistic expectations. Validation study results on the person-environment fit portion of the structured recruitment interview indicate that the person-environment fit portion failed to predict training performance in the logistics job family (Girard, 2009).

The education score is calculated using information about an applicant's highest completed level of education, additional upgrading through individual courses, and GPA. Sources include high school, college, and/or university transcripts. The Canadian Forces has not conducted predictive validation research on the education component as it is calculated at the recruiting centre. A study by Girard (2004) found that years of

education was a significant predictor of resource management clerk training performance among Anglophone students ($r = .24, p < .01$) and Francophone students ($r = .31, p < .01$).

Although published research suggests that personality and person-environment fit are predictors of performance in various organizations (e.g., Huffcutt et al., 2001), studies on the Canadian Forces structured recruitment interview suggest that measures of personality and person-environment fit contained therein are poor predictors of training performance. Moreover, the structured recruitment interview questions were designed to predict general military performance; they were not designed to specifically predict performance on Maritime Surface and Subsurface training phases or subsequent job performance in the occupation. Education was found to be a predictor of training and job performance in organizations (e.g., Ng & Feldman, 2009), in addition to training performance in the Canadian Forces (Girard, 2004). The combined personality, person-environment fit, and education constructs in this study are expected to correlate with Maritime Surface and Subsurface officer training performance.

Hypothesis 3a. Scores of military potential from the Canadian Forces Recruiting Centre will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 3b. Scores of military potential from the Canadian Forces Recruiting Centre will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Naval Officer Assessment Board Component – Information-processing, Training and Experience, and Naval Officer Potential.

Information-processing. Information-processing refers to the human brain's capacity to code, store, and retrieve environmental inputs (Arthur, Doverspike, & Bell, 2004). Ackerman's (1988) work suggests that information-processing involves specific abilities that underlie general cognitive ability (e.g., speed of cognitive processes). Similarly, Erkstrom, French, and Hartman (1979) refer to information-processing as basic, cognitive factors (cited in Arthur, Doverspike, et al., 2004). The main components of the information-processing system include short-term sensory store, perception, decision making (also labelled response selection), short-term memory (also labelled working memory), long-term memory, response execution, and attention (Arthur, Doverspike, et al., 2004). Short-term memory store is a mechanism that accepts external stimuli and retains the information for less than one second. Perception is a mechanism that organizes information and distinguishes important information from irrelevant or distracter stimuli. Decision making is a mechanism through which appropriate decisions are taken based on several thought processes (e.g., memory retrieval and solution comparisons). Short-term memory is a mechanism that stores information for brief periods in order of presentation; due to its limited size and capacity, it is a bottleneck within the information-processing system. In contrast, long-term memory stores information for longer periods and arranges information into various structures or schemas. Response execution is a mechanism that regulates response selection and execution. Lastly, attention is a mechanism that filters information to the conscious level

for subsequent use; attention is also a bottleneck for the information-processing system (Arthur, Doverspike, et al., 2004).

Traditional methods of information-processing testing are done via paper-and-pencil tests. Although information-processing tests have not been used widely in civilian selection settings, they have been used extensively for pilot selection in the US Air Force (see Carretta, 2000). Tests of specific information-processing aptitudes (i.e., tailored to specific jobs or occupations) have shown predictive validities in the .20 to .50 range with task, job, and simulator performance (Arthur, Doverspike, et al., 2004). Validities for information-processing tests are often equal to cognitive ability tests (Arthur, Doverspike, et al., 2004).

Information-processing and the Maritime Officer Selection Test. A job analysis of first tour Maritime Surface and Subsurface officers identified several information-processing abilities that were critical to successful performance as a Maritime Surface and Subsurface officer: a) perception (labelled as flexibility of closure and perceptual speed), b) memory, c) decision making (labelled as deductive reasoning), and attention (labelled as selective attention; Rodgers, 1986; Rodgers & Zuliani, 1985). To assess naval candidates for information-processing abilities identified in the job analysis, researchers modified the Flight Planning Test used by the U.S. Army Research Institute in helicopter pilot selection. To adapt the test to a naval environment, nautical terminology replaced air terminology in the test to reflect elements of a nautical passage plan. The test was eventually renamed the Maritime Officer Selection Test and incorporated into the Naval Officer Assessment Board process.

Commonly labelled a "complex cognitive perceptual test" (Blanc, 2003; Okros, 1988a), the Maritime Officer Selection Test is designed to measure naval relevant information-processing abilities such as working memory, selective attention, and deductive reasoning (Okros, 1988a; Scholtz, 2003). A study by Blanc (2003) argues that the Maritime Officer Selection Test measures three different factors: a) speediness, b) spatial scanning, and c) general reasoning. Nevertheless, the factors identified by Blanc (2003) constitute underlying elements of information-processing tests (Arthur, Doverspike, et al., 2004). Despite the debate over the factor structure of the test, it has not been altered from its original form.

Previous research studies revealed that the Maritime Officer Selection Test was a significant predictor of various performance aspects of Maritime Surface and Subsurface officer training phases. Bradley (1990) found that the Maritime Officer Selection Test was correlated with final grades on Phase III training ($r = .21$) and with final grades on Phase IV training ($r = .30$). Hain (2003) found that the Maritime Officer Selection Test was correlated with academic grades on Phase III training ($r = .48$) and with academic grades on Phase IV training ($r = .51$). A study by Okros (1988) showed that the Maritime Officer Selection Test correlated .33 with scores on the practical portion of Phase III training. Based on the abovementioned research findings, it is expected that information-processing abilities—as measured by the Maritime Officer Selection Test—will predict training performance in the current study.

Hypothesis 4a. Information-processing, as measured by the Maritime Officer Selection Test, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 4b. Information-processing, as measured by the Maritime Officer Selection Test, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Training and Experience. The purpose of training and experience evaluations is to assess an applicant's relevant training, work experience, and educational achievement in relation to the target job. The information is usually provided by the applicant using a standardized form and/or a résumé. Results from a meta-analysis on training and experience evaluations, and job performance show that training and experience evaluations have an overall mean validity of .12 (.17 corrected for range restriction; McDaniel, Schmidt, & Hunter, 1988a).

Other studies examined work experience and educational achievement separately in the context of training and experience evaluations (education was covered in a previous section and will not be repeated here). A meta-analysis by Quiñones, Ford, and Teachout (1995) found a mean validity of .22 (.27 corrected for range restriction) between work experience and job performance. Another study found a mean validity of .20 (.28 corrected for range restriction) between work experience and high complexity jobs (McDaniel, Schmidt, & Hunter, 1988b).

A common method for assessing training and experience is the point method (Porter, Levine, & Flory, 1976). Using a pre-established rating scheme, raters allocate points based on the amount and recency of the applicant's job relevant training, education, and experience. Ideally, a job analysis determines the proportion of points dedicated to areas of training, education or experience (Gatewood, Field, & Barrick, 2008). McDaniel et al. (1988a) found that the point method had a low mean validity (.07 uncorrected and .11 corrected for range restriction). Yet, mean validities were higher when applicants had low mean levels of job experience (.16 uncorrected and .29 corrected for range restriction; McDaniel et al., 1988a).

Training and experience, and the file review for naval officer candidates. Raters on the Naval Officer Assessment Board use a point method to rate individual candidates based on a review of information contained in applicant files. Points are allocated for educational achievement, employment experience, military or paramilitary experience, and miscellaneous activities, interests or achievements. The rating scale does not include descriptive anchors.

Hain's (2003) study found that training and experience—as measured by the file review—was correlated .48 (uncorrected) with grades on Maritime Surface and Subsurface Phase III training. Another study by Bradley (1990) revealed that the file review was correlated with grades on Phase III training (.20; .23 corrected for range restriction) and grades on Phase IV training (.28; .32 corrected for range restriction). Based on the abovementioned research findings, training and experience is expected to predict training performance in the current study.

Hypothesis 5a. Training and experience, as measured by scores on the file review, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 5b. Training and experience, as measured by scores on the file review, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Naval officer potential. At the Naval Officer Assessment Board, three constructs assess naval officer potential: a) leadership, b) self confidence, and c) oral communication.

Leadership. The leadership literature contains over 221 definitions of leadership (Rost, 1993). Leadership is commonly defined by the particular style (e.g., transformational, transactional) or broader leadership category (e.g., trait, behavioural, situational) under study. Leadership theories commonly attempt to explain what makes leaders effective or ineffective. In the Canadian Forces, effective leadership is defined as "directing, motivating, and enabling others to accomplish the mission professionally and ethically, while developing or improving capabilities that contribute to mission success" (Canadian Forces Leadership Institute, 2005, p.30).

Kickul and Neuman (2000) found that extroversion, openness to experience, and cognitive ability predicted emergent leadership behaviours among undergraduate students. Meta-analytic results show a mean correlation of .21 (.27 corrected for range restriction) between intelligence and leadership (Judge, Colbert, & Ilies, 2004). In a meta-

analysis of personality and leadership, results indicate that the five-factor model of personality had a multiple correlation of .48 with leadership. Correlations between leadership and the five dimensions were Neuroticism = -.24, extraversion = .31, openness to experience = .24, agreeableness = .08, and conscientiousness = .28 (Judge, Bono, Ilies, & Gerhardt, 2002).

In a longitudinal study of United States Military Academy students, researchers found correlations between results in a leadership class and tests of comprehension (.14), logic (.10), ideational fluency (.15), social judgment in organizational scenarios (.10), and work orientation (.08; Milan, Bourne, Zazanis, & Bartone, 2002). Scores on overall leadership development were correlated with measures of complex problem-solving skills (i.e. problem construction in military scenarios, .14 and social judgment in organizational scenarios, .10), tacit knowledge for military leaders (.12), background and life experiences (.11), dominance (.09), energy (.10), traditional values (.09), and work orientation (.10; Milan et al., 2002). A separate study of military officer cadets revealed that the best predictors of leader emergence were physical fitness (.20), prior influence experience (.18), self-esteem (.14), and the leader potential index (.15); leader effectiveness was predicted by physical fitness (.22) and prior influence experience (.24; Atwater, Dionne, Avolio, Camobreco, & Lau, 1999). Cognitive ability was not a significant predictor of either leader emergence or leader effectiveness (Atwater et al., 1999).

Common approaches to assessing leadership in selection include leadership questionnaires, interviews, and group scenario observations. The Naval Officer

Assessment Board uses a highly structured interview format composed of situational judgment questions. Recent meta-analytic results indicate that situational judgment tests have a mean predictive validity of .21 (.28 corrected for range restriction) with leadership (Christian, Edwards, & Bradley, 2010). In a similar vein, Huffcutt et al. (2001) found a mean overall validity of .26 (.47 corrected for range restriction) between interview ratings of leadership and job performance. In highly structured interviews, the mean validity was .22 (.40 corrected for range restriction) between leadership ratings and job performance (Huffcutt et al., 2001).

Self Confidence. When assessing candidates for leadership or management positions, measures of self confidence may be included in an employment interview. Ratings of self confidence are frequently based on general impressions of the candidate's nonverbal cues during a selection interview. Certain nonverbal cues such as eye contact, smiling, hand gestures, and body orientation may be indicators of a candidate's self confidence. Research indicates that visually based ratings of job applicants during an interview were correlated .32 with ratings of job performance (Motowidlo & Burnett, 1995). Another study by Burnett and Motowidlo (1998) showed that, taken together, visually based interview ratings (i.e. ratings based on nonverbal cues) and response content ratings correlated .35 with job performance. In a study of air traffic controller trainees, ratings of self-efficacy by candidates were correlated .42 with simulator performance (Ackerman & Kanfer, 1993). In a study of military pilot candidates, ratings of self confidence using a structured interview with behaviourally anchored rating scales correlated minimally (.04) with job performance (Walters, Miller, & Ree, 1993).

Communication Skills. Oral communication skills of job candidates often play an important role in selection interviews. Research shows that raters are influenced by the paralinguistic cues of applicants; paralinguistic cues include intonation, speech fluency, and volume (Burnett & Motowidlo, 1998). One study found that aurally based interview ratings correlated .33 with ratings of job performance (Motowidlo & Burnett, 1995). Meta-analytic results indicate a mean validity of .14 (.26 corrected for range restriction) between communication skills assessed during structured interviews and job performance; the mean validity rose to .17 (.31 corrected for range restriction) in highly structured interviews (Huffcutt et al., 2001). Despite their importance in managerial performance, communication skills are most frequently assessed in an informal fashion in employment interviews, thereby reducing the reliability of the ratings (Bambacas & Patrickson, 2009). In low structure interviews, the mean validity was only .05 between ratings of communication skills and job performance (Huffcutt et al., 2001).

Naval officer potential and the Naval Officer Assessment Board. During the Naval Officer Assessment Board, raters assess the naval officer potential of candidates using a structured selection interview. The structured Naval Officer Assessment Board interview is designed to measure three constructs within naval officer potential: a) leadership, b) self confidence, and c) oral communication. Leadership is assessed using situational interview questions that tap into elements of decisiveness, achievement, and motivation. Raters score each answer using behaviourally anchored rating scales. During the interview, the measurement of self confidence and communication skills has

little structure; scores are based on individual rater impressions and general guidelines, but no behaviourally anchored rating scales.

Previous validation studies found no significant correlation between scores on the structured Naval Officer Assessment Board interview and performance on Maritime Surface and Subsurface training phases (Bradley, 1990; Hain, 2003). However, issues of small sample size and question problems may explain the lack of predictive validity. Based on the aforementioned body of research, naval officer potential is expected to predict training performance in the current study.

Hypothesis 6a. Naval officer potential, as measured by combined scores on the constructs of leadership, self-confidence, and oral communication of the Naval Officer Assessment Board interview, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 6b. Naval officer potential, as measured by combined scores on the constructs of leadership, self-confidence, and oral communication of the Naval Officer Assessment Board interview, will be significantly correlated with first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Multistage Selection

A multistage selection system—also termed a multiple hurdle model—is designed much like a multilevel video game; job applicants must successfully complete each stage in sequential order before moving on to the next stage. To pass a stage, the applicant must meet the minimum cut-off for the predictor or set of predictors (i.e., assessment

measures) at a given stage. Failure to pass a stage results in the removal of the applicant from the selection process, thereby narrowing the applicant pool after the completion of each stage.

The order of stages is normally determined by ability requirements and logistic considerations, whereby the early stages in the selection process include predictors that are less costly and easier to administer (e.g., paper-and-pencil tests), and the assessment of applicants is based on minimum job requirements (Finch, Edwards, & Wallace, 2009). Multistage selection systems are especially useful as a cost-saving measure when managing a large applicant pool; they are also appropriate when testing for specific job required attributes when high scores on one predictor cannot compensate for low scores on another predictor (Catano et al., 2005). For organizations like the military, where entry-level jobs such as Maritime Surface and Subsurface officer require long and expensive training programs, a multistage selection system is particularly relevant (Das, 2007). In terms of disadvantages, multistage selection systems are generally time-consuming and difficult to validate because of the restriction of range of applicant scores in later stages (Catano et al., 2005). In addition, the length of time involved in the sequential administration of each selection stage may result in the loss of good applicants, especially where the demand for qualified applicants is high (Das, 2007).

Multistage selection and incremental validity. An assessment measure is said to have incremental predictive validity when results indicate that it adds to the prediction of the criterion (e.g., training performance) beyond that of another predictor or set of predictors. More specifically, it is defined as "the extent to which additional predictors

enhance the proportion of overlapping variance with the criterion" (McGrath, 2008, p. 195). Within multistage selection systems, each additional stage in the selection process should demonstrate significant incremental predictive validity over the previous stage(s); otherwise, there is little reason to include additional stages. An analysis of incremental validity begins with the examination of the correlations between the predictors in all stages of selection. Next, analyses can determine whether each stage offers incremental validity over the previous stage.

Multistage selection for Maritime Surface and Subsurface officers. Given the significant expenditure of resources involved in training naval officers, several countries have adopted a multistage selection system to assess potential naval officers beyond the initial recruiting centre screening process (Boswell, 1993). In the Canadian Forces, a dual hurdle selection system is used to select Maritime Surface and Subsurface officers. The first hurdle occurs at the various Canadian Forces Recruiting Centres across Canada, while the second hurdle takes place at the Naval Officer Assessment Board in Esquimalt.

One of the specific purposes of the Naval Officer Assessment Board is to select Maritime Surface and Subsurface officers, while the primary purpose of Recruiting Centres is to select individuals for general military service. As discussed above, different constructs are measured at each stage of the selection process. Cognitive ability and military potential are measured at the Canadian Forces Recruiting Centre; the Naval Officer Assessment Board measures information-processing, training and experience, and naval officer potential. If Naval Officer Assessment Board predictors are to provide incremental validity above recruiting centre predictors, they should not be too highly

correlated. The research literature suggests that relationships exist between the constructs in this study, as discussed below.

Cognitive ability and information-processing. Research studies incorporating measures of cognitive ability and various elements of information-processing reveal a moderate relationship between cognitive ability and information-processing. For instance, Vernon and Jensen (1984) found a multiple correlation of .47 between cognitive ability variables and information-processing variables. Another study by Ackerman and Kanfer (1993) showed correlations between measures of information-processing and verbal skills (.40 to .44), spatial ability (.53 to .55), and problem-solving (.53). Lastly, Okros (1988a) found a correlation of .44 between the Maritime Officer Selection Test and the General Classification test—a measure of cognitive ability upon which the Canadian Forces Aptitude Test was built. A subsequent study by Bradley (1990) found no significant correlation between the Maritime Officer Selection Test and the General Classification test. Based on the aforementioned research, cognitive ability is expected to be correlated to information-processing in this study.

Hypothesis 7. Cognitive ability, as measured by the Canadian Forces Aptitude Test, will be significantly correlated with information-processing, as measured by the Maritime Officer Selection Test.

Training and experience, and cognitive ability. The majority of Maritime Surface and Subsurface officer candidates have no training or experience in the naval environment. In addition, a large proportion of applicants are recent high school, college,

or university graduates who have very little work experience. The Naval Officer Assessment Board file review gives equal weight to: a) academic achievement, b) employment history, c) military and para-military experience, and d) related activities and interest. However, candidates are likely to score lower on the latter three areas than on academic achievement. As a result, academic achievement (as measured by GPA) may constitute the largest proportion of the overall training and experience scores on the Naval Officer Assessment Board file review. Studies indicate that high school grades correlate between .50 and .70 with cognitive ability (Jensen, 1998). Other studies found correlations between cognitive ability and university GPA between .16 (Lievens, Buyse, & Sackett, 2005) and .38 (Rohde & Thompson, 2007). Research by Bradley (1990) revealed that training and experience measured using a file review during the Naval Officer Assessment Board was correlated .36 with cognitive ability. Based on the abovementioned research findings, training and experience is expected to be correlated with cognitive ability.

Hypothesis 8. Training and experience, as measured by the Naval Officer Assessment Board file review, will be significantly correlated with cognitive ability, as measured by the Canadian Forces Aptitude Test.

Training and experience, and military potential. A portion of the military potential interview conducted at the recruiting centre (i.e. at stage one of selection) assesses person-environment fit. The person-environment fit assessment considers, in part, the relevant training and experience of candidates in relation to the target occupation. Hence, there appears to be some overlap in constructs measured during the

military potential interview at the recruitment centre and the file review at the Naval Officer Assessment Board. It is expected that this overlap will be reflected in the correlation between training and experience scores, and military potential scores.

Hypothesis 9. Training and experience, as measured by the Naval Officer Assessment Board file review, will be correlated with military potential, as measured during the recruiting centre structured interview.

Incremental validity and the Naval Officer Assessment Board. Given the specific focus of the Naval Officer Assessment Board on selecting candidates for the navy environment and for the Maritime Surface and Subsurface occupation in particular, Naval Officer Assessment Board measures should provide incremental validity above recruiting centre predictors. The recruiting centre measures cognitive ability and military potential; the Naval Officer Assessment Board measures information-processing, training and experience, and naval officer potential. Research suggests that cognitive ability is one of the best predictors of training and job performance (Schmidt & Hunter, 1998), whereas the constructs of personality and person-environment-fit are weaker predictors of training and job performance. Therefore, incremental validity above cognitive ability may be more difficult to achieve.

Although validities for information-processing tests are often equal to cognitive ability tests, they appear to add very little incremental validity above cognitive ability (Arthur, Doverspike, et al., 2004). Bradley (1990) found that only the file review—a measure of training and experience—provided incremental predictive validity above

stage one measures (i.e., cognitive ability and military potential assessed at the recruiting centre) for Phase III training. However, the study by Bradley (1990) did not include the Maritime Officer Selection Test or the revised structured interview within the Naval Officer Assessment Board. Based on the aforementioned research findings, it is expected that the Naval Officer Assessment Board constructs (i.e., information-processing, training and experience, and naval officer potential) will show incremental predictive validity above the Recruiting Centre constructs (i.e., cognitive ability and military potential).

Hypothesis 10a. The Naval Officer Assessment Board measures at stage two of the selection process will show incremental predictive validity above the Canadian Forces Recruiting Centre measures at stage one of the Maritime Surface and Subsurface selection process, against first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

Hypothesis 10b. The Naval Officer Assessment Board measures at stage two of the selection process will show incremental predictive validity above the Canadian Forces Recruiting Centre measures at stage one of the Maritime Surface and Subsurface selection process, against first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

Method

Data Set

Analyses for the current study were based on archival data obtained from the Directorate of Military Personnel Operational Research and Analysis, the Naval Officer

Training Center, and the Director of Maritime Training and Education. In total, training performance was obtained for 142 officers who attempted the Maritime Surface and Subsurface training. Of these, 33 cases were excluded because their Military Potential score was either missing or based on a previous scoring format that is not comparable with the current scheme. Of the remaining 109 cases, five cases were excluded for the following reasons: one member was removed from training for medical reasons, one member released voluntarily from the military for personal reasons, and three members are still undergoing training. The remaining 104 officers attended the Maritime Surface and Subsurface Phase III training program between 2007 and 2010. Of these 104 officers, 74 went on to attend the Maritime Surface and Subsurface Phase IV training program between 2007 and 2010.

Among the 104 officers, 82.7% were males and 94.2% were Anglophone. Divided by entry plan, 69.2% were Direct Entry Officers (two of which were component transfers from the Reserves), 29.8% enrolled under the Continuing Education Officer Training Plan, and 1% was not specified.

The officers represent candidates chosen from Naval Officer Assessment Boards conducted between 2005 (serial 0503) and 2008 (serial 0804). Following an assessment at a Canadian Forces Recruiting Centre, the candidates were sent to the next available Naval Officer Assessment Board serial at the Naval Officer Training Centre in Esquimalt, British Columbia. The Naval Officer Assessment Board was conducted over five days. During the first day, candidates received briefings on the Navy, recruit training, the selection process, benefits, career progression, and the Military Family

Resource Centre. Concurrently, assessment board members completed a review of each candidate's application file. On the second day, candidates were taken on a full-day sail in the Juan de Fuca Strait and attended a briefing on submarines.

Over the next two days, the candidates were divided into groups to facilitate the rotation of activities. While one group wrote the Maritime Officer Selection Test, another group participated in various tours of the base, and a third group underwent the panel interview. By the end of the fourth day, all candidates were assessed and had attended all the orientation briefings. On the fifth day, candidates completed a feedback survey on the Naval Officer Assessment Board process, they received a copy of all the week's briefings, and they were individually briefed on the board's findings regarding their application. If a candidate was not found suitable by the board, he or she had immediate access to a military career counsellor to discuss other career options.

Predictors

Canadian Forces Aptitude Test. Cognitive ability is measured at the first stage of Maritime Surface and Subsurface selection using the Canadian Forces Aptitude Test. The Canadian Forces Aptitude Test is a 60-item speeded cognitive ability test composed of three subscales: verbal skills, spatial ability, and problem solving. Scores on each subscale are summed to form a total cognitive ability score. The raw total score obtained on the Canadian Forces Aptitude Test is converted to a percentile. The current study used percentile scores for the total Canadian Forces Aptitude Test score, as well as for the specific abilities subscales. Candidates can choose to write the Canadian Forces Aptitude Test in English or French; percentile norms are applied according to the test language

chosen by the candidates. Norms are also broken down by rank category (i.e., non-commissioned member and officer); the officer norms are applied to Maritime Surface and Subsurface officer candidates. Both paper-and-pencil and electronic (non-adaptive computer-based) versions of the test are available at Canadian Forces Recruiting Centres across Canada.

Psychometric analyses conducted on the Canadian Forces Aptitude Test indicate internal consistency reliabilities ranging from .78 to .80 for the verbal skills scale, .64 to .70 for the spatial ability scale, and .88 to .90 for the problem solving scale for both Anglophone and Francophone populations (Donohue, 2005; Vanderpool, 2003b). Cronbach's alpha for the Canadian Forces Aptitude Test as a whole is estimated at .90 for both language groups (Donohue, 2005; Vanderpool, 2003b). Test-retest reliabilities for the Anglophone population on the verbal skills scale, spatial ability scale, problem-solving scale, and overall were .74, .66, .86, and .88 respectively. For the Francophone population, test-retest reliabilities on the verbal skills scale, spatial ability scale, problem-solving scale, and overall were .79, .66, .86, and .88 respectively.

Validation studies of the Canadian Forces Aptitude Test against military occupation training outcomes provide mixed results. Validity estimates range from .16 to .44 (uncorrected for range restriction) for a number of Non-Commissioned Member occupations (Boswell & Kuschnereit, 2009). To date, the Canadian Forces Aptitude Test has not been validated against training outcomes for officer occupations.

Recruiting interview and education rating. Military potential is assessed during the first stage of Maritime Surface and Subsurface officer selection using a structured recruitment interview and an education rating form. The structured interview is conducted at Canadian Forces Recruiting Centres by Military Career Counsellors. It should be noted that the structured interview questions were changed in 2009; all data from the current study include officers who were assessed using the pre-2009 format. The structured interview questions used in this study were designed to assess personality traits and person-environment fit. Applicant responses to 10 questions were rated using 5-point behaviourally anchored rating scales. A total of 30 points was allocated to the personality portion and 30 points were allocated to the person-job fit portion. The education score is based on 15 points. Scores on all three constructs were summed to provide a total military potential score. Separate construct scores were not available for the current study. There have been no studies measuring the reliability of the military potential structured interview.

Maritime Officer Selection Test. Information-processing is measured at the second stage of Maritime Surface and Subsurface officer selection (i.e., during the Naval Officer Assessment Board) using the Maritime Officer Selection Test. The Maritime Officer Selection Test is a paper-and-pencil test that contains 60 multiple-choice questions divided into five timed sections and three progressive levels of difficulty. Candidates must answer a minimum of 24 questions correctly in order to pass the test; the cut-off score was established using a modified version of the Angoff (1971) method for setting testing standards (Scholtz, 2003). Candidates who do not achieve the minimum

cut-off score continue the Naval Officer Assessment Board process; while they are unlikely to be found suitable, their files are red flagged for further discussion by the board members after all assessment scores have been collected.

The Maritime Officer Selection Test makes up 35% of the total Naval Officer Assessment Board score, which is scored out of 100 points. The total raw score for the Maritime Officer Selection Test is converted into a weighted score by multiplying the raw score by 0.58. The weight of 35% was established using both statistical and judgemental methods (Okros et al., 1988; F.A.J. Boyes, personal communication, February 26, 2010). Previous research suggests that the test showed sufficient internal consistency (.85; Stouffer, 1996). Scholtz (2003) reports a mean raw score of 31.77 and a standard deviation of 7.63, while a separate report by Blanc (2003) found a mean raw score of 29.59 (SD = 7.98) for women and 27.30 (SD = 6.92) for men.

File review. Each Maritime Surface and Subsurface officer candidate provides training and experience information during their recruiting application process. Applicant files are sent to the Naval Officer Assessment Board where raters conduct a file review to assess the training and experience of candidates. Using a nine-point anchored rating scale ranging from 1 to 5 (half marks included), four board members rate each candidate on the following dimensions: a) educational achievement; b) employment history; c) other activities, interests or achievements; and d) military or paramilitary experience. Scores are summed, and then divided by the number of assessors to provide a raw average file review score for each candidate; the maximum raw file review score is 20. A total file review score is computed by adding the military potential score and the

cognitive ability score—from the recruiting centre measures at stage one—to the initial file review score.

The file review score makes up 35% of the total Naval Officer Assessment Board score. The weighted score is calculated by multiplying the total file review score by 1.17. The weight was established using both statistical and judgemental methods (Okros et al., 1988; F.A.J. Boyes, personal communication, February 26, 2010). In the current study, the net file review score (i.e., the file review score without the military potential and cognitive ability scores) was used in order to examine only the portion of the file review that assessed training and experience by board members.

During the file review process, there is no requirement for board members to discuss each candidate's file or arrive at a consensus on the final score. However, if the point spread among assessors is greater than one point per dimension, all board members discuss the file to ensure that all pertinent information was included when assigning scores. Enough consensus must be reached to have the final scores fall within an average of one raw score per dimension (within 4 points for the total file review score). Also, if any assessor gives a candidate a score of 1 (poor) on any element of the file rating, the file is red-flagged for later discussion regarding the candidate's suitability (i.e., once the results from the interview and Maritime Officer Selection Test are available).

The reliability of the file review was estimated using only one index of reliability—interrater reliability. The interrater reliability index (i.e., the intraclass correlation coefficient) estimates the extent to which assessors rank order candidates in a

manner that is consistent with other assessors. Another index of reliability, interrater agreement, estimates the degree of consensus across assessors for each candidate's score; high interrater agreement indicates that scores are equivalent or interchangeable in terms of their absolute value (LeBreton & Senter, 2008). Interrater agreement (i.e., using rwg(*j*) calculations) could not be assessed for this study given a lack of data at the item level for each rater.

The interrater reliability estimates for the file review were calculated using data from two separate Naval Officer Selection Boards—serials 0903 and 1001. In both cases, there were four board members providing individual ratings for each of the candidates. The intraclass correlation coefficient for consistency among assessors—interrater reliability estimates—ranged from .83 (serial 1001) to .89 (serial 0903). Reliability estimates between .71 and .90 are considered strong, but estimates above .90 are recommended for important individual decisions such as hiring (LeBreton & Senter, 2008).

Panel interview. During the Naval Officer Assessment Board, each candidate undergoes a short structured interview. The interview is conducted by a panel of four assessors: three senior naval officers and a senior Personnel Selection officer. They are the same four assessors who conduct the file review. The interview is designed to assess naval officer potential based on three constructs: a) leadership, b) self confidence, and c) oral communication. Following an icebreaker question, candidates must respond to a series of structured situational interview questions designed to assess leadership. Responses to each question are scored using a behaviourally anchored rating scale

ranging from 1 (poor) to 5 (excellent). Ratings of self confidence and oral communication are based on the impressions of individual raters who follow minimal rating guidelines; ratings for both constructs range from 1 (poor) to 5 (excellent). Scores from each assessor are summed, and then an average is computed to form a raw score for each candidate; the maximum raw score is 25. It should be noted that prior to autumn 2009, the interview contained five questions; one question was removed in 2009 due to problems associated with the scale anchors. Data for this study includes only scores from the old interview format; criterion data are not yet available for applicants from 2009 and onward. Thus, the maximum raw score for each candidate in this study is 30. The interview makes up 30% of the total Naval Officer Assessment Board score. The weight was established using both statistical and judgemental methods (Okros et al., 1988; F.A.J. Boyes, personal communication, February 26, 2010).

For the panel interview, there is no requirement for board members to discuss each candidate's responses or arrive at a consensus on the final score. However, if the point spread among assessors is greater than one point per dimension, all board members discuss the responses to ensure that all pertinent information was included when assigning scores. Enough consensus must be reached to have the final scores fall within an average of one raw score per dimension (within 5 points for the total interview score). Also, if any assessor gives a candidate a score of 1 (poor) on any element of the interview, the file is red-flagged for later discussion regarding the candidate's suitability.

Data were not available to estimate the reliability of the old interview format. Instead, reliability estimates for the new format were calculated. As with the file review,

the only reliability index calculated for the current study was the interrater reliability (interrater agreement could not be calculated). The intraclass correlation coefficient for consistency among assessors—interrater reliability estimates—ranged from .90 (serial 0903) to .92 (serial 1001). Reliability estimates of this magnitude are considered strong to very strong (LeBreton & Senter, 2008). Given that a poorly performing question was removed from the old format, reliability estimates would likely have been lower for the old format.

Criterion

Criterion Selection. The best criterion measures are "relevant, reliable, and uncontaminated" (SIOP, 2003, p. 14). The relevance of the measure is a primary factor in validation research. In this study, the criterion is the pass/fail result of students attending the Maritime Surface and Subsurface officer training phases III and IV. Although a dichotomous measure of performance is restricted in variance, often attenuating the strength of the predictor-criterion relationship, it nonetheless represents an important outcome in the context of training in the Canadian Forces. Passing Maritime Surface and Subsurface training phases, and an additional 24 months (on average) of on-the-job training after completion of Phase IV, implies that officers have reached the operationally functional point; upon achieving this milestone, a Maritime Surface and Subsurface officer receives his/her first posting to a ship and the CF starts getting a return on its investment in the officer. On the other hand, failure of Maritime Surface and Subsurface training phases represents further training for the member in the Maritime Surface and Subsurface occupation or transfer to another occupation.

The purpose of Maritime Surface and Subsurface Phase III training is to introduce students to basic knowledge and skills of ship handling, navigation, and watch-keeping. Phase III training is divided into three successive phases. First, students undergo a seven-week academic phase to learn the theory of relative velocity, rules of the road, astronomics and tidal theory, and chartwork. This is followed by a week-long sea phase where students have the opportunity to apply their theoretical knowledge and develop their professional leadership qualities. Lastly, students complete a three-week simulator phase where they learn the key responsibilities of Navigation Officer, Officer of the Watch, and Fixing Officer.

Shortly after completion of Phase III training, students begin their Phase IV training. The purpose of Phase IV training is to prepare students for their Bridge Watch-Keeping qualification, which is obtained during a subsequent ship posting. The 85-day Phase IV course includes three successive phases: a) academic, b) simulator, and c) sea. The five-week academic phase includes radar theory, navigation, ship handling, engineering, stability, meteorology, joining from ahead, introductory electronic chartwork, collision regulations, communications, and Officer of the Watch. During the five-week simulator phase, students practice Officer of the Watch manoeuvres and Second Officer of the Watch skills. Finally, students complete an eight-week sea phase where they are assessed as Officer of the Watch and Second Officer of the Watch.

Previous research indicated pass rates of approximately 75% for Phase III training and 67% for Phase IV training (Bradley, 1990). Recent data suggests that overall pass rates are much lower: 52% for Direct Entry Officers and 68% for officers enrolled under

the Continuing Education Officer Training Plan (Boyes, 2008). In the current study, the final pass rate for Phase III training was 83% and the pass rate for Phase IV training was 92%; there was no significant difference in final pass rates between Direct Entry Officers and officers enrolled under the Continuing Education Officer Training Plan for both training phases. Given that officers are often provided a second and possibly a third training attempt in each phase, initial pass rates may be a better reflection of performance. In the current study, the initial pass rate for Phase III training was 67% and 82% for Phase IV training. Of those who failed their first attempt on Phase III, the majority (71%) failed the academic portion of the course. Among those who failed their first attempt at Phase IV, the majority (92%) failed the practical portion of the course. There was no significant difference in initial pass rates for both training phases between Direct Entry Officers and officers enrolled under the Continuing Education Officer Entry Plan. Analyses in this study included both the first pass/fail and the final attempt pass/fail results for Phase III and Phase IV training.

Criterion Problems. One problem in validation studies is criterion contamination, which occurs when extraneous variables affect the measured outcome. In the case of Maritime Surface and Subsurface Phase III and IV training, all courses are conducted according to established training standards. Extraneous factors may include the change in training staff, varying weather conditions at sea, the time of year, and the student cohort. Of these variables, only the time of year has been statistically shown to relate to training outcomes; Okros (1988b) observed higher initial pass rates in training serials that began in August, compared to serials that began in January or June. This

pattern of different pass/fail rates by time of year for Phase III was re-examined in the current study; no significant differences were found among training serials by month.

Results

Preliminary Analyses

Preliminary analyses included tests of the assumptions inherent to logistic regression and confirmed that the following assumptions were met: a) absence of multicolinearity among the predictors, b) no specification errors regarding the inclusion of all relevant predictors and exclusion of irrelevant predictors, and c) the scale of measurement for the predictors are either summative, interval, ratio, or categorical (Meyers, Gamst, & Guarino, 2006). Descriptive statistics confirmed that scores for the five predictors have a normal distribution.

Correlational Analyses

To test hypotheses 1 through 9, bivariate correlations were computed for all variables. Means, standard deviations, and intercorrelations for the criterion and predictors related to hypotheses 1 and 3 through 9 are presented in Table 1. Means, standard deviations, and intercorrelations for the criterion and variables related to hypotheses 2a and 2b are presented in Table 2. A consolidated table of all variables is presented in Appendix A. Corrections for range restriction in the predictor were calculated for the correlations between the Maritime Officer Selection Test and the criterion variables using the population standard deviation provided in Scholtz (2003).

The population standard deviation for the remaining predictors was unknown, therefore corrected correlations were not calculated.

Cognitive ability. Hypothesis 1a predicted that cognitive ability, as measured by the Canadian Forces Aptitude Test, would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. The total score percentile on the Canadian Forces Aptitude Test was significantly correlated with first attempt pass/fail results ($r = .28, p < .01$) on Phase III training and with final pass/fail results ($r = .34, p < .01$).

Hypothesis 1b predicted that cognitive ability, as measured by the Canadian Forces Aptitude Test, would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. The total score percentile on the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results ($r = -.04, p = .74$), nor with final pass/fail results ($r = .01, p = .93$) on Phase IV training.

Verbal Skills. Hypothesis 2a predicted that the verbal skills subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. The verbal skills subscale of the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results ($r = 0.01, p = .94$), nor with final pass/fail results ($r = .07, p = .47$), on Phase III training.

Hypothesis 2b predicted that the verbal skills subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface

and Subsurface Phase IV training. The verbal skills subscale of the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results ($r = -.12$, $p = .30$), nor with final pass/fail results ($r = .04, p = .73$), on Phase IV training.

Spatial ability. Hypothesis 2a predicted that the spatial ability subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. The spatial ability subscale of the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results on Phase III ($r = .16, p = .15$), but was significantly correlated with final pass/fail results on Phase III ($r = .21, p < .05$).

Hypothesis 2b predicted that the spatial ability subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. The spatial ability subscale of the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results ($r = .13$, $p = .27$), nor with final pass/fail results ($r = .21, p = .09$) on Phase IV training.

Problem-solving. Hypothesis 2a predicted that the problem-solving subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase training. The problem-solving subscale of the Canadian Forces Aptitude Test was significantly correlated with first attempt pass/fail results ($r = .23, p < .05$) and with final pass/fail results ($r = .27, p < .01$) on Phase III training.

Hypothesis 2b predicted that the problem-solving subscale of the Canadian Forces Aptitude Test would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. The problem-solving subscale of the Canadian Forces Aptitude Test was not significantly correlated with first attempt pass/fail results (r = -.02, p = .90), nor with final pass/fail results (r = -.004, p = .97) on Phase IV training.

Military potential. Hypothesis 3a predicted that scores of military potential from the Canadian Forces recruiting centre would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. Military potential was not significantly correlated with first attempt pass/fail results on Phase III training (r = -.04, p = .74), nor with final pass/fail results on Phase III training (r = -.05, p =.64).

Hypothesis 3b predicted that scores of military potential from the Canadian Forces recruiting centre would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. Military potential was not significantly correlated with first attempt pass/fail results (r = -.15, p = .21), nor with final pass/fail results (r.= -.14, p = .25) on Phase IV training.

Information-processing. Hypothesis 4a predicted that information-processing, as measured by the Maritime Officer Selection Test, would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. Information-processing was significantly correlated with first attempt pass/fail results (r = .36, p < .01) and with final pass/fail results (r = .33, p <.01) on Phase III training.

Table 1

Means, Standard Deviations, and Intercorrelations for Maritime Surface and Subsurface Training Phases and Performance Predictors[a]

	M	*SD*	1	2	3	4	5	6	7	8	9
1. MARS III (final fail/pass)[b]	0.83	0.38	—								
2. MARS III (fail/pass on first attempt)[b]	0.67	0.47	—	—							
3. MARS IV (final fail/pass)[b]	0.92	0.28	—	—	—						
4. MARS IV (fail/pass on first attempt)[b]	0.82	0.38	—	—	—	—					
5. CFAT[c]	60.90	24.43	.34**	.28**	.01	-.04	—				
6. Military potential[d]	53.28	5.57	-.05	-.04	-.14	-.15	.08	—			
7. MOST[eh]	31.08	6.84	.33** (.36)	.36** (.40)	.29* (.32)	.30** (.33)	.43**	-.06	—		
8. NOAB File review[f]	12.69	1.42	.07	.08	-.01	-.002	.34**	.52**	.11	—	
9. NOAB interview[g]	22.95	2.91	.03	-.05	-.11	-.08	.25*	.28**	.11	.44**	—

Note. [a] $n = 104$ for MARS III officers and predictors, $n = 74$ for MARS IV officers. MARS = Maritime Surface and Subsurface. CFAT = Canadian Forces Aptitude Test (percentile score). MOST = Maritime Officer Selection Test. NOAB = Naval Officer Assessment Board.

[b]Dichotomous criterion variables coded as 0 = fail, 1 = pass.

Constructs measured: [c]cognitive ability; [d]personality, person-environment fit, and education; [e]information-processing; [f]training and experience; and [g]naval officer potential based on leadership, self confidence, and oral communication.

[h]Corrections for range restriction are shown in brackets for the MOST.

*$p < .05$. **$p < .01$.

Table 2

Means, Standard Deviations, and Intercorrelations for Maritime Surface and Subsurface Training Phases[a] *and Specific Abilities*[b]

	M	*SD*	1	2	3	4	5	6	7
1. MARS III (final fail/pass)[c]	0.83	0.38	—						
2. MARS III (fail/pass on first attempt)[c]	0.67	0.47	—	—					
3. MARS IV (final fail/pass)[c]	0.92	0.28	—	—	—				
4. MARS IV (fail/pass on first attempt)[c]	0.82	0.38	—	—	—	—			
5. Verbal skills	67.90	22.14	.07	-.01	.04	-.12	—		
6. Spatial ability	62.74	26.26	.21*	.16	.21	.13	.26**	—	
7. Problem-solving	58.94	25.95	.27**	.23*	-.004	-.02	.26**	.42**	—

Note. [a] *n* is between 63 and 100 for MARS III officers and predictors due to missing scores, *n* = 71 for MARS IV officers.
[b]Percentile scores were used for the specific abilities subscales.
[c]Dichotomous criterion variables coded as 0 = fail, 1 = pass.
MARS = Maritime Surface and Subsurface.
*$p < .05$. **$p < .01$.

Hypothesis 4b predicted that information-processing, as measured by the Maritime Officer Selection Test, would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. Information-processing was

significantly correlated with first attempt pass/fail results ($r = .30, p < .01$) and with final pass/fail results ($r = .29, p < .05$) on Phase IV training.

Training and experience. Hypothesis 5a predicted that training and experience as measured by scores on the file review would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. Training and experience was not significantly correlated with first attempt pass/fail results on Phase III training ($r = .08, p = .45$), nor with final pass/fail results on Phase III training ($r = .07, p = .49$).

Hypothesis 5b predicted that training and experience as measured by scores on the file review would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase IV training. Training and experience was not significantly correlated with first attempt pass/fail results ($r = -.002, p = .98$), nor with final pass/fail results ($r = -01, p = .91$), on Phase IV training.

Naval officer potential. Hypothesis 6a predicted that naval officer potential, as measured by combined scores on the constructs of leadership, self-confidence, and oral communication of the Naval Officer Assessment Board interview, would be significantly correlated with pass/fail results on Maritime Surface and Subsurface Phase III training. Naval officer potential was not significantly correlated with first attempt pass/fail results ($r = -.05, p = .61$), nor with final pass/fail results ($r = .03, p = .75$).

Hypothesis 6b predicted that naval officer potential, as measured by combined scores on the constructs of leadership, self-confidence, and oral communication of the Naval Officer Assessment Board interview, would be significantly correlated with

pass/fail results on Maritime Surface and Subsurface Phase IV training. Naval officer potential was not significantly correlated with first attempt pass/fail results ($r = -.08, p = .50$), nor with final pass/fail results ($r = -.11, p = .36$) on Phase IV training.

Cognitive ability and information-processing. Hypothesis 7 predicted that cognitive ability, as measured by the Canadian Forces Aptitude Test, would be correlated with information-processing, as measured by the Maritime Officer Selection Test. Cognitive ability was significantly correlated with information-processing ($r = .43, p < .01$).

Training and experience, and cognitive ability. Hypothesis 8 predicted that training and experience, as measured by the Naval Officer Assessment Board file review, would be correlated with cognitive ability, as measured by the Canadian Forces Aptitude Test. Training and experience was significantly correlated with cognitive ability ($r = .34$, $p < .01$).

Training and experience, and military potential. Hypothesis 9 predicted that training and experience, as measured by the Naval Officer Assessment Board file review, would be correlated with military potential, as measured by the recruiting centre structured interview. Training and experience was significantly correlated with military potential ($r = .52, p < .01$).

Incremental Validity Analyses

With a dichotomous (pass/fail) outcome, hierarchical logistic regression analyses were conducted to test hypotheses 10a and 10b. Hypothesis 10a predicted that the Naval

Officer Assessment Board predictors at stage two of the selection process would show incremental predictive validity above the Canadian Forces Recruiting Centre predictors at stage one of the Maritime Surface and Subsurface selection process, against first and final pass/fail results on Maritime Surface and Subsurface Phase III training.

To analyze the incremental prediction for first attempt outcomes, the Phase III first attempt pass/fail outcome was entered as the dependent variable. In the first step, cognitive ability and military potential were entered as the covariates. In the second step, information-processing, training and education, and naval officer potential were entered as the covariates. The -2 Log likelihood (-2LL) at step one was 108.493. The -2LL at step two was 99.663. The significance of the change in -2LL was tested using the chi-square likelihood test (L-squared). The L-squared test of incremental prediction was significant: L-squared(1) = 8.83, $p < .005$.

To analyze the incremental prediction for final attempt outcomes, the Phase III final attempt pass/fail outcome was entered as the dependent variable. In the first step, cognitive ability and military potential were entered as the covariates. In the second step, information-processing, training and education, and naval officer potential were entered as the covariates. The -2 Log likelihood (-2LL) at step one was 82.636. The -2LL at step two was 76.777. The significance of the change in -2LL was tested using the chi-square likelihood test (L-squared). The L-squared test of incremental prediction was significant: L-squared(1) = 5.859, $p < .05$.

Hypothesis 10b predicted that the Naval Officer Assessment Board predictors at stage two of the selection process would show incremental predictive validity above the Canadian Forces Recruiting Centre predictors at stage one of the Maritime Surface and Subsurface selection process, against first and final pass/fail results on Maritime Surface and Subsurface Phase IV training.

To analyze the incremental prediction for first attempt outcomes, the Phase IV first attempt pass/fail outcome was entered as the dependent variable. In the first step, cognitive ability and military potential were entered as the covariates. In the second step, information-processing, training and education, and naval officer potential were entered as the covariates. The -2 Log likelihood (-2LL) at step one was 67.116. The -2LL at step two was 58.667. The significance of the change in -2LL was tested using the chi-square likelihood test (L-squared). The L-squared test of incremental prediction was significant: L-squared(1) = 8.449, $p < .005$.

To analyze the incremental prediction for final attempt outcomes, the Phase IV final attempt pass/fail outcome was entered as the dependent variable. In the first step, cognitive ability and military potential were entered as the covariates. In the second step, information-processing, training and education, and naval officer potential were entered as the covariates. The -2 Log likelihood (-2LL) at step one was 40.147. The -2LL at step two was 33.520. The significance of the change in -2LL was tested using the chi-square likelihood test (L-squared). The L-squared test of incremental prediction was significant: L-squared(1) = 6.627, $p < .05$.

Discussion

The purpose of the current study was to examine the criterion-related validity of a multistage selection process for selecting Maritime Surface and Subsurface officers into the Canadian Forces. More specifically, the present study examined the contribution of both the Canadian Forces Recruiting Centre measures and the Naval Officer Assessment Board measures in predicting performance on Maritime Surface and Subsurface officer training phases. In addition, the study sought to determine whether constructs measured in the second selection stage (i.e. during the Naval Officer Assessment Board) predicted success on Maritime Surface and Subsurface training phases above constructs measured in the first selection stage (i.e. at Canadian Forces Recruiting Centres). The results showed that one construct in stage one (i.e., cognitive ability) and one construct in stage two (i.e., information-processing) were valid predictors of Phase III training performance. Only one predictor (i.e., information-processing at stage two) was a valid predictor of Phase IV training performance. In addition, the Naval Officer Assessment Board predictors at stage two of selection showed incremental predictive validity over the Canadian Forces Recruiting Centre predictors at stage one of selection for both Phase III and Phase IV training.

Cognitive Ability as a Predictor of Training Performance

Based on military and civilian research (e.g., Boswell & Kuschnereit, 2009, Salgado, 1998, Schmidt & Hunter, 1998), cognitive ability was expected to be one of the best predictors of training performance. Results showed that cognitive ability, as measured by the Canadian Forces Aptitude Test at stage one of the selection process,

significantly predicted first attempt and final performance on Phase III training. Hence, hypothesis 1a was supported and results were consistent with previous findings on Phase III training (Bradley, 1990; Hain, 2003). The correlations of .28 with first attempts at Phase III and .34 with final attempts at Phase III are also consistent with previous research on the relationship between cognitive ability and training performance, showing uncorrected correlations of .18 to .46 (Olea & Ree, 1994; Salgado, 1995).

In contrast, cognitive ability was not a significant predictor of first attempt or final performance on Phase IV training; hypothesis 1b was not supported. Although the finding is consistent with previous research on Phase IV training (Bradley, 1990; Hain 2003), the reason for this result remains unclear. Perhaps the restriction of range in the predictor and in the pass/fail criterion, coupled with a small sample size, resulted in a lack of variance sufficient enough to detect a significant correlation between the Canadian Forces Aptitude Test and Phase IV training outcomes. Another possible explanation relates to the degradation of predictive validity over time. Studies suggest that prediction deterioration for cognitive ability is pervasive in skill acquisition (see Keil & Cortina, 2001 for a review); this might account, in part, for the lack of predictive validity of the Canadian Forces Aptitude Test in Phase IV in the current study. The prediction deterioration for cognitive ability does not follow a smooth curve in long-time-span studies (i.e., six months to five years) depending on the nature of the task (Keil & Cortina, 2001). In this study, the time-span issue is confounded by the fact that the length of time required to complete both phases varies by student. Depending on re-tests, medical delays, and training serial start dates, the time-span between the start of Phase III

and the completion of Phase IV could vary as much as 3 years among students. As a result, the Canadian Forces Aptitude Test may lose predictive validity over time for Phase III and Phase IV training, but this explanation would require further testing using training time-span data for each student. The data in the current study are not sufficient to fully support the prediction deterioration model.

Lastly, the results suggest that cognitive ability may not be the best predictor of Phase IV training. Given the course's heavy demands on leadership and bridgemanship, success might depend on non-cognitive factors. For instance, a study by Ackerman and Kanfer (1993) on air traffic controller training performance showed that self-efficacy was correlated .42 with training performance, surpassing certain elements of cognitive ability in predicting training success.

Specific Abilities as Predictors of Training Performance

The results of the current study indicate that some specific abilities predict Maritime Surface and Subsurface training outcomes, but others do not. Therefore, hypotheses 2a and 2b were not fully supported. Based on previous studies on the Canadian Forces Aptitude Test (Girard, 2004, Hodgson, 2005, Scholtz, 2004), it was expected that verbal skills would be a significant predictor of Phase III and Phase IV training. In this study, verbal skills was not a significant predictor of either phase. It is possible that the lack of a significant relationship between verbal skills and training performance in the current study is related to differences in assessment between Maritime Surface and Subsurface courses and occupations included in previous studies. More specifically, studies by Girard (2004), Hodgson (2005), and Scholtz (2004) included

occupations (i.e., clerks, military police, and stewards) that require a lot of non-technical written assessments (e.g., memos, incident reports), whereas Maritime Surface and Subsurface training involves more technical writing assessments. As a result, verbal skills might play a less important role in Maritime Surface and Subsurface training phases than in other occupations.

Spatial ability predicted final Phase III performance, but it did not predict first attempts at Phase III performance, nor did it predict Phase IV outcomes. The correlation between spatial ability and final attempt at Phase III was identical to the correlation between spatial ability and final attempt at Phase IV ($r = .21$), but the correlation was only significant for Phase III. The lack of statistical power (i.e., the lack of a sample size large enough to detect a significant relationship) may explain the lack of significance for Phase IV.

Problem-solving predicted both first and final attempts at Phase III, but it did not predict Phase IV outcomes. Previous research (e.g., Carretta, 2009; Kanfer et al., 2010) suggests that problem-solving might predict academic performance, but not procedural performance. Therefore, the lack of predictive validity for problem-solving in Phase IV might be because procedural performance is assessed to a greater extent than academic performance. The practical portion of Phase IV was the predominant source of course failures (92%), supporting the view that problem-solving is not a good predictor of Phase IV because the training on that phase places greater emphasis on the practical aspects, as opposed to the academic aspects of training.

Military Potential and Training Performance

The military potential (as defined in this study) assessment at stage one of the Maritime Surface and Subsurface selection process is designed to assess personality, person-environment fit, and education. Results of this study showed that, taken together, some personality traits, person-environment fit, and education did not significantly predict performance on Maritime Surface and Subsurface training phases; hypotheses 3a and 3b were not supported. These results are somewhat inconsistent with published research (e.g., Driskell et al., 1994; Huffcutt et al., 2001; Marcus et al., 2007) which suggests there is a correlation between each of the three predictors and training performance. At the same time, the results of the current study are consistent with research by Girard (2009) and Skomorovsky (2009) that found the military potential interview components (i.e., personality and person-environment fit) did not predict training outcomes. The lack of construct validity of the personality component may partially explain the results. In 2010, a series of revised questions was introduced into the military potential structured interview and researchers began pilot testing a new personality inventory for use in selection (i.e., the TSD-PI). These new measures should be included in subsequent predictive validity analyses for Maritime Surface and Subsurface officer selection.

Information-Processing and Training Performance

Information-processing was measured using the Maritime Officer Selection Test. Consistent with previous studies (e.g., Bradley, 1990; Carretta, 2000), information-processing was a significant predictor of training performance in the current study. The

Maritime Officer Selection Test predicted first and final attempts for both Phase III and Phase IV training; hypotheses 4a and 4b were supported. Results of uncorrected correlations indicate that the true correlation between information-processing and phase training outcomes could be as high as .40 for first attempts at Phase III training. The correlation between information-processing and cognitive ability in this study ($r = .43, p < .01$), thus hypothesis 7 was supported; the finding was consistent with previous research findings (e.g., Okros, 1988a; Vernon & Jensen, 1984). These results suggest that the Maritime Officer Selection Test, administered at stage two of the selection process, is not a redundant measure, even though it is related to cognitive ability which is measured at stage one of the selection process.

In addition, information-processing is the only construct that predicted performance in both phases of training. This finding is not surprising for two reasons. First, the Maritime Officer Selection Test was designed to tap into information-processing elements that were specifically identified through a job analysis of Maritime Surface and Subsurface officers (see Rodgers, 1986; Rodgers & Zuliani, 1985). Given that the last job analysis was conducted 25 years ago, the findings of the current study suggest that information-processing is related to long-standing, fundamental elements of Maritime Surface and Subsurface officer training. Second, research on perceptual speed—an underlying element of information-processing—suggests that it maintains predictive validity over time in some circumstances (Keil & Cortina, 2001). Therefore, the results in this study appear to be consistent with previous findings related to elements of information-processing.

Training and Experience, and Training Performance

In the current study, training and experience were assessed using the point method (Porter, Levine, & Flory, 1976) during a file review at the second stage in the Maritime Surface and Subsurface officer selection process. Results showed that the file review was not a significant predictor of Phase III or Phase IV training performance; hypotheses 5a and 5b were not supported. This finding is not consistent with previous studies on the Naval Officer Assessment Board file review (e.g., Bradley, 1990), nor with published studies on training and experience (e.g., McDaniel, Schmidt, & Hunter, 1988b). The lack of predictive validity could be related to the reliability of the file review and/or to the relevance of the file review dimensions to the role of a Maritime Surface and Subsurface officer. This issue requires further exploration.

As expected, the correlation between the file review and the Canadian Forces Aptitude Test was significant ($r = .34, p < .01$); hypothesis 8 was supported. In addition, the correlation between the file review and military potential was significant ($r = .52, p < .01$); hypothesis 9 is supported.

Naval Officer Potential and Training Performance

During a structured interview at the Naval Officer Assessment Board (i.e., at stage two of the selection process), a panel of interviewers assess candidates on the constructs of leadership, self-confidence, and oral communication, which are combined into a score of naval officer potential. Results of the current study indicate that naval officer potential was not a predictor of first or final attempts on Phase III and Phase IV training; hypotheses 6a and 6b were not supported. This finding is consistent with

previous research on the Naval Officer Assessment Board interview (Bradley, 1990; Hain, 2003).

There are several possible explanations for the current results. First, the job analysis conducted by Rodgers and Zuliani (1985) included oral expression as one of the top 18 critical abilities for successful performance as a Maritime Surface and Subsurface officer; leadership and self confidence were not among the top 18 abilities. Thus, one might conclude that leadership and self confidence may not be critical to performance as a Maritime Surface and Subsurface officer. However, given that leadership and bridgemanship (i.e., a term referring to assertiveness and self confidence on the bridge of the ship) are assessed during Phase III and Phase IV training, this explanation seems somewhat unlikely, suggesting that the job analysis no longer reflects all the current critical Maritime Surface and Subsurface officer abilities. Second, there may be issues regarding the reliability and validity of the structured interview. The construct validity of the structured interview has never been established. If the structured interview questions are not effective measures leadership, self confidence, and oral communication—constructs deemed critical in Phase III and Phase IV training—this may explain the lack of predictive validity of the structured interview. Lastly, measures of self confidence and oral communication have little structure in the interview; ratings are based on the impressions of the raters who follow minimal rating guidelines. Interviews with a low structure tend to produce lower mean validities (see Huffcutt et al., 2001).

Incremental Validity and Multistage Selection

Within multistage selection systems, each additional stage in the selection process should demonstrate significant incremental predictive validity over the previous stage(s). Results of the current study indicate that the second stage of the two-stage selection process for Maritime Surface and Subsurface officer provides incremental validity over the first stage. More specifically, the constructs measured at the Naval Officer Assessment Board (i.e., information-processing, training and education, and naval officer potential) provide incremental prediction over constructs measured at the Canadian Forces Recruiting Centre (i.e., cognitive ability and military potential) for both Phase III and Phase IV training. Therefore, hypotheses 10a and 10b were supported.

Limitations of the Current Study

Criterion problem. Although pass/fail on Maritime Surface and Subsurface training phases is a valid criterion, it lacks the variance that could be offered by criterion such as academic grades and practical assessment grades on Phase III and Phase IV training. Therefore, a pass/fail criterion is a less precise measure of training performance. Endeavours are currently underway to facilitate the collection of detailed performance data on each student attending Maritime Surface and Subsurface officer training at the Naval Officer Training Centre in Esquimalt, British Columbia (Lieutenant-Commander R.J. St-Pierre, personal communication, April 12, 2010). A database containing grades for all training assessments will enable researchers to conduct predictive validity analyses on the separate academic and practical training components of Phase III and Phase IV training.

Predictor problems. Two of the predictor methods in this study contain two or more predictor constructs. The military potential assessment (as defined in this study) conducted at the recruiting centre is designed to measure personality, person-environment fit and education. The naval officer potential interview conducted at the Naval Officer Assessment Board is designed to measure leadership, self confidence, and oral communication. Unfortunately, data for these individual constructs were not available for the current study. This issue is less of a concern with the military potential component because it has subsequently been modified. Nonetheless, future research on Maritime Surface and Subsurface officer selection should examine the predictive validity of the military potential constructs separately. The naval officer potential interview, on the other hand, remains relatively unchanged except for the removal of one leadership question. Without detailed information on the scores by construct, the current study is limited in its analysis of the predictive validity of each construct. To determine which constructs are the most predictive of Phase III and Phase IV training, future research should examine the naval officer potential constructs separately.

Implications for Future Research

The results of the current study are consistent with previous studies showing that cognitive ability is one of the best predictors of performance (e.g., Schmidt & Hunger, 1998), that information-processing is a significant predictor of training performance (Arthur, Doverspike, et al., 2004), and that both constructs vary in their predictive validity over time (Keil & Cortina, 2001). Keil and Cortina (2001) called for further research that would test their model of validity degradation over time by examining both

cognitive ability and information-processing across different types of tasks (i.e., consistent and inconsistent). Keil and Cortina's (2001) model combines Ackerman's (1987, 1988) model of validity degradation over time with a cusp catastrophe model of skill acquisition. More specifically, the model posits that there is a smooth, continuous relationship between experience and performance for those with low ability or for those performing simple, consistent tasks. In contrast, the model suggests that that performance follows a step function when individuals start with a higher ability or for those performing complex, inconsistent tasks. The Maritime Surface and Subsurface officer training phases represent an ideal environment for testing Keil and Cortina's (2001) model. Phase III and Phase IV training include both simple and complex tasks, consistent and inconsistent tasks, and students with a good range of scores on cognitive ability and information-processing.

Implications for Maritime Surface and Subsurface Officer Selection

It takes up to seven years to train Maritime Surface and Subsurface officers to the operationally functional point. This represents a substantial investment of human and monetary resources by the Canadian Forces. Therefore, selecting the right people is critical for the return on investment to the Canadian Forces through trained naval officers. Also, as the second stage in a multistage selection system, the Naval Officer Assessment Board represents an additional cost beyond selection at the recruiting centre. The cost of a multistage selection system is justified when predictors at each successive stage offers incremental predictive validity over previous selection stages.

Results of the current study indicate that some constructs at both stages of Maritime Surface and Subsurface officer selection are valid predictors of training performance, while others are poor predictors of training performance. At first glace, the finding that the file review and the Naval Officer Assessment board interview are poor predictors of Phase III and Phase IV training may make it difficult to justify the cost of sending candidates to the Naval Officer Assessment Board for selection purposes. After all, the Maritime Officer Selection Test is a timed paper-and-pencil test that could be administered at the recruiting centre along with the Canadian Forces Aptitude Test. However, the results show that, overall, the Naval Officer Assessment Board measures at stage two offer incremental validity over the Canadian Forces Recruiting Centre measures at stage one. Moreover, research suggests that the realistic job preview offered by the Naval Officer Assessment Board is an important component in promoting success on training (Phillips, 1998).

Research studies also suggest that the other constructs measured at the Naval Officer Assessment Board (i.e., training and experience, leadership, self confidence, and oral communication) are valid predictors of performance in both civilian and similar military settings (e.g., Huffcutt et al., 2001; McDaniel, Schmidt, & Hunter, 1988b; Walters, Miller, & Ree, 1993). In some cases, non-cognitive measures may be better predictors of certain aspects of performance than cognitive ability (e.g., Atwater et al., 1999). Moreover, the constructs of leadership, self confidence, and oral communication are all important elements in Phase III and Phase IV training. Therefore, the lack of predictive validity of the file review and the naval officer potential interview does not

appear to stem from inappropriate constructs; results from an on-going job analysis may serve to confirm this. Rather, the problem may be related to how the constructs are measured (i.e., the reliability and construct validity of the file review and the interview) or to the measurement methods.

Recommendations

1. The Canadian Forces should retain the Maritime Officer Selection Test as a selection tool for Maritime Surface and Subsurface officers.

2. The Canadian Forces should retain the Canadian Forces Aptitude Test as a selection tool for Maritime Surface and Subsurface officers.

3. If on-going job analysis findings support the use of training and experience, leadership, self confidence, and oral communication constructs in the selection of Maritime Surface and Subsurface officer, further research should be conducted to assess the reliability and construct validity of the file review and the interview components of the Naval Officer Assessment Board during future boards. Using well established measures of the four constructs in conjunction with existing measures would allow testing of both the construct validity of the existing measures and the predictive validity of individual constructs.

4. In light of the new military potential selection measures introduced at recruiting centres, future Maritime Officer Surface and Subsurface officer selection validation studies should examine the predictive validity of the new measures against Phase III and Phase IV training outcomes.

5. Future Maritime Surface and Subsurface officer selection validation studies should include academic and practical assessment grades, in addition to the pass/fail criterion for Phase III and Phase IV training.

6. Future research should be conducted to assess the contribution of the realistic job preview component of the Naval Officer Assessment Board in predicting training performance on Phase III and Phase IV training.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117,* 288-318.

Ackerman, P.L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. *Journal of Applied Psychology, 78*, 413-432.

Albert, J. (1998). *Cognitive Measures: Comparison of the Canadian Forces Aptitude Test (CFAT), Raven's Progressive Matrices and the Wonderlic Personnel Test.* Technical Note 1998-05, Ottawa, ON: Director Personnel Research Team.

Angoff, W.H. (1971). Norms, scales, and equivalent scores. In R.L. Thorndike (Ed.). *Educational Measurement* (2nd Ed.). Washington, DC: American Council on Education.

Arthur, W., Jr., Bell, S.T., Villado, A.J., & Doverspike, D. (2006). The use of person-organization fit in employment decision making: An assessment of its criterion-related validity. *Journal of Applied Psychology, 91*, 786-801.

Arthur, W., Doverspike, D., & Bell, S. (2004). Information-processing tests. In J.C. Thomas (Ed.), *Comprehensive handbook of psychological assessment, Vol. 4: Industrial and organizational assessment* (pp. 56-74). Hoboken, NJ US: John Wiley & Sons Inc.

Arthur, W., Jr., & Villado, A.J. (2008). The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice. *Journal of Applied Psychology, 93*, 435-442.

Atwater, L.E., Dionne, S.D., Avolio, B., Camobreco, J.E., & Lau, A.W. (1999). A longitudinal study of the leadership development process: Individual differences predicting leader effectiveness. *Human Relations, 52*, 1543-1562.

Bambacas, M., & Patrickson, M. (2009). Assessment of communication skills in manager selection: Some evidence from Australia. *Journal of Management Development, 28*, 109-120. doi: 10.1108/02621710910932070.

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

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

Besetsny, L.K., Earles, J.A., & Ree, M.J. (1993). Little incremental validity for a special test for Air Forces intelligence operatives. *Educational and Psychological Measurement, 53*, 993-997.

Besetsny, L.K., Ree, M.J., & Earles, J.A. (1993). Special tests for computer programmers? Not needed. *Educational and Psychological Measurement, 53*, 507-511.

Blanc, S. J.-R. (2003). Naval officer selection in Canada: An evaluation of the Maritime Officer Selection Test (Maritime Officer Selection Test). M.Sc. dissertation, Saint Mary's University, Canada.

Boswell, R.A. (1993). *Review of naval officer selection methods in TTCP countries.* Technical Note 21/93, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

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

Boyes, F.A.J. (2008, December 8). *Naval Officer Assessment Board.* Briefing note prepared for the Director General Maritime Personnel and Readiness, RDIMS #161755, Ottawa, ON: Director Maritime Personnel.

Bradley, J.P. (1990). *A validation study of the Naval Officer Assessment Board's ability to predict Maritime Surface and Subsurface officer training success.* Working Paper 90-7, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Brown, K.G., Le, H., & Schmidt, F.L. (2006). Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment, 14*(2), 87-100.

Burnett, J.R., & Motowidlo, S.J. (1998). Relations between different sources of information in the structured selection interview. *Personnel Psychology, 51*, 963-983.

Canadian Forces Leadership Institute (2005). *Leadership in the Canadian Forces – Conceptual Foundations*. Kingston, ON: Canadian Defence Academy.

Carretta, T.R. (2000). U.S. Air Force pilot selection and training methods. *Aviation, Space, and Environmental Medicine, 71*, 950-956.

Carretta, T.R. (2009). *Predictiv e Validity of the Air Force Officer Qualifying Test for USAF Air Battle Manager Training Performance*. Interim Report AFRL-RH-WP-TR-2009-0007, Wright-Patterson Air Force Base, Ohio: Air Force Research Laboratory.

Catano, V., Wiesner, W., Hackett, R., & Methot, L. (2005). *Recruitment and Selection in Canada*. Toronto, ON: Nelson.

Christian, M.S., Edwards, B.D., & Bradley, J.C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83-117.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Crocker, L., & Algina, J. (2006). *Introduction to Classical & Modern Test Theory.* Mason, OH: Thompson Wadsworth.

Das, H. (2007). *Recruitment, selection and deployment of human resources: a Canadian perspective.* Toronto, ON: Pearson Education Canada.

Dayan, K., Kasten, R., & Fox, S. (2002). Entry-level police candidate assessment center: An efficient tool or a hammer to kill a fly? *Personnel Psychology, 55*, 827-849.

Dean, M.A., Conte, J.M., & Blankenhorn, T.R. (2006). Examination of the predictive validity of Big Five personality dimensions across training performance criteria. *Personality and Individual Differences, 41*, 1229-1239.

Digman, J.M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*, 116-123.

Dilchert, S., & Ones, D.S. (2009). Assessment center dimensions: Individual differences correlates and meta-analytic incremental validity. *International Journal of Selection and Assessment, 17*, 254-270.

Donohue, J.J. (2005). Standard Error of Measurement (SEM) of the Canadian Forces Aptitude Test (CFAT) for Canadian Forces Applicants (2003 – 2004). Technical Note 2005-07. Director Human Resources Research and Evaluation, National Defence Headquarters, Ottawa, Ontario, Canada.

Driskell, J.E., Hogan, J., Salas, E., & Hoskin, B. (1994). Cognitive and personality predictors of training performance. *Military Psychology, 6*(1), 31-46.

Erkstrom, R.B., French, J.W., & Hartman, H.H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monograph, 79*(2), 1-84.

Finch, D.M., Edwards, B.D., & Wallace, J.C. (2009). Multistage selection strategies: Simulating the effects on adverse impact and expected performance for various predictor combinations. *Journal of Applied Psychology, 94*, 318-340.

Gatewood, R.D., Feild, H.S., & Barrick, M. (2008). *Human Resource Selection (6th Edition)*. Mason, Ohio : Thomson/South-Western.

Girard, M.L. (2004). *Validation of the CFAT for Vehicle Technician Selection.* Technical Note 2004-02, Ottawa, ON: Director Military Human Resource Research and Evaluation.

Girard, M.L. (May 2009). *Principle Components Analysis of Interview Questions Used in Selection of Canadian Forces Personnel.* Technical Note 2009-003, Ottawa, ON: Director General Military Personnel Research and Analysis.

Grubb, W.L.III, Whetzel, D.L., & McDaniel, M.A. (2004). General mental ability tests in industry. In J.C. Thomas (Ed.), *Comprehensive handbook of psychological assessment, Vol. 4: Industrial and organizational assessment* (pp. 7-20). Hoboken, NJ US: John Wiley & Sons Inc.

Hain, C.A. (August 2003). *Naval Officer Assessment Board Validation Study to Predict MARS Officer Training Success*. Unpublished draft report, Ottawa, ON: Director Human Resource Research and Evaluation.

Hodgson, K.E. (2005). *Validation of the CFAT and Establishment of Cutoff Scores for Military Police Selection*. Technical Note 2005-03, Ottawa, ON: Director Military Human Resource Research and Evaluation.

Hough, L.M., & Oswald, F.L. (2000). Personnel selection: Looking toward the future—Remembering the past. *Annual Review of Psychology, 51*, 631-664.

Huffcutt, A.I., Conway, J.M., Roth, P.L., & Stone, N.J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology, 86*, 897-913.

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

Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Judge, T.A., Bono, J.E., Ilies, R., & Gerhardt, M.W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765-780.

Judge, T.A., Colbert, A.E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology, 89*, 542-552.

Kanfer, R., Wolf, M.B., Kantrowitz, T.M., & Ackerman, P.L. (2010). Ability and trait complex predictors of academic and job performance: A person-situation approach. *Applied Psychology: An International Review, 59*, 40-69. doi: 10.1111/j.1464-0597.2009.00415.x.

Keil, C.T., & Cortina, J.M. (2001). Degradation of validity over time: A test and extension of Ackerman's model. *Psychological Bulletin, 127*, 673-697.

Kickul, J., & Neuman, G. (2000). Emergent leadership behaviors: The function of personality and cognitive ability in determining teamwork performance and KSAs. *Journal of Business and Psychology, 15*, 27-51.

Leahy, J.M. (2008). *Examining the Construct Validity of the Canadian Forces Aptitude Test*. Unpublished Master's Thesis, Saint Mary's University, Halifax, Nova Scotia.

LeBreton, J.M., & Senter, J.L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815-852. doi: 10.1177/1094428106296642.

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

Lievens, F., Buyse, T., & Sackett, P.R. (2005). Retest effects in operational selection settings: Development and test of a framework. *Personnel Psychology, 58*, 981-1007.

Marcus, B., Goffin, R., Johnston, N., & Rothstein, M. (2007). Personality and Cognitive Ability as Predictors of Typical and Maximum Managerial Performance. *Human Performance*, *20*(3), 275-285. doi:10.1080/08959280701333362.

McDaniel, M.A., Schmidt, F.L., & Hunter, J.E. (1988a). A meta-analysis of the validity of methods for rating training and experience in personnel selection. *Personnel Psychology, 41*, 283-314.

McDaniel, M.A., Schmidt, F.L, & Hunter, J.E. (1988b). Job experience correlates of job performance. *Journal of Applied Psychology, 73*, 327-330.

McGrath, R.E. (2008). Predictor combination in binary decision-making situations. *Psychological Assessment, 20*(3), 195-205. doi: 10.1037/a0013175.

Meyers, L.S., Gamst, G., & Guarino, A.J. (2006). *Applied multivariate research.* Thousand Oaks, CA: Sage Publications.

Motowidlo, S.J., & Burnett, J.R. (1995). Aural and visual sources of validity in structured employment interviews. *Organizational Behavior and Human Decision Processes, 61*, 239-249.

Ng, T.W.H., & Feldman, D.C. (2009). How broadly does education contribute to job performance? *Personnel Psychology, 62*, 89-134.

Oakes, D.W., Ferris, G.R., Martocchio, J.J., Buckley, M.R., & Broach, D. (2001). Cognitive ability and personality predictors of training program skill acquisition and job performance. *Journal of Business and Psychology, 15*, 523-548.

Okros, A.C. (1988a). *Development of the Passage Planning Test: A Maritime Surface and Subsurface officer specific selection test.* Technical Note 10/88, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Okros, A.C. (1988b). *Analysis of Maritime Surface and Subsurface classification training results.* Technical Note 1/88, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Okros, A.C., Johnston, V.W., & Rodgers, M.N. (1988). *An evaluation of the effectiveness of the Naval Officer Selection Board as a predictor of success on the basic officer training course.* Working Paper 88-1, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Okros, A.C., & Lynn, D.G. (1989). *Development of the problem sensitivity test: A Maritime Surface and Subsurface officer specific selection test.* Technical Note 3/89, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Olea, M.M., & Ree, M.J. (1994). Predicting pilot and navigator criteria: Not much more than *g*. *Journal of Applied Psychology, 79*, 845-851.

Personnel Psychology Directive 102 (2006). A-PM-241-001 *Personnel Selection Services Manual*. Ottawa, ON.

Phillips, J.M. (1998). Effects of realistic job previews on multiple organizational outcomes: A meta-analysis. *Academy of Management Journal, 41*, 673-690.

Porter, W.R., Levine, E.L., & Flory, A. III. (1976). *Training and experience evaluation.* Tempe, AZ: Personnel Services Organization.

Quiñones, M.A., Ford, J.K., & Teachout, M.S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology, 46*, 887-910.

Ree, M.J., & Earles, J.A. (1991). The stability of *g* across different methods of estimation. *Intelligence, 15*, 271-278.

Rodgers, M.N. (1984). *Overview of the Naval Officer Production Systems*. Technical Note 12/84, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Rodgers, M.N. (1986). *The identification of ability requirements for MARS officer selection.* Technical Note 23/86, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Rodgers, M.N., & Johnston, V.W. (1985). *An evaluation of the effectiveness of Naval Officer Interview Board serials 8305 to 8404.* Technical Note 18/85, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Rodgers, M.N., & Zuliani, R.A. (1985). *A critical incident analysis and ability assessment of first-tour MARS officer tasks.* Technical Note 19/85, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

Rohde, T.E., & Thompson, L.A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83-92.

Rost, J.C. (1993). *Leadership for the Twenty-First Century.* Westport, Connecticut: Praeger.

Roth, P.L., BeVier, C.A., Switzer, F.S.III, & Schippmann, J.S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology, 81*, 548-556.

Sacket, P.R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology, 85*, 112-118.

Salgado, J.F. (1995). Situational specificity and within-setting validity variability. *Journal of Occupational and Organizational Psychology, 68*, 123-132.

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*, 262-274.

Schmidt, F.L., Shaffer, J.A., & Oh, I.S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology, 61*, 827-868.

Scholtz, D. (2003). *Establishing defensible cutoff scores for the Naval Officer Assessment Board.* Research Report, Ottawa ON: Director of Maritime Command.

Scholtz, D.C. (2004). *Validation of the CFAT and Establishment of Cutoff Scores for Steward Selection.* Technical Note 2004-01, Ottawa, ON: Director Military Human Resource Research and Evaluation.

Skomorovsky, A. (May 2009). *Structured Interview Questionnaire: Examining Convergent and Predictive Validity in a Sample of Privates*. Technical Memorandum 2009-04, Ottawa, ON: Director General Military Personnel Research and Analysis.

Society for Industrial and Organizational Psychology, Inc. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.

Stouffer, J.M. (1996). *Psychometric Evaluation of the Canadian Forces Aptitude (Canadian Forces Aptitude Test) and the Passage Planning (Passage Planning Test) Tests*. Technical Note 1/96, Willowdale, ON: Canadian Forces Personnel Applied Research Unit.

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

Vanderpool, M.A. (2003a). *Aptitude Test for Selecting Canadian Aboriginal Peoples*. Technical Note 2003-04, Ottawa, ON: Director Military Human Resource Research and Evaluation.

Vanderpool, M.A. (2003b). *Determining if the Canadian Forces Aptitude Test is Adversely Impacting Canadian Aboriginal Peoples*. Technical Note 2003-03, National Defence Headquarters, Ottawa, ON: Director Military Human Resource Research and Evaluation.

Vernon, P.A., & Jensen, A.R. (1984). Individual and group differences in intelligence and speed of information-processing. *Personality and Individual Differences, 5*, 411-423.

Walters, L.C., Miller, M.R., & Ree, M.J. (1993). Structured interviews for pilot selection: no incremental validity. *The International Journal of Aviation Psychology, 3*, 25-38.

Appendix A

Means, Standard Deviations, and Intercorrelations for Demographic Variables, Maritime Surface and Subsurface Training Phases, and Performance Predictors[a]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Sex (male/female)[b]	—														
2. Language[c]	—	—													
3. Entry Plan (DEO/CEOTP)[d]	—	—	—												
4. MARS III (final fail/pass)[e]	—	—	—	—											
5. MARS III (fail/pass on first attempt)[e]	—	—	—	—	—										
6. MARS IV (final fail/pass)[e]	—	—	—	—	—	—									
7. MARS IV (fail/pass on first attempt)[e]	—	—	—	—	—	—	—								
8. CFAT[f]	-.22*	.06	.03	.34**	.28**	.01	-.04	—							
9. Verbal Skills	-.22*	.11	.10	.07	-.01	.04	-.12	.48**	—						
10. Spatial Ability	-.21*	.11	.17	.21*	.16	.21	.13	.61**	.26**	—					
11. Problem-solving	-.21*	.08	.09	.27**	.23*	-.004	-.02	.80**	.26**	.42**	—				
12. Military potential[g]	.15	.08	-.36**	-.05	-.04	-.14	-.15	.08	.05	.01	-.01	—			
13. MOST[h]	-.02	.09	.001	.33**	.36**	.29*	.30**	.43**	.01	.36**	.41**	-.06	—		
14. NOAB File review[i]	.15	.09	-.11	.07	.08	-.01	-.002	.34**	.21*	.10	.24*	.52**	.12	—	
15. NOAB interview[j]	.06	.22*	-.07	.03	-.05	-.11	-.08	.25*	.25*	.03	.09	.26**	.12	.44**	—
M	.17	.06	.30	.83	.67	.92	.82	60.90	67.90	62.74	58.94	53.28	31.08	12.69	22.95
SD	.38	.23	.46	.38	.47	.28	.38	24.43	22.14	26.26	25.95	5.57	6.84	1.42	2.91

Note. [a] $n = 104$ for MARS III officers and predictors, $n = 74$ for MARS IV officers.
MARS = Maritime Surface and Subsurface. CFAT = Canadian Forces Aptitude Test (percentile scores for total score and subscales). MOST = Maritime Officer Selection Test. NOAB = Naval Officer Assessment Board.
[b]Sex: 0 = male, 1 = female. [c]Language: 0 = English, 1 = French. [d]Entry Plan: 0 = DEO (Direct Entry Officer), 1 = CEOTP (Continuing Education Officer Training Plan). [e]Dichotomous criterion variables coding: 0 = fail, 1 = pass.
Constructs measured: [f]cognitive ability; [g]personality, person-environment fit, and education; [h]information-processing; [i]training and experience; and [j]naval officer potential based on leadership, self confidence, and oral communication.
*$p < .05$. **$p < .01$.