Big Five Personality Research in the Military: A Meta-analysis

Wendy Darr
Selection and Assessment
Personnel Generation Research

DGMPRA TM 2009-023
November 2009

Defence R&D Canada
Director General Military Personnel Research & Analysis

Chief Military Personnel

National Defence Défense nationale

Canada

Big Five Personality Research in the Military: A Meta-analysis

Wendy Darr
Selection and Assessment
Personnel Generation Research

Director General Military Personnel Research & Analysis

Technical Memorandum
DGMPRA TM 2009-023
November 2009

Author

(Original signed by)

Wendy Darr, PhD

Approved by

(Original signed by)

Douglas Pelchat, BA

Section Head – Personnel Generation Research

Approved for release by

(Original signed by)

Kelly Farley, PhD

Chief Scientist – Director General Military Personnel Research and Analysis

The opinions expressed in this paper are those of the authors and should not be interpreted as the official position of the Canadian Forces, nor of the Department of National Defence.

Abstract

Despite established validity of the Big Five factors of personality, validity generalization is limited by evidence for situational specificity. Unique requirements and expectations demanded of military personnel warrant the examination of personality-outcomes associations gathered in such samples. Using accumulated research into the validity of the Air Force Self-descriptive Inventory (AFSDI) across several military organizations, this examination involved a meta-analysis to determine the validity of the Big Five personality factors as measured by the AFSDI or variants of it. Effects are comparable with mainstream meta-analytic estimates, being slightly higher for some effects. Conscientiousness was confirmed to be a strong predictor of military performance, and generalized across samples. Although moderated by the type of measure (i.e., original/adapted length) and/or the type of military stream (non-commissioned members/officers), Neuroticism and Extraversion were the next best predictors of performance and officer training, respectively.

Résumé

Malgré la validité établie des cinq grands facteurs de la personnalité, la généralisation de la validité est restreinte par l'évidence d'une spécificité situationnelle. Des exigences et des attentes uniques demandées au personnel militaire garantissent l'examen d'associations axées sur les résultats de la personnalité rassemblées dans de tels échantillons. Utilisant la recherche réalisée sur la validité de la *Self-descriptive Inventory* de la force aérienne des États-Unis (AFSDI) à l'intérieur de nombreuses organisations militaires, l'examen renferme une méta-analyse visant à déterminer la validité des cinq grands facteurs de la personnalité, mesurée par l'AFSDI ou ses versions. Les effets sont comparables aux estimations principales de la méta-analyse, étant légèrement supérieurs dans certains cas. Le souci du travail bien fait a été confirmé comme prédicteur important du rendement militaire et généralisé dans des échantillons. Bien que modérés par le type de mesure (longueur originale/adaptée) et/ou le grade des militaires (militaires du rang/officiers), le névrosisme et l'extraversion ont été, avec la formation des officiers, dans cet ordre, les meilleurs prédicteurs du rendement.

This page intentionally left blank.

Executive summary

Big Five Personality Research in the Military: A Meta-analysis:

Wendy Darr; DGMPRA TM 2009-023; Defence R&D Canada – DGMPRA.

It is now well known that personality is related to important job outcomes such as performance and training. However, consolidated findings from 15 meta-analytic studies (e.g., Barrick, Mount, & Judge, 2001) suggest that the strength of some personality-outcome associations depend on situational factors. Indeed, Tett, Jackson, and Rothstein (1991) argued for situational specificity, demonstrating the influence of various situational or study characteristics (e.g., exploratory versus confirmatory research approach, occupational group, type of criterion) on the validity of personality factors. Unique requirements and expectations demanded of military personnel (e.g., Soloman, 1954; French & Ernest, 1955) suggest likely trait differences between military and civilian samples that echo underlying motivational differences, which could in turn strengthen or weaken certain personality-outcome linkages in military settings.

The Big Five measure of personality known as the Air Force Self-descriptive Inventory (AFSDI), originally developed by the US Air Force, has been used or adapted for use in selection research by several military organizations. With more than a decade of accumulated research, the objective of this study was to quantitatively synthesize personality-performance associations accumulated through research across military settings, to determine whether findings can be generalized across various studies, and to examine comparability with mainstream personality research. A total of 117 effects from 20 independent samples (17 studies) were available for analyses.

Findings suggest that adaptive versions of the AFSDI moderate the predictor-outcome associations for the factors of Neuroticism and Extraversion. Accounting for the moderating effects of adaptations to the measure, overall findings offer evidence for the validity of the Big Five personality factors. Job performance is most strongly related to Conscientiousness and this association generalizes across all military samples, which is also consistent with mainstream research. When the 163-item measure is used, the effects for Neuroticism and job performance are also stronger than those obtained in mainstream research. With respect to training performance, a small generalizable effect was obtained for Conscientiousness, followed by Neuroticism. The negative effect for Neuroticism and training performance is not altered by the type of measure or sample (i.e., NCM versus Officer). Extraversion (when measured by the 163-item version) has a strong positive association with Officer training performance compared to NCM training.

This consolidation of military findings paints a clearer picture for militaries seeking to incorporate specific personality factors in their selection decisions. It also identifies areas in which further local development or validation work is required. Recommendations for improving the validity of personality and guidance for future validation research are provided.

Sommaire

Big Five Personality Research in the Military: A Meta-analysis:

Wendy Darr; DGMPRA TM 2009-023; R & D pour la défense Canada – DRASPMNovembre 2009.

Il est bien connu que la personnalité est liée aux résultats important au plan du travail, comme le rendement et la formation. Toutefois, les conclusions réunies de 15 études méta-analytiques (p. ex., Barrick, Mount, & Judge, 2001) suggèrent que la force de certaines associations en matière de personnalité dépend de facteurs situationnels. En fait, Tett, Jackson et Rothstein (1991) appuient la spécificité situationnelle, démontrant l'influence de certaines caractéristiques situationnelles ou études (p.ex., l'approche exploratoire par rapport à l'approche de recherche confirmative, groupe professionnel, type de critère) sur la validité des facteurs de personnalité. Des exigences et des attentes uniques demandées au personnel militaire (p.ex., Soloman, 1954; French & Ernest, 1955) suggèrent des différences de personnalité probables entre les échantillons militaires et civils qui reflètent des différences motivationnelles sous-jacentes, pouvant également renforcer ou affaiblir certains liens avec la personnalité dans les organisations militaires.

La mesure des cinq grands facteurs de personnalité, connue sous le non de *Self-descriptive Inventory* de la force aérienne (AFSDI), élaborée à l'origine par la force aérienne des États-Unis, a été utilisée ou adaptée pour être utilisée dans la recherche de sélection par plusieurs organisations militaires. Comptant plus d'une décennie de recherche cumulée, l'objectif de cette étude était de synthétiser quantitativement des associations axées sur le rendement de la personnalité accumulées par le biais de la recherche dans diverses études, et d'examiner la comparabilité avec la recherche principale sur la personnalité. Au total, 117 effets tirés de 20 échantillons indépendants (17 études) ont servi aux analyses.

Les conclusions suggèrent que les versions adaptatives de l'AFSDI modèrent les associations axées sur les résultats des prédicteurs pour les facteurs de névrosisme et d'extraversion. Justifiant les effets modérateurs des adaptations à la mesure, les conclusions générales offrent la preuve de la validité des cinq grands facteurs de personnalité. Le rendement au travail est fortement lié au souci du travail bien fait. Cette association se retrouve dans tous les échantillons militaires, un aspect cohérent avec la recherche principale. Quand la mesure de 163 articles est utilisée, les effets liés au névrosisme et au rendement au travail sont également plus forts que ceux obtenus dans la recherche principale. Concernant le rendement à l'entraînement, un petit effet généralisable a été obtenu à l'égard du souci du travail bien fait, suivi du névrosisme. L'effet négatif du névrosisme et du rendement à l'entraînement n'est pas modifié par le type de mesure ou d'échantillon (p.ex., militaire du rang par rapport à officier). L'extraversion (quand elle est mesurée à l'aide de la mesure à 163 articles) a une forte association positive avec le rendement à l'entraînement des officiers, comparativement à l'entraînement des militaires du rang.

Le regroupement des conclusions militaires donne une image plus claire aux militaires qui veulent intégrer les facteurs particuliers de la personnalité dans leurs décisions de sélection. Il détermine également les domaines nécessitant de plus amples développements ou travaux de validation. On y trouve enfin des recommandations visant à améliorer la validité de la personnalité et à fournir des conseils à l'égard d'autres études de validation.

Table of contents

List of tables

1 Introduction

1.1 Mainstream Personality Research

It is now well-known that personality is related to important job outcomes such as performance and training. Barrick, Mount, and Judge (2001) consolidated findings from 15 meta-analytic studies to demonstrate the validity of the Big Five personality factors. While their secondary meta-analysis demonstrated that variations in some personality-outcome effects can be explained by differences in sample size, there was also evidence for situational specificity. In other words, additional variance remained to be explained even after accounting for sample size and artifacts such as measurement error, suggesting that the strength of some personality-outcome associations epend on other factors. Indeed, Tett, Jackson, and Rothstein (1991) argued for and demonstrated that various situational or study characteristics (e.g., exploratory versus confirmatory research approach, occupational group, type of criterion) can influence the validity of personality factors.

In particular, Tett *et al*. (1991) found slightly higher validities for military samples in comparison to civilian ones. These findings were counter to that hypothesized, and was explained by the influence of one large study of military recruits. The use of military samples was hypothesized to diminish the validities of personality because of similar findings examining the validity of cognitive ability in military samples (e.g., Vineberg & Joyner, 1982), where the problematic use of subjective criteria by military evaluators was explained as the contributing factor. As explained below, differences between military and civilian organizational settings or systems lead one to question the generalizability of mainstream personality research to military settings, the majority of which has involved civilian settings or samples (Tett *et al*., 1991).

1.2 The Military Context

Soloman's (1954) sociological account of military recruits' transformation from civilian to soldier highlights the "distinctive requirements" (p. 87) of military life, which involves incorporating and internalizing new behaviours and attitudes, and conforming to expectations. Military jobs are also described as being more physically and psychologically demanding than most civilian jobs. (Krueger, 2001). Pulakos, White, Oppler, and Borman (1989) factor analysed performance ratings across 19 military occupational specialties to find three main factors, two of which according to Sackett & Dubois (1999), are not typically emphasized in civilian organizations. These included personal discipline (adhering to rules, exercising self control, and demonstrating integrity) and military bearing (maintaining appropriate military appearance).

In Holloman's (1967) examination of leadership, he found differential informal expectations for military versus civilian supervisors working within the same rigidly defined military structure. Military supervisors were expected to be lower in consideration and higher in structure than civilian supervisors. The implication of this finding to personality validity research is that interpersonal traits (reflective of consideration) might be more relevant to managerial performance in civilian samples than in military samples. In their examination of retention, Capon, Chernyshenko, Stark (2007) found that contrary to civilian research findings, military personnel experiencing work-life conflict were not more dissatisfied and were not more likely

to leave the organization, suggesting that certain individual characteristics might buffer the effects of such stressors in military personnel more strongly than they do in civilian personnel.

Personality is thought to influence outcomes primarily through underlying motivational processes (e.g., Barrick, Mount, & Strauss, 1993; Mount & Barrick, 1995), and motivational differences across military and civilian samples offers another reason for expecting differential predictor-outcome effects. In Doll and Gunderson's (1969) examination of volunteers who signed up for a year-long assignment to Antarctica, a stronger job satisfaction- job performance relationship for a group of civilian scientists versus Navy enlistees was explained by underlying differences in motivation. The scientists reported being oriented towards accomplishing tasks and goals relevant to their scientific projects, whereas the military enlistees offered a variety of non job-specific reasons (e.g., saving money, adventure, experience) for signing up.

Clearly, there are unique requirements and expectations of military personnel in comparison to their civilian counterparts. French and Ernest (1955) discussed that individuals predisposed to authoritarianism might be more accepting of the authoritarian military ideology or environment. Kurpius and Lucart's (2000) comparison of military and civilian undergraduates revealed that military students did indeed have higher levels of authoritarianism, suggesting that likely trait differences between military and civilian samples could echo underlying motivational differences that could strengthen or weaken certain personality-outcome linkages in military settings.

1.3 Personality Research in Military Organizations

Personality research in military organizations has progressed equally well, beginning with the early identification of five personality factors (Tupes & Christal, 1961) through work conducted for the United States Air Force, and subsequent research and development of the Air Force Self Descriptive Inventory (AFSDI) by Christal and colleagues (e.g., Christal, Barucky, Driskill, & Collis, 1997). The AFSDI was developed as a 163-item computer-administered measure of the Five Factor Model (FFM) of personality. Through a collaborative initiative within the Technical Cooperation Program (TTCP), several militaries including Canada, United Kingdom, New Zealand, and Australia have adapted the AFSDI for selection research within their own organizations.

Adaptations of the AFSDI include the reduction of its length, replacement of the original Arch scale with a Likert-type rating scale, conversion to paper-and-pencil format, Anglicization of some of items, and translation into the French language. Strong convergence, in the range of .92 to .97, has been established between factors assessed by the computer-administered and paper-and-pencil versions of the AFSDI (Christal *et al*., 1997), and between existing personality measures and various later adaptations of the original AFSDI. For example, in examining the 163-item paper-and-pencil version, Black and Skomorovsky (2007) reported strong correlations between factors of the AFSDI and Costa and McCrae's (1994) NEO Personality Inventory (NEO-PI) (N = .87, E = .76, O = .68, A = .60, and C = .71). In addition, Boyes' (2005) examination of a 75-item version showed strong convergence with respective factors assessed by the NEO-PI (N = .82, E = .63, O = .75, A = .64, and C = .78) and Lee and Ashton's (2004) HEXACO (N = .64, E = .79, O = .75, A = .43, and C = .79).

In a narrative review of criterion-related military research on the AFSDI and its variants, Syed and Klammer (2001) reported mixed evidence for the ability of this measure to predict job-

related outcomes. They attributed these mixed findings to variations in sample size, criterion measure, and sample demographics (i.e., non-commissioned members,NCM/Officer) across studies conducted by these five countries. With the accumulation of validity research spanning over a decade is an opportunity to determine the generalizability of the AFSDI or its variants in predicting military performance.

1.4 Objectives

Using meta-analytic techniques, the objective of this study is to quantitatively synthesize personality-performance associations accumulated through research across military settings, to determine whether findings can be generalized across various studies, and also to examine comparability with mainstream personality research. Findings from this paper will provide information on the relative strength of each personality factor in predicting various outcomes, which is particularly useful to those looking to use the AFSDI or its variants in making personnel selection decisions.

2 Methodology

2.1 Literature Search

A list of relevant empirical studies (i.e., ones that used variants of the AFSDI) was first compiled from that reviewed in Syed and Klammer (2003). In addition, a request for more recent studies was sent to the all military organizations participating on TTCP in the fall of 2008. The International Military Testing Association (IMTA) conference proceedings was accessed online, and searched for relevant studies using keywords such as "personality", and inventory names used by other militaries, namely "Trait Self Descriptive Inventory" or "OCEAN." Another search was undertaken within the PsychArticles database using the same search keywords.

2.2 Data Compilation

In order to be included in this meta-analysis, a study had to measure personality using the AFSDI or an adapted version. It had to be empirical in nature, and had to report zero-order correlations (*r*) between any of the five personality factors and some performance criterion of interest. In addition to recording basic statistics (e.g., *r*, reliability), each study was coded on several variables of interest: length of TSD measure (original 163 item or revised to another length), sample [non-commissioned member (NCM) or officer), type of criterion [job performance, training performance, counterproductive behaviour (includes turnover, and intent to leave), and leadership], and recruit/incumbent samples. The criterion was also coded to determine whether a self- or other- report measure was used. For studies reporting criterion-relevant group comparisons on personality factor scores, a *d*-estimate of effect size was first computed and then converted to *r* (there was only one such study). A coding sheet (see Annex A) was developed for compiling this data.

2.3 Meta-Analytic Procedure

The Hunter-Schmidt Meta-analysis program, version 1.1 (Schmidt & Le, 2004) was used to synthesize the compiled effects.

2.3.1 Artifact Corrections

Effects were corrected for measurement error using artifact distributions (i.e., aggregate estimates of the reliabilities). Corrections to the predictor were based on factor scale reliability estimates reported in studies included in the present meta-analysis. Estimates for each factor were available for between 11-13 samples, and the average reliability for each factor is reported in Table 1. With respect to criterion reliability, estimates were available from only 3 studies, all of which used self-report criterion measures. Therefore, corrections for unreliability in the criterion were based on the approach used in Tett *et al*. (1991). Specifically, Rothstein's (1990) average single-rater interrater reliability estimate (to represent reliabilities for other-report criteria) and the average estimates reported or calculated for self-report criteria in the present study were used to calculate a sample-weighted average estimate (see Table 1). The sample-weighted average is, therefore, the sum of these two estimates, weighted by the number of

effects using other-report criteria (89) and self-report criteria (28). Effect sizes were not corrected for range restriction because of difficulties in (i) obtaining estimates of variance in the restricted and unrestricted samples, and (ii) identifying other non-personality variable(s) likely contribute to range restriction (i.e., cognitive measure, medical criteria, other criteria).

Table 1: Average Reliability Estimates used for Attenuation Corrections

Measure	*k*	Average Reliability	SD
Predictor Measures			
Neuroticism	12	.903	.051
Extraversion	13	.855	.121
Openness	11	.874	.053
Agreeableness	11	.865	.105
Conscientiousness	13	.794	.223
Criterion Measures			
Single rater interrater average reliability (other-report criteria) (Rothstein, 1990)	72	.508	.071
Average reported or composite reliability (present examination, self-report criteria)	3	.772	.052
Sample-weighted average reliability	117	.571	.066

2.3.2 Independence Assumptions

In meta-analyzing findings using the Hunter and Schmidt (1990) approach, there is a requirement for the accumulated estimates to be statistically independent. Studies having fully replicated designs (i.e., measures and associations are reported for different organizations), conceptually similar measures, and sub-sample analyses often report more than one estimate. Based on Hunter and Schmidt, estimates from studies with fully replicated designs (i.e., correlations are reported for each organization separately) were treated as being independent and were all included in the meta-analysis. For studies using conceptually similar multiple measures (e.g., supervisor ratings of task and contextual performance), a composite correlation (based on the intercorrelation of the conceptually similar measures and their associations with the target variable), or the average correlation (whenever intercorrelations were not reported) was used. Composite correlations were calculated using procedures outlined in Nunnally (1978) and Hunter and Schmidt (2004). Composite correlations were not calculated for cases in which multiple estimates of effect size enabled a comparison across certain variables of interest (e.g., NCM and Officer samples). Because personality has been shown to have different relationships with various criteria, composite correlations were calculated within, but not between, multiple measures of job performance, training performance, leadership, and counter-productive work behaviours (CWB).

2.3.3 Credibility and Confidence Intervals

Credibility and confidence intervals were calculated in accordance with Whitener (1990). Credibility interval computations utilized the posterior distribution of effects (i.e., those corrected for sampling and measurement error), and were calculated before the calculation of confidence intervals. Confidence interval computations depended on the width of the credibility interval (homogeneous or heterogeneous), and were based upon observed or sample size-weighted effects. Non-zero confidence intervals provided a test for the significance of a particular effect size, while the Z-test (Hunter & Schmidt 1990a, p. 437) was used to compare two effect sizes.

2.3.4 Moderator Analyses

In determining the presence of moderators in this synthesis, a conjunctive combination of the 75 percent rule and credibility intervals was used (Cortina, 2003). In other words, the operation of substantial moderators was deemed likely whenever the amount of variance explained by artifacts was less than 75 percent and the credibility intervals were wide or contained zero values. There is presently no established criterion for determining how wide a credibility interval should be to signal the presence of moderators. A credibility interval ranging from .04 to .48 was interpreted by Whitener (1990) as being positive and thus the search for additional moderators was deemed unnecessary. Yet, Kuncel, Hezlett, and Ones (2001) simply examined the lower value of the credibility interval to determine whether it was positive or negative. In their meta-analysis, a positive lower credibility value with substantial unexplained variance was interpreted as signaling some fluctuation in the effect's magnitude across settings. In the present study, the percent of variance explained by artifacts dictated the use of a particular confidence interval (CI) formula. When this percent exceeded 75, CI computations were based on a homogeneous sample.

3 Results

A total of 117 effects from 20 independent samples (17 studies) was available for analyses. Of these, three were based on studies conducted in the Royal Army, three from the Royal Navy, one from the US AirForce, one from the Australian Navy, and 12 from the Canadian Forces. To maintain the independence assumptions, multiple studies using the same sample and criterion measures could not be included; in such cases, the larger sample was retained for analysis. There were four such cases. Very few studies reported the average age or gender make-up of their samples to permit an adequate description of the sample included in this meta-analysis. Seventy eight percent of the studies included in this meta-analysis used recruit samples, with a relatively equal number of NCM and Officer samples. Obtained average observed effects are described below for various sub-groups and outcomes.

3.1 Length of the Measure

As indicated earlier, the original 163-item AFSDI has been revised and adapted for use across various military organizations. For example, the Canadian Forces reduced the measure to 75 items, whereas the Royal Navy has reported use of a 148- and 172-item measure. Consequently, there is need to compare effects based on the original length of the measure with those based on other altered lengths. Synthesizing across all criteria, effects were grouped by length of the measure (163 versus other). Note that the direction of effects involving CWB criteria was reversed for this particular analysis.

The effects for each factor are reported in Table 2. All effects appear to be comparable with the exception of that for Neuroticism and Extraversion, where the average effect for these factors obtained using the 163-item measure is almost twice as large as that compared to effects obtained when an adapted version is used. However, a Z-test of comparison revealed a significant difference only for Extraversion ($Z = 2.54$, $p < .05$). Nevertheless, effects based on the 163-item versus other adapted versions are analyzed separately (whenever a sufficient number of effects permit such a comparison) for Extraversion and Neuroticism.

3.2 Work-related Outcomes

To compare the relative strength of each personality factor in predicting various work outcomes, effects were synthesized by criterion variables (i.e., grouping common criteria together). Training performance (which comprised basic training results) was examined the most frequently, followed by job performance. Only a couple of studies examined CWB, turnover, and intent to leave; these were grouped together because of their small number and also because they are thought to represent behaviours or intentions that are counter to an organization's legitimate interest (Sackett & DeVore, 2001). Table 3 presents the results of these analyses for each personality factor.

Table 2: Personality-criterion Associations for Length of TSD Measure

	k (#effects)	N	Observed r	Corrected r	90 Percent Cred Int	Percent Artifacts	95 Percent CI
Neuroticism							
TSD 163	12	4606	-.12	-.17	-.32 to -.01	35.28%	-.17 to -.07
TSD other	11	2131	-.06	-.08	-.38 to .23	23.46%	-.15 to .03
Extraversion							
TSD 163	13	4791	.12	.18	.07 to .30	54.91%	.09 to .16
TSD other	11	2131	.03	.05	-.13 to .22	49.87%	-.03 to .09
Openness							
TSD 163	12	4687	.04	.05	-.04 to .15	61.36%	.00 to .07
TSD other	11	2131	.01	.02	-.11 to .15	62.78%	-.04 to .07
Agreeableness							
TSD 163	11	4225	.02	.03	-.11 to .17	42.05%	-.03 to .07
TSD other	11	2131	.01	.02	-.29 to .31	24.92%	-.07 to .10
Conscientiousness							
TSD 163	14	5253	.13	.20	-.02 to .42	18.08%	.06 to .20
TSD other	11	2131	.13	.20	-.06 to .47	33.74%	.06 to .21

Table 3: Criterion-related Validity Estimates for the Big Five Personality Factors

	k (#effects)	N	Observed r	Corrected r	90 Percent Cred Int	Percent Artifacts	95 Percent CI
Neuroticism							
Job Perf.	6	1774	-.16	-.22	-.34 to -.10	54.05%	-.22 to -.09
TSD 163	3	1352	-.20	-.28	-	100.00%	-.15 to -.25
TSD other	3	422	-.02	-.03	-	100.00%	-.11 to .08
Training Perf.	12	2744	-.09	-.13	-.24 to -.02	66.01%	-.14 to -.04
TSD 163	7	2002	-.09	-.13	-.01 to -.26	55.34%	-.03 to -.15
TSD other	5	742	-.09	-.12	-.06 to -.18	90.93%	-.02 to -.16
CWB	3	965	.23	.33	.18 to .48	42.09%	.15 to .32
Ldrship potential	2	1254	.01	.01	-	100.00%	-.05 to.06
Extraversion							
Job Perf.	6	1774	.13	.19	.04 to .34	47.60%	.06 to.20
TSD 163	3	1352	.17	.24	-	100.00%	.11 to .22
TSD other	3	422	.03	.04	-.07 to .16	75.24%	-.08 to .14
Training Perf.	12	2744	.05	.07	-.09 to .23	50.70%	-.01 to .10
TSD 163	7	2002	.08	.11	-.02 to .25	54.08%	.02 to .14
TSD other	5	742	-.04	-.06	-	100.00%	-.12 to .03
CWB	3	965	-.09	-.13	-.30 to.04	38.66%	-.19 to .01
Ldrship potential	3	1439	.15	.22	-	100.00%	.10 to .20
Openness							
Job Perf.	5	1393	-.01	-.02	-.12 to .08	65.01%	-.08 to .05
Training Perf.	12	2744	.05	.07	-	100.00%	.01 to .09
CWB	4	1427	-.10	-.15	-	100.00%	-.15 to -.05
Ldrship potential	2	1254	-.05	-.08	-	100.00%	-.11 to .00

	k (#effects)	N	Observed r	Corrected r	90 Percent Cred Int	Percent Artifacts	95 Percent CI
Agreeableness							
Job Perf.	5	1393	.09	.13	-.02 to .28	47.30%	.01 to .17
Training Perf.	12	2744	-.02	-.02	-	100.00%	-.05 to .02
CWB	3	965	-.15	-.21	-.43 to .01	27.57%	-.27 to -.03
Ldrship potential	2	1254	-.09	-.13	-	100.00%	-.15 to -.04
Conscientiousness							
Job Perf.	6	1774	.23	.35	.30 to .39	94.72%	.18 to .27
Training Perf.	12	2744	.12	.18	-	100.00%	.08 to .15
CWB	4	1427	-.19	-.28	-.47 to -.10	41.71%	-.28 to -.10
Ldrship potential	3	1439	-.02	-.03	-.12 to .07	59.67%	-.09 to .05

As seen in Table 3, the effects for Neuroticism and Extraversion were analyzed separately for the 163-item measure versus other adapted versions. Although the type of measure did not moderate the effects of Neuroticism on training performance, it did influence the Neuroticism-job performance effect, being much stronger when the 163-item version was used ($Z = 19.51$, $p < .01$). Despite the small number of effects, Neuroticism also demonstrates a strong positive association with counter-productive work behaviours (CWBs). With respect to Extraversion, effects were analyzed separately (whenever there were enough effects to do so), because the length of the measure proved to be a moderator for this factor. Extraversion predicts job performance and training performance, but only with the 163-item version. Perhaps, removal of content relevant to the prediction of these two criteria may be responsible for the lack of an effect using versions of alternate lengths. Extraversion also has a significant positive association with leadership potential, although the number of studies contributing to this effect is small.

Of the criteria predicted by Openness to Experience, the effects for training performance and CWB are the only significant ones. Agreeableness predicts job performance and leadership potential equally well, being negatively related to the latter. Using the 75-percent rule, the strength of the agreeableness-job performance association is likely to be influenced by other factors. Finally, in comparison to the other personality factors, Conscientiousness appears to be the strongest predictor of job performance and training performance. Conscientiousness also negatively predicts CWBs, although this effect is likely to depend on other factors, as the percent of variance explained by artifacts is less than 75 percent.

3.3 Non-commissioned Member (NCM) and Officer Samples

To determine whether the predictive validity of personality differs across NCM and Officer samples, effect sizes were compared across these two samples. However, this comparison was possible only for training performance, because there were a number of estimates within each group to permit this analysis. Based on the results reported in Table 3, it made sense to perform these sub-group analyses only for those training performance effects for which less than 75% of the variance was explained by artifacts (i.e., effects for Neuroticism and Extraversion measured by the 163-item TSD which suggested the presence of additional moderators). Nevertheless, for the sake of comparison the effects for all personality factors are presented by NCM and Officer samples in Table 4.

A Z-test comparing the effects across NCM and Officer samples reveals a significant difference only for Extraversion. In other words, the association between training performance and Extraversion using the 163-item TSD measure is moderated by type of sample, being stronger and significant for the Officer sample. Officer training differs from that for NCMs in that it includes a leadership component. This finding is consistent with that reported in Table 3; of all the personality factors, Extraversion was found to have the strongest positive association with leadership potential.

Table 4: Effects for training performance across NCM and Officer samples

	k (#effects)	N	Observed r	Corrected r	90 Percent Cred Int	Percent Artifacts	95 Percent CI
Neuroticism							
Training Perf.	12	2744	-.09	-.13	-.24 to -.02	66.01%	-.14 to -.04
NCM	5	1184	-.12	-.17	-.21 to -.14	94.52%	-.18 to -.07
Officer	6	661	-.09	-.13	-.32 to -.05	59.12%	-.17 to -.02
Extraversion							
Training Perf.	12	2744	.05	.07	-.09 to .23	50.70%	-.01 to .10
TSD 163	7	2002	.08	.11	-.02 to .25	54.08%	.02 to .14
NCM	3	824	.01	.02	-.09 to .13	63.32%	-.07 to .10
Officer	3	278	.16	.23	—	100.00%	.05 to .28
Openness							
Training Perf.	12	2744	.05	.07	—	100.00%	.01 to .09
NCM	6	1482	.04	.06	.04 to .07	98.64%	-.01 to .09
Officer	6	660	.06	.09	—	100.00%	-.01 to .14
Agreeableness							
Training Perf.	12	2744	-.02	-.02	—	100.00%	-.05 to .02
NCM	5	1184	-.01	-.02	—	100.00%	-.07 to .05
Officer	6	660	-.06	-.08	—	100.00%	-.13 to .02
Conscientiousness							
Training Perf.	12	2744	.12	.18	—	100.00%	.08 to .15
NCM	5	1184	.11	.17	.11 to .24	87.29%	.06 to .17
Officer	6	660	.11	.16	—	100.00%	.03 to .18

3.4 Comparison with Published Meta-analytic Estimates

Given the large accumulation of independent meta-analytic examinations of the Big Five personality factors in mainstream research, validity estimates obtained in this study are compared with that accumulated thus far (see Table 5). To ensure comparable criteria, the effects for job performance, training performance, and leadership potential were compared against Barrick *et al*'s (2001) coefficients for supervisor ratings, training performance, and managerial performance, respectively. For CWB, a study weighted average of estimates for deviant behaviour and turnover reported in Salgado's (2002) meta-analysis was used as a comparison, as these most closely correspond to the nature of criteria used in the present analysis.

As seen in Table 5, the effects for Neuroticism are generally larger than that reported in published meta-analyses across all criteria except leadership potential. With respect to Extraversion, the effects in the present meta-analyses are generally larger with the exception of that for training performance. A comparable effect for this criterion is observed only for the Officer sample when the 163-item measure is used. For Openness, the largest discrepancies were observed for the CWB effect which was negative and larger in the present examination (r = -.15, k= 4) compared to published estimates (r= .05, k = 12), and for training performance where the published effect was stronger (r = .24, k = 18) than that obtained in the present examination even when the 163-item measure were synthesized separately (r= .11, k = 7). Effects for Agreeableness are comparable for job performance and CWB. The direction of the effect for leadership potential (although based on two effects) is negative in the present examination, but positive in mainstream research. With the exception of leadership potential, effects for Conscientiousness are comparable across all criteria, being slightly larger for job performance in the present examination.

Table 5: A Comparison of the Big Five Personality Validity Coefficients

	Present Meta-analytic Validity Estimates				Published Meta-analytic Validity Estimates			
	k (#effects)	**N**	**Observed r**	**Corrected r**	**k (#effects)**	**N**	**Observed r**	**Corrected r**
Neuroticism								
Job Perf.	6	1774	-.16	-.22	167	23687	-.07	-.12
TSD 163	3	1352	-.20	-.28				
TSD other	3	422	.03	.04				
Training Perf.	12	2744	-.09	-.13	25	3753	-.05	-.08
CWB	3	965	.23	.33	19	3661	.08	.12
Ldrship potential	2	1254	.01	.01	63	11591	-.05	-.08
Extraversion								
Job Perf.	6	1774	.13	.19	164	23785	.07	.11
TSD 163	3	1352	.17	.24				
TSD other	3	422	.03	.04				
Training Perf.	12	2744	.05	.07	21	3484	.13	.23
NCM (163)	3	824	.01	.02				
Officer (163)	3	278	.16	.23				
CWB	3	965	-.09	-.13	16	2937	-.03	-.04
Ldrship potential	3	1439	.15	.22	67	12602	.10	.17
Openness								
Job Perf.	5	1393	-.01	-.02	116	18535	.03	.05
Training Perf.	12	2744	.05	.07	18	3177	.14	.24
CWB	4	1427	-.10	-.15	12	1975	.03	.05
Ldrship potential	2	1254	-.05	-.08	44	8678	.05	.07

	Present Meta-analytic Validity Estimates				Published Meta-analytic Validity Estimates			
	k (#effects)	**N**	**Observed r**	**Corrected r**	**k (#effects)**	**N**	**Observed r**	**Corrected r**
Agreeableness								
Job Perf.	5	1393	.09	.13	151	22193	.06	.10
Training Perf.	12	2744	-.02	-.02	24	4100	.07	.11
CWB	3	965	-.15	-.21	13	1853	-.14	-.21
Ldrship potential	2	1254	-.09	-.13	55	9864	.04	.08
Conscientiousness								
Job Perf.	6	1774	.23	.35	185	33312	.15	.26
Training Perf.	12	2744	.12	.18	20	3909	.13	.23
CWB	4	1427	-.19	-.28	18	7024	-.18	-.27
Ldrship potential	3	1439	-.02	-.03	60	11325	.12	.21

4 Discussion

4.1 Comparison with Mainstream Research

One of the objectives of this research was to furnish validity estimates for each Big Five factor (based on military samples) to enable comparisons with published meta-analytic estimates that would address Tett *et al*'s (1991) concerns about situational specificity. Tett *et al*. argued that existing published meta-analyses (e.g., Barrick & Mount, 1991) do not adequately account for variations in study characteristics (e.g., military versus civilian settings), limiting the generalizability of accumulated personality validity estimates across these samples. With a focus on military samples, the present validity synthesis provides evidence for personality-outcome effects that are comparable to that obtained in mainstream research, substantiating the use of the AFSDI or its variants as a valid measure of personality.

The stronger associations for Conscientiousness and Neuroticism compared to mainstream meta analytic findings (the majority of which has focussed on civilian samples) are thought to result from stronger trait-role linkages. In other words, there is a higher degree of correspondence between features of the military role and certain individual requirements. The possibility that these higher coefficients could also be explained by common method variance is ruled out as the majority of these effects utilized other-report criteria. Therefore, Conscientiousness may be more predictive of performance in military personnel because military performance requirements such as exerting effort, persevering under adverse conditions, adhering to rules/regulations, exercising self-control, and demonstrating integrity (Pulakos *et al*., 1989) correspond well with conscientious individuals' capacity to work hard and to do whatever it takes to succeed at some goal (Roberts *et al*., 2005). In fact, Quinn, Sinha, Keeney, and Schmitt's (2009) recent examination of mediators demonstrated that effortful and ethical behaviours were responsible for the Conscientiousness-performance effect. Gade, Lakhani, and Kimmel (1991) also found that individuals high on traits similar to those that make up the Conscientiousness factor (i.e., self-confidence, self-discipline, and independence) tended to rate their experience with the military as being valuable compared to those low on these traits.

The stronger negative effect for Neuroticism and job performance is explained in terms of the greater physical and psychological demands of military jobs in comparison to civilian ones (Krueger, 2001), such that personnel who succeed at buffering themselves from such demands are more capable of performing well. Because individuals low in Neuroticism are also described as emotionally stable, they are more likely to deal effectively with the stressful demands of military jobs. They are, perhaps, more capable of demonstrating "toughness and practical hard-headedness" (French & Ernest, 1955, p.181) required to maintain effectiveness within a military work environment. This explanation is consistent with Capon *et al*'s (2007) finding, contrary to civilian research, that military personnel experiencing work-life conflict were not more dissatisfied and were not more likely to leave the organization.

4.2 Validity Generalization across Military Organizations

A second objective of this study was to examine the generalizability of the validity of personality across military organizations utilizing adapted versions of the original United States Air Force Self Descriptive Inventory (AFSDI). Findings suggest that the type of measure does moderate the validity of Neuroticism and Extraversion, with the 163-item version having stronger criterion-related associations than other versions (e.g., 75, 148, or 172 items). Perhaps, revisions to the original measure resulted in the removal of item content relevant to certain performance outcomes. For example, with respect to Neuroticism, Darr's (2009) facet-level examination of a 75-item adaptation of the AFSDI revealed that Neuroticism had the strongest convergence with facets of anxiety, depression, and vulnerability as measured by the commercial NEO Personality Inventory (NEO-PI). Yet, her review of criterion-related evidence for NEO-PI Neuroticism facets suggested that the facets of vulnerability and impulsiveness were more frequently related to performance outcomes of interest. Hence, removal of relevant content offers one explanation for the lower validities obtained with adapted versions of this personality measure.

With respect to training performance, a small generalizable effect was obtained for Conscientiousness followed by Neuroticism. The negative effect for Neuroticism and training performance is not altered by the type of measure or sample (i.e., NCM versus Officer). Extraversion (when measured by the 163-item version) has a strong positive association with Officer training performance compared to NCM training. This finding is not surprising given that the distinguishing feature of Officer training is its leadership component, and research has demonstrated a link between Extraversion and managerial performance/potential (e.g., Barrick *et al*., 2001; Craik, Ware, Kamp, O'Reilly, Staw, & Zedeck, 2002). Extraversion was also found to have the strongest association with leadership potential, although only three studies examined this outcome.

With respect to counter-productive work behaviours (CWBs), Neuroticism and Conscientiousness have equally strong effects, but the strength of these effects is likely to vary across military samples. Due to the small number of studies that examined CWBs as an outcome, it was not possible to identify likely moderators. Agreeableness was also found to have a small, negative relationship with CWBs, but this effect did not generalize across samples. Even though Agreeableness is conceptualized as an interpersonal dimension of personality (Wiggins, 1979), and has stronger empirical links to teamwork than to deviant behaviour (e.g., Barrick *et al*., 2001; Salgado, 2002), there is conceptual reason to expect Agreeableness to predict deviant behaviours, as one of the underlying bases for Agreeableness is the effortful control or self-regulation of emotions (Darr, 2009; Graziano & Eisenberg, 1997).

4.3 Limitations

One of the limitations of this meta-analytic examination was its focus on a single measure of personality, which perhaps limits generalizability across military organizations utilizing other personality measures. However, even with this limited focus, variations in effects were observed and attributed to adapted versions. Consequently, a source of variation attributed to the type of measure has been minimized in the present examination, strengthening the inferences for military organizations using variants of this specific measure. A second possible limitation is the small

number of studies that were included in this examination. With limited accessibility to research studies conducted in other military organizations, it is difficult to estimate what proportion of validation studies on the AFSDI or adaptive versions remain excluded.

4.4 Conclusions and Future Directions

In summary, there is evidence for the use of this measure in predicting relevant performance outcomes. The predictive utility of a particular personality factor, however, depends on the nature of performance criteria examined and adapted version used. This consolidation of military findings paints a clearer picture for militaries seeking to incorporate specific personality factors in their selection decisions. It also identifies areas in which further local development or validation work is required. For example, those assessing Neuroticism or Extraversion with adapted versions of the AFSDI might wish to re-examine scale content to improve the prediction of job performance and Officer training performance, respectively. In predicting counterproductive work behaviours, it may be beneficial to examine the potential of Agreeableness in adding to variance explained by Neuroticism and Conscientiousness for this particular outcome. In addition, unexplained variance for these effects suggests the presence of moderators which can be further explored.

In recommending approaches to improve the validity of personality in military settings, I draw upon Barrick *et al*'s (2001) suggestion for the need to first identify and understand specific criteria of interest, and then examine facet level linkages of each personality factor with these criteria. Understanding why some personality factor influences an outcome of interest will enable the identification of item content most relevant to the prediction of the outcome, consequently allowing for the customization of personality measures to maximize predictor-criterion correspondence. For example, Hough and Oswald (2008) discussed the creation of compound traits which are formed by grouping items from various facets together to maximize the amount of variance explained in some outcome of interest. Using findings from the present examination and those from Darr (2009), CWB is likely to be best predicted by a composite based on relevant facets of Conscientiousness, Neuroticism and Agreeableness. Consequently, future research may focus on the identification of such compound traits for predicting military relevant criteria.

As highlighted in the introduction and through some of the present findings, the unique expectations and demands of military work warrant the continued treatment and examination of military/civilian settings as an important contextual factor in personnel research. Consequently, there is a need to accumulate military research in a way that is accessible to those pursuing such examinations. Barrick *et al*. (2001) suggested the development of taxonomic frameworks within which hypotheses and related findings can be accumulated. For example, personality validation research in the military can be examined and accumulated within criterion categories such as that utilized in the present study (i.e., job performance, training performance, CWB, and leadership). As observed through the present examinations, examinations of CWB and leadership remain underexplored in comparison to those of training and job performance. Perhaps, future military research might expand their efforts to include such criteria in personality validation research. Finally, through data compilation efforts for the present study, it became apparent that the reporting of research findings requires much improvement. Reports often failed to include the full zero-order correlation matrices of all predictor and criterion variables (not simply ones that are significant), reliability estimates for criterion variables, reliability estimates for each sample

examined, and descriptions of the study's research design (whether predictive, concurrent, or postdictive). Such details are crucial to future secondary research efforts such as validity generalization studies.

5 References

* included in meta-analysis

[1] *Abram, C.M. & Elshaw, C.C. (1997). *Predicting voluntary withdrawal from the Royal Military Academy Sandhurst*. DERA/CHS/HS3/CR97071/1.0. Farnborough, Hampshire: Defence Evaluation and Research Agency.

[2] *Allender, C. & Greig, J. (2000). *Does the Army need an OCEAN*? Paper presented at the annual conference of the International Military testing Association, Edinburgh, United Kingdom.

[3] 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? *International Journal of Selection and Assessment, 9*, 9-30.

[4] Barrick, M.R., Mount, M.K., & Strauss, J.P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78*, 715-722.

[5] *Bhat, C.S. (1999). *Does personality influence training success in the Royal Australian Navy?* Unpublished master's thesis, Monash University, Australia.

[6] Black, L.E. & Skomorovsky, A. (2007). *An investigation of the psychometric properties of the Trait Self Descriptive Inventory using an Air Force Sample*. Technical Report 2007-11. Directorate of Human Resources Research and Evaluation, Ottawa, Ontario, Canada.

[7] *Boyes, F.A.J. (2005). *Personality as a predictor of military performance and counterproductive behaviour*. Unpublished master's thesis. Saint Mary's University, Halifax, Nova Scotia, Canada.

[8] Capon, J., Chernyshenko, O.S., & Stark, S. (2007). Applicability of civilian retention theory in the New Zealand military. *New Zealand Journal of Psychology, 36*, 50-56.

[9] *Christal, R.E., Barucky, J.M., Driskill, W.E., & Collis, J.M. (1997). *Informal technical final report. The Air Force Self Description Inventory (AFSDI): A summary of continuing research*. F33615-91-D-0010. Armstrong Laboratories, Brooks AFB.

[10] *Collis, J.M. (1995). *The Trait-Self-Description Inventory: Measurement of the "Big Five" personality factors at the Admiralty Interview Board*. DRA/CHS/HS3/CR95050/01. Farnborough, Hampshire: Defence Research Agency.

[11] Collis, J.M., & Barucky, J.M. (1996). '*Big Five" personality factors research using Christal's Self Description Inventory*. Paper presented at the annual conference of the International Military Testing Association, San Antonio, Texas.

[12] Cortina, J.M. (2003). Apples and oranges (and pears, oh my!): The search for moderators in meta-analysis. *Organizational Research Methods, 6*, 415-439.

[13] Costa, P.T., & McCrae, R.R. (1994). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI): Professional Manual*. Florida: Psychological Assessment Resources, Inc.

[14] Craik, K.H., Ware, A.P., Kamp, J., O' Reilly, C.O. III, Staw, B., & Zedeck, S. (2002). Explorations of construct validity in a combined managerial and personality assessment programme. *Journal of Occupational and Organizational Psychology, 75*, 171-193.

[15] Darr, W.A. (2009). *The Trait Self Descriptive (TSD) Inventory: A facet-level examination*. Technical Memorandum 2009-010. Directorate of Military Personnel Research and Analysis, Ottawa, Ontario, Canada.

[16] Doll, R.E. & Gunderson, E.K.E. (1969). Occupational group as a moderator of the job satisfaction-job performance relationship. *Journal of Applied psychology, 53,* 359-361.

[17] *Elshaw, C. & Abram, M. (1999). Predicting early voluntary wastage in Army officer training. *Journal of Defence Science, 4(1),* 2-8.

[18] French, E.G. & Ernest, R.R. (1955). The relation between authoritarianism and acceptance of military ideology. *Journal of Personality, 24*, 181-191.

[19] Gade, P.A., Lakhani, H., & Kimmel, M. (1991). Military Service: A good place to start? *Military Psychology, 3,* 251-267.

[20] Hough, L.M. & Oswald, F.L. (2008). Personality testing and industrial-organizational psychology: Reflections, progress, and prospects. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1,* 272-290.

[21] *Grandmaison, L.J. (2006). *Assessing the incremental validity of personality on direct leadership in the Canadian Forces.* Unpublished doctoral dissertation. Carleton University, Ottawa, Ontario.

[22] Graziano, W.G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.). *Handbook of Personality Psychology*. New York: Academic Press.

[23] *Hampson, A., Vincent, M., & Jacobs, N. (1997). *The use of the Trait Self-Description Inventory to predict training wastage in Phase I training*. DERA/CHS/HS3/CR97262/2.0. Farnborough, Hampshire: Defence Evaluation and Research Agency.

[24] Holloman, C.R. (1967). The perceived leadership role of military and civilian supervisors in a military setting. *Personnel Psychology,* 20, 199-210.

[25] Hunter, J.E., & Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

[26] Hunter, J.E., & Schmidt, F.L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.). Thousand Oaks, California: Sage.

[27] *Jones, D., Uggerslev, K., Paquet, S. Kline, T., & Sulsky, L. (2000a). *Validation of the Trait Self Descriptive Inventory factors using basic and advanced MOC course results and self-reported job performance*. Contractor's Report 00-08. Director of Human Resource Research and Evaluation, Ottawa, Ontario, Canada.

[28] Krueger, G.P. (2001). Military psychology: United States. *International Encyclopedia of the Social & Behavioral Sciences* (pp. 9868-9873). Oxford, England: Elsevier Ltd/Pergamon Press.

[29] Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*, 162-181.

[30] Kurpius, S.E. & Lucart, A.L. (2000). Military and civilian undergraduates: Attitudes toward women, masculinity, and authoritarianism. *Sex Roles, 43,* 255-265.

[31] Lee, K., & Ashton, M.C. (2004). The HEXACO Personality Inventory: A new measure of the major dimensions of personality. *Multivariate Behavioral Research, 39*, 329-358.

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

[33] Nunally, J.C. (1978). *Psychometric Theory*. New York, New York: McGraw Hill.

[34] *O'Keefe. D.F. (1998). *Investigating the use of occupational personality measures in the Canadian Forces Selection System*. Technical Note 98-14. Directorate of Human Resources Research and Evaluation, Ottawa, Ontario, Canada.

[35] *O'Keefe, D.F. (1999). *Development of an optimal Trait Self Descriptive (TSD) inventory profile for military police applicants*. Technical Note 99-2. Directorate of Human Resources Research and Evaluation, Ottawa, Ontario, Canada.

[36] *Perkins, A., & Corr, P. J. (1997). *Ability-personality interactions predict military leadership potential.* PLSD/CHS/HS3/CR97027/1.0. Farnborough, Hampshire: Defence Evaluation and Research Agency.

[37] Pulakos, E.D., White, L.A., Oppler, S.H., Borman, W.C. (1989). Examination of race and sex effects on performance ratings. *Journal of Applied Psychology, 74,* 770-780.

[38] Quinn, A., sinha, R., Keeney, J., & Schmitt, N. (2009). *Why Conscientiousness predicts performance; mediation by effortful and ethical behaviors*. Paper presented at the annual convention of the Society for Industrial/Organizational psychology, New Orleans, Louisiana.

[39] Rothstein, H.R. (1990). Inter-rater reliability of job performance ratings: growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75,* 322-327.

[40] Sackett, P.R. & DuBois, C.L.Z. (1991). Rater-ratee race effects on performance evaluation: Challenging meta-analytic conclusions. *Journal of Applied Psychology, 76,* 873-877.

[41] Sackett, P.R. & DeVore, C.J. (2001). Counterproductive behaviors at work. In N. Anderson, D. Ones, H. Sinangil, C. Viswesvaran (Eds), *Handbook of Industrial, Work, and Organizational Psychology,* vol.1, Thousand Oaks, CA: Sage.

[42] Salagado, J.F. (2002). The Big Five personality dimensions and counterproductive behaviours. *International Journal of Selection and Assessment, 10,* 117-125.

[43] *Schmit, V. P. (1997). *An initial assessment of motivational and personality measures for RN Officer selection*. DERA/CHS/HS3/CR97027/1.0. Farnborough, Hampshire: Defence Evaluation and Research Agency.

[44] Schmidt, F.L. & Le, H.A. (2004). Hunter-Schmidt Meta-analysis Programs. Iowa City, Indianapolis: University of Iowa.

[45] *Schwartz, S. (1999). *Validation of the Trait Self Descriptive inventory against Basic Recruit Training and Basic Officer Training Course results*. Technical Note 99-15. Directorate of Human Resources Research and Evaluation, Ottawa, Ontario, Canada.

[46] *Skomorovsky, A. (2007). *The 75-item Trait Self Descriptive Personality Inventory (TSD-PI) English version: Examining Convergent and Predictive Validity in a Sample of Privates.* Technical Note under review. Directorate of Military Personnel Operational Research and Analysis, Ottawa, Ontario, Canada.

[47] *Smith, G. (2008). *RAFAAT and TSD-PI as predictors of aerospace control performance*. Paper presented at the annual convention of the International Military Testing Association, Amsterdam, Netherlands.

[48] Soloman, D.N. (1954). Civilian to soldier: Three sociological studies of infantry recruit training. *Canadian Journal of Psychology, 8,* 87-94.

[49] Syed, F. & J.D. Klammer (2002). *The Trait Self Descriptive Inventory: An examination of the Issues*. Technical Panel HUM-TP3 Report. Ottawa, Ontario, Canada: Director Human Resources Research and Evaluation.

[50] 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.

[51] Tupes, E.C. & Christal, R.E. (1961). *Recurrent Personality Factors Based on Trait Ratings* (ASD-TR-61-97). Lackland Air Force Base, Texas: Aeronautical Systems Division, Personnel Laboratory.

[52] Vineberg, R. & Joyner, J.N. (1982). *Prediction of job performance: Review of military studies*. Alexandria, VA: Human Resources research Organization.

[53] *Watkins, J.L. (1998). *Big Five personality dimensions and recruit training outcomes in the Royal Australian Navy: An exploratory study*. Unpublished master's thesis, Monash University, Australia.

[54] Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315-321.

[55] Wiggins, J.S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37*, 395-412.

Annex A Data Coding Sheet

TSD META - CODING SHEET

1. Study #: _____
2. Title: _____
3. Authors: _____
4. Year: _____

TSD Details:

5. TSD length (# items):_____
6. TSD alternate name: _____

Sample Details (if more than one sample present – complete the following for each sample):

Orgn./Country:_____	Orgn./Country:_____	Orgn./Country:_____
Sample size: _____	Sample size: _____	Sample size: _____
Recruits:___ Incumbent:____	Recruits:___ Incumbent: _____	Recruits:___ Incumbent: _____
NCM:____ Officer: ___	NCM:____ Officer: ___	NCM:____ Officer: ___
Occupation/breakdown:_____	Occupation/breakdown:_____	Occupation/breakdown:___
Avg age: __ or Avg. tenure:___	Avg. age: _____ or Avg. tenure: _	Avg age: __ or Avg. tenure: ___
% female:_____	% female:_____	% female:_____

Study Details:

13. Predictive:_____ Concurrent: _____ Postdictive: _____
14. Measurement Interval (b/n predictor and criterion):_____ (in months)
15. Research: ____ Administrative:____
16. Theoretical: _____ Empirical:_____

Criterion Details:

16. Soft (subjective): _____ Hard (objective):_____
17. Training: _____ On the job:____

Effect Size - Complete one for each criterion/sample

Sample: _____ Size: _____

Criterion: _____ N (r_{NC}): _____ E (r_{EC}): _____ O (r_{OC}): _____	Criterion: _____ N (r_{NC}): _____ E (r_{EC}): _____ O (r_{OC}): _____

A (r_{AC}): ________ C (r_{CC}): ________	A (r_{AC}): ________ C (r_{CC}): ________
Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________	Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________
Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________	Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________

Sample:_____________________ Size:_________________________

Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________	Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________
Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________	Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________
Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________	Criterion: ___________________ N (r_{NC}): ________ E (r_{EC}): ________ O (r_{OC}): ________ A (r_{AC}): ________ C (r_{CC}): ________

Glossary

Sources used for definitions:

Cooper, H. & Hedges, LV(1994). The handbook of research synthesis (pp. 531-542). New York, NY: Russell Sage Foundation.

Piedmont, R.L. (1998). *The Revised NEO Personality Inventory: Clinical and Research Applications* (pp. 84-92). New York, NY: Plenum Press.

Whitener, E.M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315-321.

artifact
Statistical and measurement imperfection that causes observed statistics to depart from population (parameter) values. Imperfections could result from unreliability in the predictor or criterion variable, range restriction due to selection, range restriction due to dichotomization of variables.

confidence interval
The interval within which the mean sample-weighted effect (i.e., corrected for sampling error) is expected to lie. Intervals that include zero are interpreted as meaning that an effect does not exist.

credibility interval
The interval within which the mean corrected effect (i.e., corrected for artifacts) is expected to lie. Large intervals or those including zero indicate the presence of several sub-populations, suggesting the operation of moderators.

effect
The association between two variables.

meta-analysis
The statistical analysis of a collection of analysis results from individual studies for the purpose of integrating findings.

moderator
Any factor that influences the size of a particular relationship and is itself not a consequence of the relationship.

Neuroticism
One of the Big Five factors of personality which is also referred to as Emotional Stability. High scores on this factor reflect a proneness to psychological distress, maladaptive coping, unrealistic ideas, and excessive cravings or urges.

Extraversion
One of the Big Five factors of personality which represents the need and intensity for interpersonal interaction, stimulation, and capacity for joy. Individuals who score high on this factor are thought to be sociable, active, and person-oriented.

Openness to Experience
One of the Big Five factors of personality, representing the need to proactively explore and appreciate the unfamiliar. Individuals who score high on this factor are thought to be curious, untraditional and creative.

Agreeableness
One of the Big Five factors of personality, representing one's attitudes towards other people. Individuals who score high on this factor are compassionate, trusting, forgiving, and like to help others.

Conscientiousness
One of the Big Five factors of personality, reflecting an individual's degree of organization, persistence, and motivation in goal-directed behaviour. Individuals who score high on this factor are thought to be dependable and competent.

Distribution list

Document No.: DGMPRA TM 2009-023

LIST PART 1: Internal Distribution by Centre

1	DG CORA
1	DRDC CORA Chief Scientist
1	DRDC CORA Library (1 copy)
2	DRDKIM Library (2 copies)
1	DRDC/DGSTO/DSTP
1	DGMPRA
1	DGMPRA – Chief Scientist
1	DGMPRA – Deputy DG
1	DGMPRA – Personnel Generation Research – Section Head
1	DGMPRA – Personnel and Family Support Research – Section Head
1	DGMPRA – Organizational and Operational Dynamics – Section Head
1	DGMPRA – Team Leaders
1	CMS/D Mar Strat 2-6
1	DRDC (Toronto)
1	CFHA
1	D Mar Pers
1	DLPM/G1
1	D Air Pers Mgt
1	VAC LO – Christiane Dubé
20	TOTAL LIST PART 1

LIST PART 2: External Distribution by DRDKIM

1	Library and Archives Canada
1	TOTAL LIST PART 2

21 TOTAL COPIES REQUIRED

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

DGMPRA
101 Colonel By Drive
Ottawa, Ontario K1A 0K2

2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED
(NON-CONTROLLED GOODS)
DMC A REVIEW:GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Big Five Personality Research in the Military: A Meta-analysis:

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Darr, W.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
November 2009	40	55

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Memorandum

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

DGMPRA
101 Colonel By Drive
Ottawa, Ontario K1A 0K2

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DGMPRA TM 2009-023	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Despite established validity of the Big Five factors of personality, validity generalization is limited by evidence for situational specificity. Unique requirements and expectations demanded of military personnel warrant the examination of personality-outcomes associations gathered in such samples. Using accumulated research into the validity of the Air Force Self-descriptive Inventory (AFSDI) across several military organizations, this examination involved a meta-analysis to determine the validity of the Big Five personality factors as measured by the AFSDI or variants of it. Effects are comparable with mainstream meta-analytic estimates, being slightly higher for some effects. Conscientiousness was confirmed to be a strong predictor of military performance, and generalized across samples. Although moderated by the type of measure (i.e., original/adapted length) and/or the type of military stream (non-commissioned members/officers), Neuroticism and Extraversion were the next best predictors of performance and officer training, respectively.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

DRDC CORA

www.drdc-rddc.gc.ca