ASVAB Validation Technical Report

Master at Arms (MA) Rating

Shaneka Parham, Ph.D.
Peraton

Janet Held and Tom Blanco
S &T Consulting and Program Management, LLC

Reviewed by
Jason Jacobs-Lentz
Navy Selection & Classification Office (OPNAV N-132G)

Approved and Released by
Donald H. Lambrix, Jr.
Director, Navy Selection & Classification Office
(OPNAV N-132G)

May 30, 2022

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
30-05-2022	Technical Report	Oct 2017 – Sept 2021

4. TITLE AND SUBTITLE ASVAB Validation Technical Report Master at Arms (MA) ASVAB Standards for Navy Enlisted Rating Entry	5a. CONTRACT NUMBER N00189-17-D-Z016
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Shaneka Parham, Ph.D. Janet D. Held Tom Blanco	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Perspecta, Chantilly, VA 20151 S & T Consulting and Program Management, Brentwood, CA 94513	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Director, Navy Selection and Classification Office OPNAV N132G 701 S. Courthouse Road Arlington, VA 22204	**10. SPONSOR/MONITOR'S ACRONYM(S)** OPNAV N132
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 22-02

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The Armed Services Vocational Aptitude Battery (ASVAB) is a joint-service test battery used by the U.S. military services for enlistment selection and occupational classification decisions. The military services validate the ASVAB periodically to ensure that the composites of ASVAB tests used to classify enlisted members to occupations are the most predictive of important training performance measures, and that cutscores are set to manage academically related setbacks and failures. This report contains an evaluation of the ASVAB standards for the Navy's Master at Arms (MA) rating. The purpose of the evaluation was to determine if a revision in the standards could improve (a) MA A-School performance to include what is called First Pass Pipeline Success (FPPS, graduating without an academic setback), (b) qualification rates, and (c) diversity. Analyses showed that all three could be improved by revision of the operational standards.

15. SUBJECT TERMS
ASVAB, selection and classification, Navy ASVAB validation/rating entry standards, validity coefficients, cutscores, diversity

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
UNCLASSIFIED					Donald H. Lambrix, Jr.
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UNLIMITED	**20**	**19b. TELEPHONE NUMBER** *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			(703) 604-5103

Contents

1. Introduction

The Navy Selection and Classification Office (OPNAV N132G) oversees the process for re-evaluating the Armed Services Vocational Aptitude Battery (ASVAB) entry standards for enlisted accession ratings and programs. Although training performance and recruiting data are reviewed every year for all ratings, full in-depth studies are conducted only for ratings that exhibit specific stressors such as high academic setback rates or difficulty filling goal. This report of an in-depth study for the Master at Arms (MA) rating was conducted to address a recently identified spike in ASVAB waiver rates in concert with fluctuating first pass pipeline success (FPPS), graduation rates, and academic setback rates.

All Navy ratings are reviewed annually by the Navy's Selection & Classification Office (N132G) for indicators of sub-optimal schoolhouse performance that then are discussed with the rating community to determine if a full ASVAB Validation/Standards study is required. A rating community can also preemptively determine a study requirement due to difficulty in filling rating goal, the need to issue Exception to Policy ASVAB waivers to fill goal, or a decline in school performance to the extent that not enough trainees graduate to fill Fleet requirements.

The primary goal of an in-depth ASVAB validation/standards study is to set a standard that balances acceptable performance in the schoolhouse with enough accessions qualified to fill the rating's annual goal. There is always a tension between the two that increases in a challenging recruiting environment. Setting "effective" ASVAB standards not only reduces the Navy costs of remediating or setting back a trainee to a subsequent class, but also minimizes the burden placed on recruiting to re-fill the class seats occupied by trainees who failed the course. Effective ASVAB standards also allows for improvement of the person-job fit, or Sailor-job fit, particularly if the operational ASVAB composite is replaced or augmented with one that measures additional constructs relevant to the rating's training, and by extension, the rating's job tasks and duties. Many Navy ratings have "alternative" ASVAB standards that include more aptitudes/abilities/knowledge relevant to the rating that increase opportunities for those with diverse backgrounds. The MA rating with slightly below "average" cutscore levels for its two "required" standards does not have alternative standards.

The purpose of the evaluation was to determine if a revision in the MA rating's ASVAB standards could improve (a) MA A-School performance measured as First Pass Pipeline Success (FPPS, graduating without an academic setback or failure), (b) qualification rates, and (c) diversity.

2. Description of the ASVAB and Special Classification Tests

The Armed Services Vocational Aptitude Battery (ASVAB) is a mix of aptitude/ability/ knowledge-based tests that are used by all the military services as their primary cognitive instrument for selecting military applicants and classifying them into enlisted occupations. The ASVAB was developed to predict training performance with the understanding that training content is informed by job analysis. New tests are now being considered as additions to the ASVAB, or as special classification tests, as technology and military jobs change.

Table 1 provides a brief description of the ASVAB tests, and at the lower portion of the table, the ASVAB special classification tests.

Table 1: The ASVAB and Special Classification Tests

	Test		Content
ASVAB	GS	General Science	Biological and physical sciences
	AR	Arithmetic Reasoning	Arithmetic word problems
	WK	Word Knowledge	Synonyms/meaning of words in context
	PC	Paragraph Comprehension	Written passages
	MK	Mathematics Knowledge	Algebra, geometry, fractions, decimals, exponents
	EI	Electronic Information	Electrical Principles and electronics
	AS	Auto and Shop Information	Automotive, tool, shop, practices
	MC	Mechanical Comprehension	Mechanical and physical principles
	AO	Assembling Objects	Patterns and connection point recognition
Special Classification	CS	Coding Speed	Perceptual speed/accuracy
	DLAB	Defense Language Aptitude Battery	Aptitude for learning languages
	NAPT	Navy Advanced Placement Test	Advanced physics, mathematics, and chemistry
	CT	Cyber Test	Knowledge in cyber field
	MCt	Mental Counters	Working memory

Each ASVAB test is scored on a scale referenced to the ASVAB normative youth population (Profile of American Youth, 1997, or PAY97) (Segall, 2004). The scores are standardized to have a mean score of 50 and standard deviation (SD) of 10. The bulk of ASVAB test scores fall in the range of 20 to 80. Scores on Word Knowledge (WK) and Paragraph Comprehension (PC) are combined to form the Verbal (VE) composite (also with mean 50 and SD 10) with PC weighted 1/3rd and WK, 2/3rds. VE, Arithmetic Reasoning (AR), and Mathematics Knowledge (MK) are part of the Armed Forces Qualification Test (AFQT, 2VE+AR+MK) used to qualify military applicants for service and also are part of many Navy classification composites (listed in Appendix A). The AFQT is scaled as a uniform percentile distribution with scores ranging from 1-99, which is a transformation from the sum of test scores. The Navy ASVAB composite scores are the sum of the scores on the individual ASVAB tests that form the composite.

The MA rating's operational standard is WK+AR ≥98 and WK ≥43. That is, Navy accessions can qualify for the MA rating by meeting the requirement of at least a 43 on WK and a combined score of at least 98 on the WK+AR composite.

3. Linkage of the ASVAB Tests to the MA Rating Job Description

Because of the multitude of Navy ASVAB composites (26 listed in Appendix A), many of which are highly correlated, the list was narrowed down for MA rating analysis by a rational linkage of the constructs measured by the ASVAB tests with the MA rating job skills and qualifications requirements and high level training content descriptions. The Navy MA rating's job description

card was used for the linkage analysis obtained from the military web site, (https://cool.osd.mil/usn/enlisted/rating_info_cards/ma.pdf).

The overall description of the Navy MA rating from the website is as follows:

"Those personnel who want to qualify for the MA rating should be people oriented, dedicated, resourceful, and versatile. They should possess strong writing skills, a good memory, and the ability to conduct detailed work while maintaining accurate records. Prospective Master at Arms should possess physical strength, manual dexterity and competence with tools, equipment, and machines."

The MA Class "A" School training is approximately 10 weeks with subject matter consisting of "Antiterrorism techniques, Armed Sentry/Post Standing techniques, crime prevention, military and civil law, physical fitness, communications, first aid, firearms deployment and physical restraint techniques".

Table 2 provides the linkage matrix with the job description components.

Table 2: ASVAB Test Linkages to the MA Rating

MA Rating Skills/Qualifications	AO	AR	AS	CS	EI	GS	MC	MK	PC	VE	WK
Strong writing skills									x	x	x
Strong speaking skills									x	x	x
A good memory				x			x	x			x
Ability to conduct detailed work	x	x	x				x	X			
Ability to maintain accurate records				x				x	x	x	x
Physical strength											
Manual dexterity	x		x				x				
Competence w/tools, equipment, and machines	x		x		x	x	x				

Table 2 shows the Mechanical Comprehension (MC) test to have the most checked boxes (4) in direct linkage to MA rating's job description followed by Assembling Objects (AO), Auto Shop (AS), Mathematics Knowledge (MK), and the verbal tests, PC, WK and VE. Coding Speed (CS) with 2 checked boxes was eliminated from consideration because other tests with more check boxes overlap in measuring "good memory" and "maintain accurate records" (only a fraction of MA rating requirements). Electronics Information (EI) and General Science (GS) with 1 checked box each were eliminated from formation of the set of "candidate" MA rating composites to be matched to the Navy list of 26 composites for subsequent validity coefficient analyses.

4. Data Collection and Descriptive Statistics

MA A-School training data were collected from the CeTARS database (Corporate Enterprise Training Activity Resources Systems) that included Person Event (PEVT) but no Final School Grade (FSG) or modules grades. The data were pulled for the MA A-School CDPs 1981 and 23FD over available Calendar Years (CY17-21) by Active Duty Start Date (ADSD). CY18-20 reflected full year training datasets (total sample size was 4,910). Each trainee was classified as to whether he or she:

- Graduated with no academic setbacks or failures
- Suffered an academic failure or setback
- Was a DOR (drop on request)

- Was a non-academic attrite, with no prior academic setback
- Did not complete the class, but did not suffer an attrite or academic setback

Trainees in the last two categories were eliminated from the analysis dataset as were cases with ASVAB test scores outside the legitimate score range of 11-89. Those with a DOR code (Drop on Request) were also eliminated because there was no indication of whether the request was academically related vs. non-academically related. Trainees were assigned a 'First Pass Pipeline Success' (FPPS) value of 1 for graduating with no academic setback or failure, and a 0 if they experienced either an academic failure or academic setback. A 'Graduation' variable was also computed with a value of 1 for graduated and 0 for not graduated.

Table 3 provides a breakout of the MA rating training data by performance metric and ASVAB waiver rate across the five CYs recognizing complete calendar year training data were only available for CY18, CY19 and CY20. The two most recent years of complete data, have performance metrics in bold for statistically significant differences. Practical significance was assessed by the test for proportion difference "Effect Sizes" with the usual reporting of small =.2, medium =.5, large =.8 (Cohen, 1988).

Table 3: MA Rating Data Descriptives

Performance Metrics	Calendar Year Statistics					
	Overall	2017	2018	2019	2020	2021
Waiver Rate	8.5%	1.6%	4.8%	**7.0%**	**12.7%**	14.5%
FPPS Rate	86.2%	90.9%	85.0%	**83.5%**	**88.7%**	85.2%
Graduation Rate	96.5%	98.7%	96.9%	97.2%	97.4%	90.9%
Aca Setback Rate	13%	9.1%	13.9%	**15.8%**	**10.6%**	14.0%
Aca Failure Rate	2.1%	1.3%	2.5%	1.8%	1.8%	2.7%
Sample Size	4,910	386	1,418	1,088	1,390	628

Notes.
1). Bold values for CY19/20 indicate statistical significance at <.001 probability level.
2). Effect sizes for proportion differences for CY19/20 were at the .15 level for Academic Setback rates and. .20 for the ASVAB waiver rate comparison.

Table 3 shows an upward trend in ASVAB waiver rates over CYs (1.6% in CY17 rising to 14.5% in CY21). In contrast, graduation rates and academic failure rates appear steady within a narrow range across all years except for CY21. CY21 without a full year's data input had a much lower FPPS rate (85.2%) and graduation rate (90.0%) than in CY20 and with a higher academic setback rate (14.0%). These "red flag" performance metrics are not a concern because FY20 is immature data with early reporting of non-graduates and academic setbacks with graduation trainees with FPPS =1 not yet reported. CY20 is key (with all data accounted for) for a comparison with the prior year, CY19. FPPS rose in CY20 consistent with a decline in academic setback rates, which is counterintuitive for the rise in ASVAB waiver rate (12.7% in CY20 vs. 7.0% in CY19). Conclusions from the Table 3 statistics are that the MA A-School is doing well in managing the recruiting environment challenges with lower ASVAB scores and (b) other

attributes besides aptitude/ability/knowledge, as measured by the ASVAB, influence success in MA A-School, potentially influencing the robustness of the ASVAB/FPPS relation.

5. Standards Assessment

MA A-School performance for the four measures (academic setback, academic failure, FPPS, and graduation) were broken out for the total MA rating sample, n = 4,910, by how MA rating trainees qualified for the rating or did not. This "decomposition" analysis involved computing the FPPS rate for four mutually exclusive groups of trainees: (a) qualified on both standard requirements, WK+AR ≥98 and WK ≥43, (b) waivered on WK but not on WK+AR, (c) the reverse, waivered on WK+AR but not on WK, and (d) waivered on both WK+AR and WK (did not qualify for either standard requirement).

Table 4 provides the operational MA rating standard's decomposition results.

Table 4: Decomposition of MA A-School Data

MA A-School Performance from How Trainees Qualified Sample Size =4,910 (multiyear data)						
	Mean WKAR Score	Mean WK Score	Academic Setback	Academic Failure	FPPS	Grad
Met both WK+AR and WK (n = 4,493)*	**105.8**	**53.0**	**12.6%**	**1.9%**	**86.6%**	**96.8%**
Waivered on WK (n=0)	n/a	n/a	n/a	n/a	n/a	n/a
Waivered on WK+AR (n=416)**	**94.8**	**48.7**	**17.8%**	**4.1%**	**81.5%**	**92.5%**
Waivered on both WK+AR and WK (n=1)	90.0	40.0	0.0%	0.0%	100%	100%
Total=4,910 Trainees	104.8	52.6	13.0%	2.1%	86.2%	96.5%

*Highest performance was observed for those who met both of the MA standards.
**Lowest performance was observed for those who met only the WK requirement but not the WK+AR requirement.

Notes.
1). T-tests showed significant differences in WK+AR and WK mean scores for qualified and WK+AR trainees at the <.001 probability level. Effect Sizes (standardized mean test score differences between groups) were 1.49 and .88, respectively, considered large.
2). Chi-square tests showed significant differences in Academic Setback, Academic Failure, FPPS, and Grad rates between qualified and WK+AR waivered trainees at the <.001 probability level. Effect Sizes for differences in proportions were at the .15 level for the FPPS rate and Academic Setback rate; .20 level for Graduation rate, and .50 level for Academic Failure rate.

Table 4 shows, for each mutually exclusive way that MA A-School trainees qualified for the MA rating, how they performed across the four-performance metrics of interest (Academic Setback rate, Academic Failure rate, FPPS rate, and Graduation rate). For those waivered on WK+AR, the rates show that trainee performance was subpar across the four-performance metrics compared to those who met both WK+AR and WK requirements. Note that there were not

enough trainees who were waivered on both WK+AR and WK (n =1) or waivered on only WK (n =0) for there to be a comparison among all 4 categories of the decomposition.

6. ASVAB Validity Coefficients

The objective of a criterion-related predictive validity analysis using ASVAB composites as the predictor is to determine which ASVAB composites are most predictive of scores on the performance measure of interest. Another objective is to determine how the test composition of the composite affects diversity. Final school grade (FSG) up until the variable's unavailable in CeTARS (approximately 2018) was included as the primary criterion (performance) variable, with FPPS now taking its place. FPPS is a useful broader based performance metric than FSG in that it captures costly academic setback as well as failure incidences. However, the continuous FSG measure underlies both FPPS and graduation status and maintaining this assumption allows for validity coefficient analyses with a slight modification.

Predictive validity in this report refers to the robustness of the relationship between scores on an ASVAB composite (or a single test) with scores on the FPPS. There are two ways to measure this robustness, by the validity coefficient magnitude and by prediction of FPPS via a logistic regression equation. Validity coefficients, or correlation coefficients, range from -1 to +1 with a validity coefficient of zero (0) indicating no predictive power at all (i.e., course outcomes are completely unrelated to ASVAB scores or the measure of performance is poorly designed). A validity coefficient of +1 or -1 indicates perfect prediction, which never occurs with the ASVAB because psychological tests (a) always have some measurement error associated with them, (b) do not cover the total performance construct domain, and (c) do not take into account other factors that influence performance.

The average validity coefficient magnitude across Navy ratings is about .55. The smallest validity coefficient is about .25 for the physically focused Special Operations Basic Underwater Demolition/SEAL (BUD/S) course. Validity coefficients as high as .85 have been observed for highly technical ratings like Nuclear Field. The larger the validity coefficient the stronger the relation between the ASVAB composite scores (the predictor variable) and FPPS outcome (the criterion variable) and the more impact adjustments to the cutscore will impact FPPS levels. The methods for estimating ASVAB composite (and single test) validity coefficients involves correcting the correlations observed in the training data for the effects of reduced score variance that occurs from use of an operational ASVAB standard in rating qualification (composite with cutscore). Score variance directly influences the magnitude of a correlation so that without the correction, the operational composite's validity coefficient will be smaller than those of other composites. Further, a second step correction is made to address the dichotomization of the performance variable into a 1/0 variable that has underlying it a continuous criterion variable such as FSG.

The effects of test score variance range restriction on validity coefficient analysis is described in Appendix B Figure B with reference to the multivariate correction for range restriction applied in this study. The second and last correction obtains the biserial correlation from the point-biserial correlation (Cohen & Cohen, 1983).

The validity coefficients for the operational and candidate set of composites were calculated using the following steps:

- Calculate the observed restricted in range correlation (validity coefficient) between composite score and FPPS in the school sample.
- Correct this validity coefficient for range restriction of test scores (that occurs from assigning accessions to a rating based on an ASVAB standard).
- Use the multivariate range correction method with the ASVAB full score range reference population (PAY97) serving as the validity estimation population to obtain fully range-corrected coefficients.
- Apply the correction for criterion variable dichotomization (DCC) to the range-corrected validity coefficients.

Table 5 lists the validity coefficients for the MA rating's operational (current) composite and the composites rationally formed through ASVAB linkage analysis (Table 2) matched to the existing Navy set of operational composites. Only tests with the most linkages to the MA rating job and training descriptions were considered with a heavy emphasis on verbal ability. Also included are the tests in those composites. FPPS served as the performance measure and the ASVAB full score range reference population served as the validity estimation population (Profile of American Youth, PAY97, Segall, 2004).

Table 5: MA Rating Validity Coefficients Corrected for Score Range Restriction and Criterion Variable Dichotomization Correction (DCC)

Composite/Single Test	Validity Coefficients	
	PAY97 Correction	PAY97+DCC
VE+MK+AS	0.33	0.52
VE+AR+MK+AS	0.33	0.52
VE+AR+MK+MC (Candidate)	0.33	**0.52**
PC+AR+MK	0.31	0.49
VE+MK	0.30	0.47
VE+AR+MK+AO	0.30	0.46
WK+AR (Current)	0.29	**0.46**
MC	0.31	0.48
PC	0.30	0.47
AR	0.28	0.44
VE	0.28	0.44
MK	0.28	0.43
WK	0.25	0.39
AS	0.24	0.37
AO	0.20	0.32

Note. Corrections were applied to observed validity coefficients from the MA A-School data, n=4,910. No statistical tests were performed to establish confidence intervals.

Table 5 shows the two sets of 1st and 2nd stage corrected validity coefficients for the set of candidate composites and their test derived from the rationale linkage process. Both sets of validity correction stage values are provided to demonstrate the effects of dichotomizing the criterion variable into a binary variable– the degree of performance information loss provided by the underlying continuous performance variable, school grades. The MA rating's operational composite, WK+AR, with a .46 validity coefficient is the least predictive of all of the composites and thus warrants exploration of a replacement composite with higher FPPS predictive validity.

Three composite had the largest validity coefficients at .52 (about average for the Navy ratings), VE+MK+AS, VE+AR+MK+AS, and VE+AR+MK+MC. However, the Mechanical Comprehension (MC) test with a .46 validity, the largest of all single tests consistent with a broader coverage than other tests of the MA rating requirements, is less specific to a single technical area (i.e., shop and auto mechanics) as provided by the AS test with a smaller .37 validity coefficient. The VE+AR+MK+MC composite provided an increase of .06 in predictive validity (.52 compared to .46 for WK+AR). Given the same aptitude/ability level in cutscore, a .06 level of increased predictive validity results in about a 1-to-2-percentage expected increase in FPPS (currently at 86.2%), explained in Appendix C.

The VE+AR+MK+MC composite provides a balance of the AFQT tests, with the verbal component (VE) having strong rational linkages to the MA rating's job and training constructs, and mechanical aptitude/ability. The compensatory feature of the composite allows high scores on some tests to compensate for lower scores on other tests thus broadening the qualification coverage for accessions with relevant and compensating strengths.

Although the VE+AR+MK+MC composite is recommended, it is also instructive to assess the two other Navy composites not listed in Table 5 with the same validity coefficient magnitude as potential replacements for the MA rating's operational "current" WK+AR composite. For example, GS+MC+EI+AR and VE+MK+MC+GS had validity coefficients of .53 and .52, respectively. However, GS with academic science (biological and physical) test content and EI, a test of specific electronics and electrical technical knowledge, are more appropriate for the highly technical ratings that have clear linkages to their job and training requirements (e.g., Hospital Corpsman rating for GS, and the Nuclear Field ratings for EI. Accessions with high scores on the GS and EI tests are in short supply (compared to the MC test) in the current challenging recruiting environment and so these tests should be reserved for ratings that have those tests in their operational ASVAB standards. The VE+AR+MK+MC composite is most suitable for the MA rating considering predictive validity, rational linkages to the job and training descriptions, and avoiding overuse of tests more appropriate for other ratings.

7. ASVAB Predicted FPPS in School Data and Observed from Cutscore Analysis

The magnitude of the WK+AR and VE+AR+MK+MC validity coefficients was one of two analytical methods applied in the study to gauge the power of these composites to predict FPPS outcomes. The second method involved logistic regression equations (one for each composite) developed in the MA A-School training data to predict FPPS outcomes. These equations were then be applied to Navy accession and applicant data that have a fuller ASVAB score range than observed in the training data.

Graphs of ASVAB composite scores vs. predicted FPPS is useful for observing the slope of the FPPS prediction curve, a steep slope indicating a strong relation vs. a flat slope indicating no relation. Also, for observing if an upward adjustment in the ASVAB cutscore achieves a meaningful increase in an observed low FPPS rate, or if lowering the cutscore (e.g., ASVAB Exception to Policy waiver points) risks trainees having a low probability of success.

Figure 1 is a plot of predicted FPPS in the MA training data "at" each composite score developed from logistic regression equations using the WK+AR and VE+AR+MK+MC composites in separate prediction equations (shown in Appendix D). The equations are applied to Navy applicants and accessions in a later analysis.

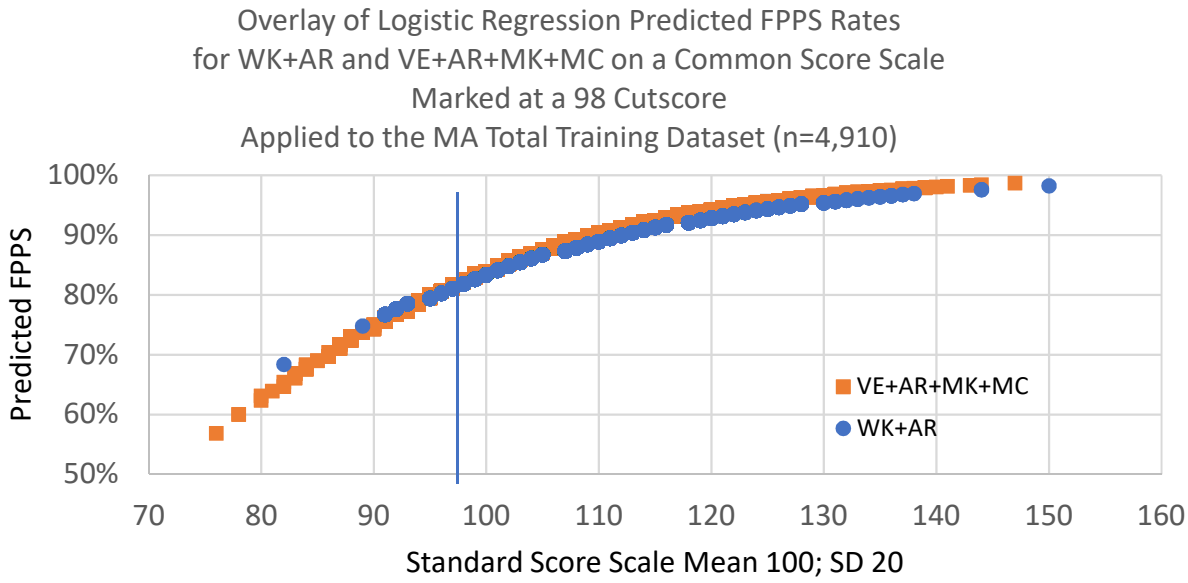


Figure 1: Predicted FPPS by standardized composite scores applying logistic regression equations developed and applied in the MA A-School data.

In order for the two prediction equations to be compared in the same Figure 1 graph, scores on the two composites were linearly transformed to a common score scale with mean score of 100 and standard deviation (SD) of 20.[1] (It is coincidence that the 98 cutscore for WK+AR on the Navy sum of test score scale is same as the standardized common score scale.) The graph shows a clear positive relation between WK+AR scores and predicted FPPS; however, with a slightly stronger relation between VE+AR+MK+MC and FPPS (indicated by a slightly elevated FPPS prediction curve above the 98 standard score cutpoint). The elevation in the curve is consistent with the VE+AR+MK+MC composite's larger validity coefficient (.52 for VE+AR+MK+MC vs. .48 for WK+AR). Scores below the 98 cutscore are not of interest because the goal is not to lower the MA rating's cutscore level, but to improve FPPS rates.

[1] The Army and Marine Corps both use the standardized score scale mean of 100, standard deviation of 20 for their ASVAB composites. The linear transformation of the sum of test scores to the new score scale requires knowing the mean and standard deviation in the PAY97 ASVAB reference population. For WK+AR those values are 100 and 18.3 whereas for VE+AR+MK+MC, 200 and 34.7.

Whereas Figure 1 depicts logistic regression predicted FPPS "at" specific cutscores in the MA rating training data, Table 7 displays an "at and above" cutscore analysis. Cutscore analysis applied to the training data can only be conducted for the operational composite(s) because any other "candidate" composite would capitalize on the aptitude/ability floor resulting from the operational ASVAB standard. (Logistic regression equations displayed in Appendix D have been formulated to mitigate the "floor" biasing effect against the operational composite.)

Table 7: Cutscore Analysis applied to the MA Rating's Training Data (N=4,910) and FY19-20 Navy Accession Qualification Rates (N=48,601)

WK+AR Score	FPPS Rate Without WK ≥43 requirement	FPPS Rate With WK ≥43 requirement	FY19-20 Accession Qual Rate Without WK ≥43 requirement	FY19-20 Accession Qual Rate With WK ≥43 requirement
110	91.8%	91.8%	34.5%	34.5%
109	91.5%	91.5%	37.2%	37.2%
108	90.5%	90.5%	40.0%	40.0%
107	90.4%	90.4%	42.8%	42.7%
106	90.1%	90.1%	45.8%	45.6%
105	90.0%	90.0%	48.7%	48.5%
104	89.7%	89.7%	51.8%	51.5%
103	89.3%	89.3%	55.0%	54.7%
102	88.6%	88.6%	58.1%	57.8%
101	88.1%	88.1%	61.3%	60.8%
100	87.8%	87.8%	64.6%	63.9%
99	87.4%	87.4%	67.7%	66.9%
98	**86.6%**	**86.6%**	**71.0%**	**69.9%**
97	86.7%	86.7%	74.2%	72.8%
96	86.6%	86.6%	77.3%	75.5%
95	86.5%	86.5%	80.2%	78.1%
94	86.4%	86.4%	83.0%	80.4%
93	86.4%	86.4%	85.8%	82.6%
92	86.2%	86.2%	88.4%	84.6%
91	86.2%	86.2%	90.7%	86.3%
90	86.2%	86.2%	92.7%	87.7%
88	86.2%	86.2%	94.6%	88.9%
84	86.2%	86.2%	97.3%	90.5%

Table 7's highlighted row shows the MA rating's operational 98 cutscore for the WK+AR composite with an 86.6% observed FPPS rate (for those who scored at and above 98). The 86.6% FPPS rate is the same with and without the WK ≥43 requirement. The WK standard does not differentiate observed FPPS either above or below the 98 cutscore. Given elimination of the WK standard results in slight increases in the FY19-20 qualification rates at and above the 98 score (larger increase below), the WK ≥43 standard is considered ineffective and should be eliminated.

Conclusions from the rational linkage, predictive validity, and cutscore analyses to this point are that (a) the VE+AR+MK+MC composite has a stronger linkage to the MA rating's training and job duties than the operational composite, WK+AR, (b) VE+AR+MK+MC also has a larger validity coefficient and higher validity in predicting MA A-School FPPS outcomes than the operational composite, WK+AR, and (c) the WK standalone requirement in addition to WK+AR does not improve FPPS with some indication it limits qualification rates.

8. Standards Evaluation: Qualification Rates, Diversity, and Predicted FPPS

Table 8 lists predicted FPPS from the logistic regression equations developed for the Figure 1 graph, for three standards: the current MA rating standard and the VE+AR+MK+MC replacement composite with two cutscore options. The 196 cutscore for VE+AR+MK+MC was determined equivalent to 98 on WK+AR (described for Figure 1). The 200 cutscore is simply a 1-score point increase per test ($1/10^{th}$ of a SD for each test). Included in the table are FY19 accession qualification rates, QR ratios for diversity analysis (females and the two minority groups, African Americans and Hispanics, most impacted by ASVAB mean score differences.

Table 8: Three Standards and How they Performed for the MA Rating

Group Assessment	WK+AR ≥98 and WK ≥43	VE+AR+MK+MC ≥196	VE+AR+MK+MC ≥200
	Qualification		
	[Current]	**[Option 1]**	**[Option 2]**
FY19 Accessions	72.4%	81.2%	74.5%
		+2,647	+620
Qualified F::M	0.81	0.85	0.79
Qualified AA::W	0.67	0.71	0.63
Qualified H::W	0.81	0.88	0.83
	Predicted FPPS		
CY18 Applicants	88.9%	90.5%	91.2%
FY19 Accessions Rating Assigned	81.9%	83.0%	83.6%

Note. Predicted FPPS rates applied to the total CY18 Navy applicant population (N=54,117) and FY19 Navy accessions (N=30,217) subjected to simulated rating assignments.

Table 8 shows that Option 1 provides the largest FY19 accession qualification rate of the three standards (81.2% QR for Option 1 vs. 72.4% for Current). Option 1 also improves diversity for each diversity category over the current standard meaning a greater proportion of the gender and minor groups are qualified within their respective total group, when compared to that proportion in the major group. The ideal is a 1:1 ratio. Option 1 also shows a gain over the current standard in predicted FPPS for the CY18 Navy applicant data (90.5% vs. 88.9%) and also for FY19 Navy accession data (83.0% vs. 81.9%) subjected to rating assignment simulation. Simulated rating assignment methods introduce rating competition for the limited number of a population's high ASVAB scorers, resulting in the expected lower predicted FPPS rates compared to when a total population is available for assignment to one rating.

Option 2 in Table 8 also improves the MA rating qualification rate when compared to the current standard, but not as large of an increase QR as observed for Option 1. Further, Option 2 only improved one of the diversity categories (Hispanic) from that of the current standard, compared to Option 1 which improved all three. Option 2 improves the predicted FPPS from the current MA rating standard for the CY18 Navy applicant data (91.2% vs. 88.9%) and also for the FY19 Navy accession data (83.6% vs. 81.9%). However, the predicted FPPS rate gains for Option 2 over those for Option 1 were not considered large enough to offset the lower MA rating qualification rate (74.5% for Option 2 vs. 81.2% for Option 1).

Of the three MA rating standard options in Table 8 (one current and two candidates), Option 1 came in first on the qualification rate and diversity metrics, and when compared to the current standard, also on the predicted FPPS metric.

9. Summary and Conclusions

1). The VE+AR+MK+MC composite has a stronger linkage to the MA rating's training and job duties than the operational composite, with the multidimensional MC test (Mechanical Comprehension) having the strongest linkage.

2). The ASVAB has substantial validity in predicting FPPS with the VE+AR+MK+MC composite having a .06 larger validity coefficient than that of WK+AR (.52 compared to .48) that should provide a 1-2-percentage point increase in the FPPS rate, all other things like recruiting and training remaining the same in the future.

3). Predicted FPPS from logistic regression was consistent with validity coefficient analysis results in that the scores on the VE+AR+MK+MC composite have a stronger relation to FPPS than scores on the WK+AR composite.

4). The WK ≥43 standalone requirement in addition to WK+AR ≥98 is not beneficial to observed FPPS rates, and below the 98 cutscores reduces the accession qualification rate.

5). The qualification rates and diversity analyses show improvements with the VE+AR+MK+MC ≥196 standard when compared to the current WK+AR ≥98 and WK ≥43 standard with the 196 and 98 cutscores comparable in aptitude/ability level tie to the standard normal curve.

6). Expectations are that replacing the current MA rating's current ASVAB standard with the VE+AR+MK+MC ≥196 standard will improve both the accession qualification rate and FPPS with the caveat that all things affecting recruiting and training in the future remain the same.

10. Recommendations

The following recommendations were presented to the ASVAB Validation Review Committee on 08 December 2021 and the standard was implemented as operational on 03 January 2022.

1). Replace the operational alternative standard (WK+AR ≥98 "and" WK ≥43) with VE+AR+MK+MC ≥196.

2). Monitor the effectiveness of the standards and provide an assessment in a retrospective study after a year of MA A-School data become available.

These revised MA A-School standards are based upon performance measurements that were logged as binary 1/0 outcomes (FPPS). As such, the performance measures do not identify specific points in the training where the largest number of trainees had insurmountable challenges and thus might benefit from even more stringent ASVAB standards than recommended in this report. The next study for the MA rating could include module level performance scores and final school grade to identify such problematic training areas and if an upward cutscore adjustment would improve FPPS rates more than what is expected from adopting the recommended standards.

11. References

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (revised edition). Hillsdale, NJ: Erlbaum

Cohen, J., & Cohen, P. (1983). Applied multivariate regression/correlation analysis for the behavioral sciences (2nd edition). Hillsdale, NJ: Erlbaum.

Held, J. & Foley, P. (1994). Explanations for accuracy of the general multivariate formulas for correcting for range restriction. Applied Psychological Measurement, 18, 355-367

Lawley, D. (1943). A note on Karl Pearson's selection formula. Royal Society of Edinburgh, Proceedings, Section A, 62, 28-30.

Segall, D. (2004). Development and evaluation of the 1997 ASVAB score scale (Technical Report No. 2004-002). Seaside, CA: Defense Manpower Data Center.

Taylor, H. C. & Russell, J. T. (1939). The Relationship of Validity Coefficients to the Practical Effectiveness of Tests: Discussion and Tables. Journal of Applied Psychology, 23, pp. 565-578.

Appendices

Appendix A: The Navy's ASVAB and Special Test Classification Composites

Table A: Navy ASVAB and Special Test Classification Composites

Composites	Composite Name
VE+MK	Administration 1
VE+MK+CS	Administration 2
PC+MK	Administration 2
VE+AR	Specialized 1
VE+MK+GS	Specialized 2
WK+AR	Specialized 3
VE+AR+MK+GS	Specialized 4
AR+MK+GS	Specialized 5
2VE+MK+GS	Specialized 6
VE+AR+MK+MC+NAPT	Specialized 7
AR+MK+EI+GS+NAPT	Specialized 8
AR+MC+AS	Mechanical 1
MK+AS+AO	Mechanical 2
AR+MK+AS	Mechanical 3
VE+MK+AS	Mechanical 4
VE+AR+MK+AS	Operations 1
AR+2MK+GS	Operations 2
VE+AR+MK+AO	Operations 3
VE+MK+MC+CS	Operations 4
VE+AR+MK+EI	Operations 5
VE+MK+MC+GS	Operations 6
VE+MK+EI	Operations 7
VE+AR+MK+MC	Technical 1
AR+MK+EI+GS	Technical 2
GS+MC+EI	Technical 3
GS+MC+EI+AR	Technical 4
CT+MK+EI	Cyber 1
CT+MK+AR	Cyber 2

Appendix B: The Restriction in Range Effects on Validity Coefficients

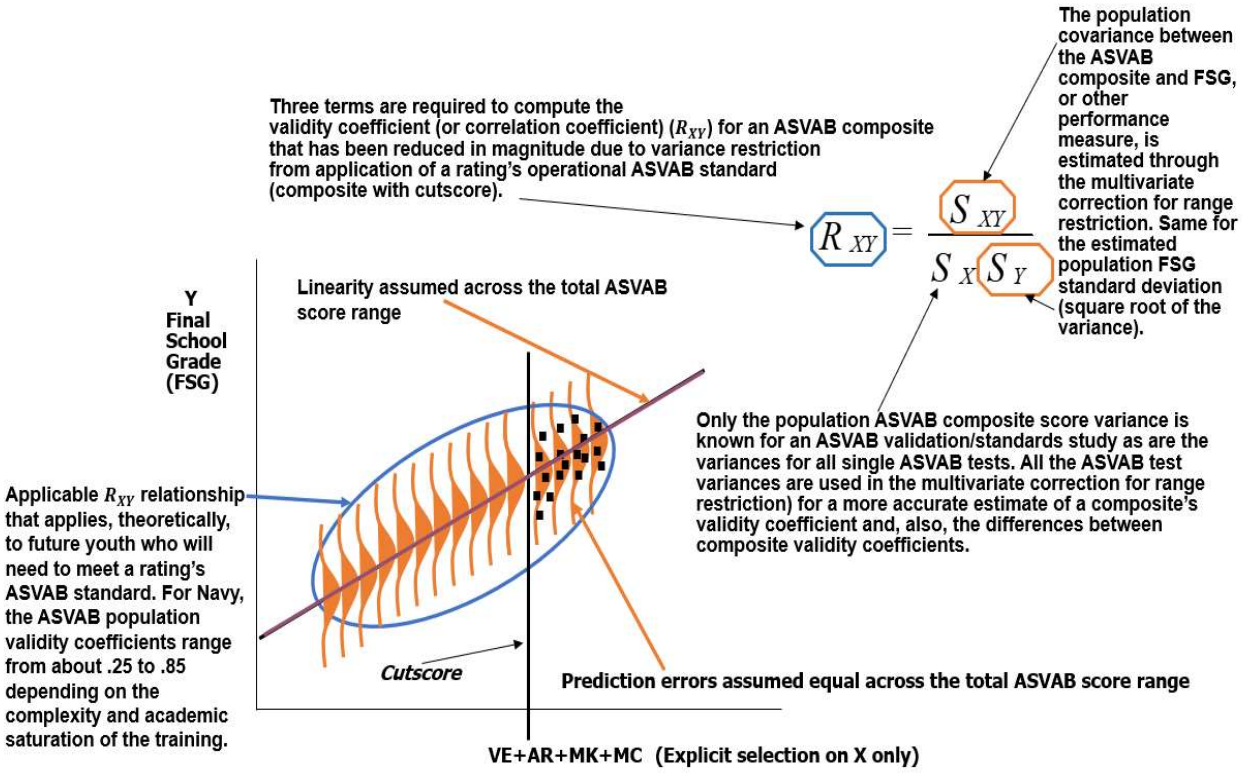


Figure B: Depiction of the validity coefficient restriction in range problem.

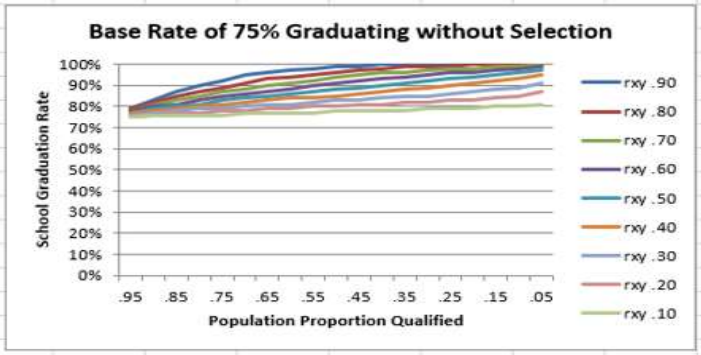
Figure B shows the restriction in range problem for the bivariate normal case of one predictor and one criterion variable. Matrix algebra computations are involved when there are more test variables available to provide potentially more accurate validity coefficient results. The multivariate correction for range restriction (Lawley, 1943) was conducted using the ASVAB test correlation matrix for the reference population, Profile of American Youth, 1997 (PAY97), as the full ASVAB score range population. Factors that influence the accuracy of the corrections are explained in Held & Foley (1994).

Appendix C: Expected FPPS Rates from ASVAB Composite Replacement and Cutscore Adjustments

R_{XY} is the Validity Coefficient

Selection Ratio (Qualification Rate),

	rxy .90	rxy .80	rxy .70	rxy .60	rxy .50	rxy .40	rxy .30	rxy .20	rxy .10
.95	79%	79%	78%	78%	77%	77%	76%	75%	75%
.90	83%	82%	81%	80%	79%	78%	77%	76%	76%
.85	87%	85%	83%	81%	80%	79%	78%	77%	76%
.80	90%	87%	85%	83%	81%	80%	79%	77%	76%
.75	92%	89%	87%	85%	83%	81%	79%	78%	76%
.70	95%	91%	88%	86%	84%	82%	80%	78%	77%
.65	96%	93%	90%	87%	85%	83%	81%	79%	77%
.60	97%	94%	91%	88%	86%	84%	81%	79%	77%
.55	98%	95%	92%	90%	87%	84%	82%	80%	77%
.50	99%	96%	94%	91%	88%	85%	83%	80%	78%
.45	99%	97%	95%	92%	89%	86%	83%	81%	78%
.40	100%	98%	96%	93%	90%	87%	84%	81%	78%
.35	100%	99%	96%	94%	91%	88%	85%	82%	78%
.30	100%	99%	97%	95%	92%	89%	85%	82%	79%
.25	100%	99%	98%	96%	93%	90%	86%	83%	79%
.20	100%	100%	98%	96%	94%	91%	87%	83%	79%
.15	100%	100%	99%	97%	95%	92%	88%	84%	80%
.10	100%	100%	99%	98%	96%	93%	89%	85%	80%
.05	100%	100%	100%	99%	97%	95%	91%	87%	81%



Data taken from Taylor, H. C. & Russell, J. T. (1939). The Relationship of Validity Coefficients to the Practical Effectiveness of Tests: Discussion and Tables. *Journal of Applied Psychology, 23,* pp. 565-578.

The Taylor-Russell (T-R) .75 Base Rate table most nearly matches the MA rating's study parameters: (a) .45 validity for WK+AR, the operational selector, (b) 87% FPPS, (c) 55% Qual Rate (Selection Ratio tied to cutscore level for a standard normal distribution of scores. A cutscore of 98 on WK+AR is equivalent to a cutscore of 196 on VE+AR+MK+MC with both transformed to a standard Z-Score of .115 (with the WK+AR PAY97 full ASVAB score range SD =18.3 for WK+AR, and 34.7 for VE+AR+MK+MC).

Highlighted in yellow - given the validity coefficient for VE+AR+MK+MC is .05 higher than that for WK+AR (roughly .50 vs. .45, respectively), there is about a 1.5% increase in the FPPS rate (interpolating ½ of the 3% expected increase going from the existing 87% FPPS rate to 90% in the same .55 Selection Ratio row). Alternatively, keeping WK+AR ≥196 standard and applying a more stringent cutscore that qualifies only 50% of the population instead of 55% shows an expected FPPS increase of 1% (88%).

Highlighted in green is a 91% expected FPPS if, in addition to adopting VE+AR+MK+MC composite with a larger validity coefficient, the 196 cutscore that equates to 98 on WK+AR is increased to 200. The expected increase in FPPS is 4-percentage points (original 87% increasing to 91%).

Highlighted in red is the expected 85% FPPS rate if the ASVAB validity coefficient drops to .40.

Appendix D: MA Rating Logistic Regression Equations

Below are the two logistic regression equations developed for the MA rating's operational WK+AR and the VE+AR+MK+MC replacement composites. Some data from the total sample, n =4,910, were eliminated for each model's equation development: below the 98 cutscore for the operational WK+AR and below the 196 equivalent score for VE+AR+MK+MC. The elimination of data below each of these cutscores was intended to mitigate equation parameter influencers, specifically (a) ASVAB waivers for which additional screening factors were not available (for model specification) and (b) screening, or selection effects imposed by an ASVAB standard that reduces the operational composite's score variance (ASVAB) and covariance (ASVAB with FPPS) resulting in a steeper sloping curve for that operational composite relative an elevated and flatter curve for the candidate replacement composite. The concern about these influencers is at the margins of the operational cutscore region where most cutscore adjustments are made.

WK+AR Model developed for 4,493 MA A-School trainees meeting the operational standard.

```
Logistic regression                             Number of obs     =      4,493
                                                LR chi2(1)        =      54.26
                                                Prob > chi2       =     0.0000
Log likelihood = -1738.9025                     Pseudo R2         =     0.0154
```

FPPS	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wkar	.0525512	.0076542	6.87	0.000	.0375492	.0675532
_cons	-3.642105	.798118	-4.56	0.000	-5.206388	-2.077823

VE+AR+MK+MC Model developed for 4,155 MA A-School trainees meeting the 196 cutscore.

```
Logistic regression                             Number of obs     =      4,155
                                                LR chi2(1)        =      62.77
                                                Prob > chi2       =     0.0000
Log likelihood = -1457.7625                     Pseudo R2         =     0.0211
```

FPPS	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
vearmkmc	.0328416	.0044816	7.33	0.000	.0240579	.0416253
_cons	-4.912336	.9408765	-5.22	0.000	-6.75642	-3.068252

Figure D: Logistic regression equations for composites predicting FPPS for the MA rating.