Enlistment Screening Test to Predict Army Aptitude Composite Scores

Richard A. Kass, Mary Weltin, Leonard Seeley and Hilda Wing

U.S. Army Research Institute
for the Behavioral and Social Sciences

All applicants for military service are required to achieve a minimum score on the Armed Forces Qualification Test (AFQT) to be eligible for enlistment. The AFQT is a composite score derived from four of the subtests of the Armed Services Vocational Aptitude Battery (ASVAB). Applicants for the Army also have to achieve minimum scores on Army Aptitude Area Composites which are various combinations of the ASVAB subtests.

The traditional Enlistment Screening Test (EST) used by recruiters for all military services to screen service applicants for potential failure has been designed to predict the AFQT score (Bayroff, Thomas & Kehr, 1959; Jensen & Valentine, 1976; Mathews, 1981; Morton & Houston, 1957). The U.S. Army Research Institute (ARI) has developed a Pre-enlistment Recruiting Test (PERT) to predict ASVAB Army Aptitude Area Composite scores, as well as scores on the AFQT. The PERT mini-battery is a shorter but parallel version of the operational ASVAB. PERT scores will be used within the Army's new Joint Optical Information Network (JOIN) system, a system using mini-computers with video display capabilities for use in Army recruiting stations. PERT-derived Aptitude Area scores will allow a recruiter to discuss particular Military Occupational Specialties (MOS) for which an applicant might be most qualified.

This paper presents preliminary results of the criterion-related validation of the PERT against the operational ASVAB.

Method

Instrument

The PERT was developed during the fall of 1980. Table 1 presents the correspondence between the subtests of the PERT and the operational ASVAB 8/9/10. The 80 items comprising the eight non-speeded PERT subtests were collected as follows. Forty items were from ASVAB Form 2 and four items from ASVAB Form 1. Twenty six items were original and ten Paragraph Comprehension items were obtained from a cross-section of three experimental Enlistment Screening Test (EST) booklets (Mathews, 1981). All of the 72 Coding Speed (CS) items were original. None of the PERT items is in the operational ASVAB.

There are three differences between the PERT and ASVAB. First, the PERT has no equivalent subtest for Numerical Operations (NO). The Coding Speed (CS) subtest in PERT substituted for both the CS and NO in the operational ASVAB. While the correlation between CS and NO in the operational ASVAB is only .64, the patterns of correlations between these two speeded subtests with the non-speeded ASVAB subtests are quite similar (see Table K-3 in Sims and Truss, 1980). Second, the PERT subtests have fewer items than the ASVAB subtests. Third, the items for the PERT subtests are not presented as separate content areas nor are they independently timed. The one exception is the PERT subtest for Coding Speed which is presented in a separate booklet (Test Book II) and has a five minute time limit. The other PERT subtest items are sequentially presented in Test Book I such that three items from a subtest are followed by three items from the next subtest. This sequencing continues until nine items from each of the subtests are presented. The last eight items of the 80 items in Test Book I are the successive presentations of the 10th items in each of the eight PERT subtests. The major advantage is that the recruiter does not need to monitor time limits nor give specific directions for each subtest. The examinee is provided with initial directions and example problems and allowed to complete Test Book I (all subtests except CS) in 50 minutes. In the event a slower examinee does not finish, each subtest should be affected about equally. The examinee is then provided with Test Book II, and allowed five minutes to answer the 72 CS items.

Procedure

In March of 1981, the U.S. Army Recruiting Command (USAREC) tasked each of the 57 District Recruiting Commands (DRC's) to select 10 of the recruiting stations in their districts to participate in the research project. This would yield a total of 570 recruiting stations. They were selected in the following manner. First, all recruiting stations which had station commanders who were not on production were identified. (Non-production station commanders are not directly involved in soliciting applicants). Second, within each DRC, up to 10 recruiting stations were randomly selected. The DRC's and their designated recruiting stations participated in the study until each station commander tested 15 applicants and forwarded to the DRC 15 complete and usable answer sheets. Each DRC then forwarded all answer sheets to ARI.

Table 1

Correspondence Between the PERT and

ASVAB Subtests

Subtest Name	Content	Number of Items ASVAB	Number of Items PERT
Word Knowledge (WK)	Understanding the meaning of words, i.e. vocabulary	35	10
Arithmetic Reasoning (AR)	Word problems emphasizing mathematical reasoning rather that mathematical knowledge	30	10
Paragraph Comphrehension (PC)	Understanding the meaning of paragraphs	15	10
Numerical Operations (NO)	A speeded test of four arithmetic operations, i.e., addition, subtraction, multiplication and division	50	None
Mechanical Comprehension (MC)	Knowledge of general mechnical and physical principles	25	10
Electronics Information (EI)	Knowledge of electronics and radio principles	20	10
Auto-Shop Information (AS)	Knowledge of auto mechanics, shop practices and tool functions.	25	10
Mathematics Knowledge (MK)	Knowledge of algebra, geometry and fractions	25	10
General Science (GS)	Knowledge of the physical and biological sciences	25	10
Coding Speed (CS)	A speeded test of matching words and numbers	84	72

During the last week in April, each of the designated station commanders was sent a packet of materials which included: a letter explaining the purpose of the research project; PERT Test booklets and IBM scorable answer sheets; test administration instructions. The station commanders were instructed to administer the PERT to all applicants. An individual was defined as an applicant when a recruiter completed a USAREC Form 714 for the individual. (A Form 714 is completed on applicants who are intending to be tested by the operational ASVAB). After that time, and before the applicant received the operational ASVAB, the station commander administered the PERT and had the applicant record all responses on a single answer sheet. The testing room normally used by the recruiters at the recruiting station was used for this purpose.

Subjects

This procedure yielded 2,921 answer sheets returned to ARI. Nineteen of these were dropped due to missing social security numbers. The distribution of the remaining 2,902 responses by DRC and Regional Recruiting Command (RRC) are presented in Tables 2 and 3, respectively. A large portion of the PERT testing was completed in May with the remainder finishing in June. Accordingly, for data analyses the May applicants were treated as the developmental sample and the June applicants as the validation sample. The criteria, the actual ASVAB test results, were obtained from the Military Processing Command (MEPCOM) for the May and June applicants. When the social security numbers were matched against those on the MEPCOM tapes, 1,058 and 478 matches were found for the May and June applicants respectively. The applicants who had results for tests other than ASVAB 8/9/10 were dropped, yielding developmental and validation samples of 1,047 and 473 respectively. The distribution of these two samples with respect to the DRC where PERT was administered is also presented in Tables 2 and 3.

An inspection of Table 2 indicates that the applicants tested were nationally distributed with no apparent regional biases. Table 3 summarizes these data with respect to the five recruiting regions. Comparing the percent of matched returns between the developmental and validation sample in each region highlights a problem with using month of testing as the basis for dividing the total sample into a developmental and validation sample. Since the majority of the applicants in the Northeast and Southeast Regions were tested early (i.e., in May) these two regions are over-represented in the developmental sample and under-represented in the validation sample. Just the opposite occurred for the Southwest, Midwest and Western Regions. While this unequal distribution is unfortunate, it will serve to make the validation of the regression weights computed in the developmental sample a more stringent test, since the validation sample may be less similar to the developmental sample. This will tend to decrease the size of the cross-validated Regression Coefficients. To ascertain the extent to which the developmental and validation samples differed, demographic characteristics and AFQT scores were examined for each sample.

Table 2

The Number of Applicants Tested on the PERT in Each District Recruiting Command (DRC)

District Recruiting Command	Number of Applicants Tested with PERT	Matched with ASVAB: Developmental	Matched with ASVAB: Validation
Northeast Region			
Albany	28	13	--
Baltimore	33	15	4
Boston	33	9	1
Concord	25	16	--
Harrisburg	49	30	1
New Haven	28	7	--
Long Island	103	37	1
Newburg, NY	39	22	--
Ft. Monmouth, NJ	--	--	--
Niagara Falls	67	28	--
Philadelphia	81	28	2
Pittsburg	29	14	--
Syracuse	87	31	--
Total	602	250	9
Southwest Region			
Albuquerque	31	11	3
Dallas	67	29	6
Denver	18	1	3
Houston	43	17	8
Jackson	34	13	5
Kansas City	46	28	10
Little Rock	7	8	15
New Orleans	31	18	6
Oklahoma City	64	18	17
San Antonio	43	15	9
Total	384	158	82
Southeast Region			
Atlanta	34	14	--
Beckley	--	--	--
Charlotte	75	24	3
Columbia	76	17	3
Jacksonville	53	29	7
Louisville	12	8	5
Miami	81	40	6
Montgomery	58	18	5
Nashville	90	27	13
Raleigh	108	37	2
Richmond	15	6	7
San Juan	85	9	19
Total	687	229	70

Table 2 (continued)

District Recruiting Command (DRC)	Number of Applicants Tested with PERT	Matched with ASVAB: Developmental	Matched with ASVAB: Validation
Midwest Region			
Chicago	27	6	9
Cincinnati	57	23	15
Cleveland	105	31	31
Columbus	99	44	22
Des Moines	16	9	7
Detroit	27	11	7
Indianapolis	60	27	16
Lansing	39	11	11
Milwaukee	72	30	17
Minneapolis	56	21	17
Omaha	36	17	12
Peoria	78	24	26
St. Louis	56	21	11
Total	728	275	201
Western Region			
San Francisco	72	32	14
Honolulu	--	--	--
Los Angeles	40	14	4
Phoenix	72	17	19
Portland	23	5	6
Sacramento	59	10	19
Salt Lake City	25	11	5
Santa Ana	89	21	21
Seattle	45	19	19
Total	425	129	107

Note: The number of applicant test results with unidentifiable DRC was 76 for the total number of applicants tested, 6 for the developmental sample, and 4 for validation sample.

Table 3

The Number of Applicants Tested on the PERT in Each Regional Recruiting Command (RRC)

Regional Recruiting Command (DRC)	Number of Applicants Tested with PERT	Number of Applicants Matched with ASVAB Developmental	Number of Applicants Matched with ASVAB Validation
Northeast	602	250 (42%)	9 (1%)
Southwest	384	158 (41%)	82 (21%)
Southeast	687	229 (33%)	70 (10%)
Midwest	728	275 (38%)	201 (28%)
Western	425	129 (30%)	107 (25%)
	2,826	1,041 (37%)	469 (17%)

The distribution of both samples with respect to the demographic variables of Gender, Race and Education information available from the MEPCOM tapes, is presented in Table 4. The validation sample includes a slightly higher proportion of whites than the developmental sample, i.e., 71% vs. 61%. The proportion of males in both samples is identical, 78%. The distribution of applicants by education for the developmental sample is 41% for high school graduates, 14% for high school seniors and 45% for those with General Educational Diplomas (GED). The corresponding distribution for the validation sample is 47%, 11% and 42%. Thus, the developmental and validation samples are similar with respect to demographic characteristics.

A critical dimension of both samples is the range of scores on the AFQT portion of the operational ASVAB. Table 5 reveals that the developmental and validation samples are similar in respect to AFQT scores. Both samples include a large proportion (71% and 67%) of applicants who scored below the 50th percentile. These proportions are fortuitous for our purposes because the PERT predictions will be most useful for those applicants in the lower ability levels who may not qualify for all MOS. In general, comparisons between the developmental and validation samples indicate that there does not appear to be any major difference between the two samples, even though the two samples were not equally distributed among the five Recruiting Regions.

Analyses

To examine some psychometric properties of the PERT, subtest scale means, standard devitions and Cronbach's coefficient alpha's, an index of the internal consistency reliability, were computed in the developmental sample. The validation of the PERT was accomplished by computing eleven separate regression equations in the developmental sample. For each regression analysis, the PERT subtest raw scores served as the predictor variables and each of the ten Army Aptitude Area Composites and the AFQT successively served as the criterion variable. Two features of these regression analyses need to be explicated. First, the PERT was designed to predict Army Area Aptitude Composites, not ASVAB individual subtest scores. The ten Area Aptitude Composties with corresponding MOS are presented in Table 6. Second, one PERT subtest, Coding Speed (CS), was excluded as a predictor in the regression analyses. Preliminary analyses of the distribution of CS scores showed many high scores. The distribution indicated that about half of the applicants were apparently allowed to respond beyond the 5-minute time of the subtest, thus invalidating the results.

Once the regression equations were computed in the developmental sample, the regression coefficients (non-standardized beta weights) were used to compute predicted ASVAB composite scores in the validation sample. The correlation between the predicted and actual ASVAB composite scores in the validation sample constituted the "validated" multiple R.

Table 4

The Distribution of the Developmental and Validation Samples with Respect to Gender, Race and Education

			Developmental	Validation
White	HS Graduate	Males	191	114
		Female	61	41
	HS Seniors	Male	56	28
		Female	21	9
	GED	Male	227	124
		Female	30	19
			636 (61%)	335 (71%)
Black	HS Graduate	Male	103	37
		Female	60	19
	HS Senior	Male	43	5
		Female	20	9
	GED	Male	108	43
		Female	27	5
			361 (34%)	118 (25%)
Other	HS Graduate	Male	15	8
		Female	4	2
	HS Senior	Male	5	1
		Female	1	0
	GED	Male	22	8
		Female	3	1
			50 (5%)	20 (4%)
		Total	1,047 (100%)	473 (100%)

Table 5

The Distribution of AFQT Scores
on the Operational ASVAB for the
Developmental and Validation Samples

AFQT Mental Category	Percentile Score	Developmental Sample (n = 1,047) Percent	Developmental Sample (n = 1,047) Cum. Percent	Validation Sample (n = 473) Percent	Validation Sample (n = 473) Cum. Percent
1	93-99	2	100	1	100
2	65-92	18	98	20	99
3A	50-64	9	80	12	79
3B	31-49	21	71	26	67
4A	21-30	14	50	13	41
4B	16-20	12	36	12	28
4C	10-15	15	24	9	16
5	1-9	9	9	7	7

Table 6

The Area Aptitude Composite Prerequisite for Major Groups of Army MOS

Area Aptitude Composite	Component ASVAB Subtests[1]	Military Occupational Specialties (MOS)
CO (Combat)	AR+AS+MC+CS	Infantry, Armor, Combat Engineer
FA (Field Artillery)	AR+MK+MC+CS	Field Cannon and Rocket Artillery
EL (Electronics Repair)	AR+MK+EL+GS	Missles Repair, Air Defense Repair, Electronics Repair, Fixed Plant Communications Repair
OF (Operators and Food)	NO+VE[2]+MC+AS	Missles Crewmen, Air Defense Crewmen, Driver, Food Services
SC (Surveillance and Communication)	NO+CS+VE+AS	Target Acquisition and Combat Surveillance, Communication Operations
MM (Mechanical Maintenance)	NO+EI+MC+AS	Mechanical and Aircraft Maintenance, Rails
GM (General Maintenance)	MK+EI+GS+AS	Construction and Utilities, Chemical, Marine Petroleum
CL (Clerical)	NO+CS+VE	Administrative, Finance, Supply
ST (Skilled Technical)	VE+MK+MC+GS	Medical, Military Policeman, Intelligence, Data Processing, Air Control, Topography and Printing, Information and Audio Visual
GT (General Technical)	VE+AR	Not currently used for classification into a particular MOS

Note: [1] standardized subtest scores are used in computation of composites.
[2] Verbal (VE) is a standard score conversion of the sum of the raw scores for WK and PC.

Results

Table 7 presents the means, standard deviations and reliability coefficients for the PERT subtests. Mathematics Knowledge (MK) was the most difficult of the PERT power subtests. Its mean of 3.5 is substantially lower than the means for the other subtests. The mean and standard deviation for Coding Speed (CS) is misleading since, as mentioned previously, many applicants were apparently allowed to go beyond the specified time limit. The reliabilities of the power scales are quite adequate for ten-item scales.

Results of the regression analyses are presented in Table 8. The multiple R's in the developmental sample were quite high, ranging from a low of .73 for the Clerical (CL) Aptitude Area to a high of .82 for the Skilled Technician (ST) and General Technical (GT) Aptitude Areas and the AFQT. When the regression coefficients were validated in the second sample, multiple R's decreased only slightly, with the exception of predicting CL, which decreased from .73 to .63. In general, however, the PERT appears to be a useful predictor of Army ASVAB composites.

An examination of the standardized beta weights for the PERT subtests (Table 8) reveals that the most useful subtests were Arithmetic Reasoning (AR), Word Knowledge (WK), Mechanical Comprehension (MC) and Auto/Shop Information (AS) followed by Math Knowledge (MK), Paragraph Comprehension (PC) and General Science (GS). The PERT subtest Electrical Information (EI) did not achieve a single beta weight above .10. Comparisons of the distribution of beta weights across the PERT subtests in each regression equation with the actual corresponding ASVAB subtests used to compute ASVAB composites (indicated by underlinings in Table 8), suggests only a moderate relationship. This may be the result of sample fluctuations as the correlations of the operational ASVAB subtests are moderate to high (see Sims and Truss, 1980; Table K-3). It must also be noted that the PERT subtests are typically based on half (or less) of the number of items than the ASVAB subtests and, consequently, have lower reliabilities. The ASVAB subtest scale reliabilities range from .80 to .93 (Ree, Mullins, Mathews and Massey, in press). This lower reliability of the PERT scales would account for some of the coefficient attenuation.

Summary and Conclusion

In this investigation, a newly developed test was validated. The PERT was designed to be used by recruiters to test Army applicants in order to obtain an indication of the applicant's eligibility for enlistment as well as for specific MOS. One of the requirements for MOS entry is a qualifying score on the relevant Army Aptitude Area Composite. The PERT responses of 1,047 May 1981 Army applicants were used to develop regression weights to predict their ASVAB composite scores. The prediction equations were validated in an independent sample of 473 June 1981 Army applicants. The results from this research demonstrate that the PERT would be a useful tool for recruiters. The final step remaining before implementation of the PERT as an Army recruiting tool is to equate the PERT scales to the ASVAB scales. Recruiters need to know what ASVAB values are predicted by which PERT values. This step will be compeleted in the near future.

Table 7

Means, Standard Deviations and Reliabilities
of PERT Subtest Scales

PERT Subtest [1]	Mean	Standard Deviation	Coefficient Alpha
Word Knowledge (WK)	5.6	2.6	.74
Arithmetic Reasoning (AR)	5.1	2.6	.76
Paragraph Comprehension (PC)	5.7	2.6	.73
Mechanical Comprehension (MC)	5.4	2.3	.66
General Science (GS)	5.2	2.2	.62
Electronics Information (EI)	5.0	2.3	.60
Mathematics Knowledge (MK)	3.5	2.2	.64
Auto/Shop Information (AS)	4.9	2.5	.69
Coding Speed (CS)[2]	36.0	18.1	

Note: [1]All PERT subtests have 10 items apiece, except CS which has 72.

[2]Coefficient Alpha was not computed for CS since it is a speeded test.

Table 8

Regression Weights and Multiple R's for Predicting
Army Aptitude Area Composites from PERT Subtest Scores

Army Aptitude Area Composite	Multiple R: Developmental Sample	Multiple R: Validation Sample	Standardized Beta Weights for PERT Subtests[1,2]: WK	AR	PC	MC	EI	AS	MK	GS	CS
Combat (CO)	.78	.74		.31		.18		.24			___
Field Art. (FA)	.78	.72		.35		.17			.20		___
Electronics (EL)	.80	.77	.21	.28		.12	___	.12	.12		
Oper/Foods (OF)[3]	.80	.76	.18	.24	___	.16		.22			
Surv/Comm (SC)[3]	.77	.70	.17	.26	.15			.13	.11		
Motor Maint. (MM)[3]	.77	.75	.12	.22		.17	___	.28			
Gen. Maint. (GM)[3]	.78	.77	.21	.18	___	.13	___	.24	___	___	
Clerical (CL)[3]	.73	.63	.18	.26	.20				.18		___
Skilled Tech (ST)	.82	.77	.26	.20	___	.18		.11	.11	___	
Gen. Tech (GT)	.82	.75	.24	.35	.15	.10					
AFQT	.82	.76	.23	.35	.12				.16		

[1]ASVAB subtests used to compute composites are indicated by underlining.
[2]Standardized Beta Wts. less than .10 are omitted.
[3]ASVAB subtest NO also used to compute operational ASVAB Army compoiste.

References

Bayroff, A. G., Thomas, J. A. & Kehr, C. J. Evaluation of EST for Predicting AFQT Performance. (ARI Technical Research Report 1114). Alexandria, VA: Army Research Institute, February, 1959.

Jensen, H. E. & Valentine, Jr., L. D. Development of the Enlistment Screening Test - EST Forms 5 and 6. (AFHRL Technical Report 76-42). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, May, 1976.

Mathews, J. J. Development and Calibration of the Enlistment Screening Test (EST) Forms 81a and 81b. Paper presented at the meeting of the Military Testing Association, March, 1981.

Morton, M. A., Houston, T. J. & Bayroff, A. G. Development of Enlistment Screening Test, Forms 3 and 4. (ARI Technical Research Report 1102). Alexandria, VA: Army Research Institute, May, 1957.

Ree, M. J., Mullins, C. J., Mathews, J. J., & Massey, R. H. Armed Services Vocational Aptitude Battery: Item and factor analysis of Forms 8, 9, and 10. (AFHRL-TR-xx). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division, in press.

Sims, W. H. & Truss, A. R. Normalization of the Armed Services Vocational Aptitude Battery ASVAB) Forms 8, 9, and 10 using a sample of service recruits. (CNA Research Contribution 438). Alexandria, VA: Center for Naval Analyses, December, 1980.

Footnotes

The assistance of CPT David Duke of the U.S. Army Recruiting Command in coordinating the distribution and collection of tests was invaluable. John Mathews with the Air Force Human Resources Lab was very helpful in disussion of data collection procedures and providing Paragraph Comprehension (PC) items. Dr. Kent Eaton, with ARI, provided critical institutional support and coordination. The opinions expressed are those of the authors and do not necessarily reflect those of the Army Research Institute or the Department of the Army.

Copies of this paper can be obtained from the senior author, U.S. Army Research Institue, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.