CRM 90-120 / December 1990

Construction of Final Forms for a New Enlistment Screening Test

D. R. Divgi

CNA
CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

Work conducted under contract N00014-91-C-0002.

REPORT DOCUMENTATION PAGE

Form Approved OPM No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1990	Final

4. TITLE AND SUBTITLE
Construction of Final Forms for a New Enlistment Screening Test

5. FUNDING NUMBERS
C - N00014-91-C-0002
PE - 65153M
PR - C0031

6. AUTHOR(S)
D.R. Divgi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Naval Analyses
4401 Ford Avenue
Alexandria, Virginia 22302-0268

8. PERFORMING ORGANIZATION REPORT NUMBER
CRM 90-120

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commanding General
Marine Corps Combat Development Command (WF 13F)
Studies and Analyses Branch
Quantico, Virginia 22134

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Two forms, each containing 35 verbal and 30 mathematics items, have been developed for a new Enlistment Screening Test (EST) to predict Armed Forces Qualification Test (AFQT) scores of military applicants. These forms were constructed in two stages from items in discontinued versions of the Defense Department's test batteries. The first stage was to develop overlength forms from the available item pool. This research memorandum describes the second stage: constructing final forms by selecting items from the overlength forms. Item selection was based on the correlation of the item with AFQT, in a subsample of applicants with AFQT percentiles between 21 and 65. For each EST form, the AFQT score was predicted from the total score on the final EST items. The results were used to calculate expectancy tables which, for any given EST score, provide probabilities of exceeding the specified AFQT cutoffs. These probabilities are reported in tables.

14. SUBJECT TERMS
AFQT (armed forces qualification test), Aptitude tests, ASVAB (armed services vocational aptitude battery), Forms (paper), Performance (human), Performance tests, Personnel selection, Predictions, Recruiting, Regression analysis, Scoring, Tables (data), Test construction, Test scores

15. NUMBER OF PAGES 20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
CPR	CPR	CPR	SAR

CNA CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

9 January 1991

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 90-120

Encl: (1) CNA Research Memorandum 90-120, *Construction of Final Forms for a New Enlistment Screening Test*, by D.R. Divgi, Dec 1990

1. Enclosure (1) is forwarded as a matter of possible interest.

2. Two forms, each containing 35 verbal and 30 mathematics items, have been developed for a new Enlistment Screening Test (EST) to predict Armed Forces Qualification Test (AFQT) scores of military applicants. These forms were constructed in two stages from items in discontinued versions of the DOD's test batteries. The first stage was to develop overlength forms from the available item pool. This research memorandum describes the second stage: constructing final forms by selecting items from the overlength forms. Item selection was based on the correlation of the item with AFQT, in a subsample of applicants with AFQT percentiles between 21 and 65. For each EST form, the AFQT score was predicted from the total score on the final EST items. The results were used to calculate expectancy tables which, for any given EST score, provide probabilities of exceeding the specified AFQT cutoffs. These probabilities are reported in tables.

Lewis R. Cabe
Director
Manpower and Training Program

Distribution List:
Reverse page

Subj: Center for Naval Analyses Research Memorandum 90-120

Distribution List

SNDL

A1	DASN - MANPOWER
A1H	ASSTSECNAV MRA
A2A	CNR
A6	HQMC MPR & RA
	Attn: M
	Attn: MP
	Attn: MR
	Attn: MA
	Attn: MPP-54
FF38	USNA
	Attn: Nimitz Library
FF42	NAVPGSCOL
FF44	NAVWARCOL
	Attn: E-111
FJA1	COMNAVMILPERSCOM
FJA13	NAVPERSRANDCEN
	Attn: Technical Director (Code 01)
	Attn: Director, Testing Systems (Code 13)
	Attn: Technical Library
	Attn: Director, Personnel Systems (Code 12)
	Attn: CAT/ASVAB PMO
	Attn: Manpower Systems (Code 11)
FJB1	COMNAVCRUITCOM
FT1	CNET
V12	MCCDC
	Attn: Training and Education Center
	Attn: Warfighting Center (WF-13F)

OPNAV

OP-11B
OP-136

OTHER

Military Accession Policy Working Group (17 copies)
Defense Advisory Committee on Military Personnel Testing (8 copies)

CRM 90-120 / December 1990

Construction of Final Forms for a New Enlistment Screening Test

D. R. Divgi

Force Structure and Acquisition Division

CNA
CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

Two forms, each containing 35 verbal and 30 mathematics items, have been developed for a new Enlistment Screening Test (EST) to predict Armed Forces Qualification Test (AFQT) scores of military applicants. These forms were constructed in two stages from items in discontinued versions of the Defense Department's test batteries. The first stage was to develop overlength forms from the available item pool. This research memorandum describes the second stage: constructing final forms by selecting items from the overlength forms. Item selection was based on the correlation of the item with AFQT, in a subsample of applicants with AFQT percentiles between 21 and 65. For each EST form, the AFQT score was predicted from the total score on the final EST items. The results were used to calculate expectancy tables which, for any given EST score, provide probabilities of exceeding the specified AFQT cutoffs. These probabilities are reported in tables.

EXECUTIVE SUMMARY

The Enlistment Screening Test (EST) is used by military recruiters to predict how a potential applicant is likely to score on the Armed Forces Qualification Test (AFQT). Persons with low EST scores can be screened out as being unlikely to pass the AFQT standard. Persons with high EST scores can be encouraged to apply by describing available incentives such as bonuses and enlistment guarantees.

CNA has developed a new EST because the Marine Corps felt that the previous EST had become obsolete. The development had two stages: In the first stage, described in an earlier research memorandum, CNA constructed two overlength forms (containing about 50 percent more test items than would be needed in the final forms) from items in discontinued versions of the DOD's test batteries. The overlength forms were administered to applicants for military enlistment and the resulting data were sent to CNA. AFQT scores of the applicants were obtained from the Defense Manpower Data Center (DMDC). In the second stage of analysis, the data on overlength forms and AFQT were used to select items for the final EST forms and to compute performance prediction tables. This research memorandum describes the second stage, i.e., construction of final forms and calculation of prediction tables.

At first the final forms were constructed for the Marine Corps using USMC data. These forms were then printed and distributed to USMC recruiters. However, other services also expressed interest in using the new EST and provided data on overlength forms. The data from the Marine Corps, the Navy, and the Air Force were therefore analyzed together to construct a Joint Service EST. Items having high correlations with the AFQT were included in the final forms. Separate expectancy tables were developed for EST Forms A and B. These tables provide probabilities of exceeding specified cutoff scores on the AFQT from a potential applicant's EST score.

Subgroup analyses showed that, at any EST score, mean AFQT was higher for whites than for blacks. With concurrence from the Military Accession Policy Working Group, the author decided to use only the white subsample while computing prediction tables. Therefore, to some extent, the tables overpredict the AFQT scores of blacks.

The new EST forms, along with their expectancy tables, were printed in February 1989 and distributed to recruiters in all four services.

CONTENTS

INTRODUCTION

The Enlistment Screening Test (EST) is used by military recruiters to predict how a potential applicant is likely to score on the Armed Forces Qualification Test (AFQT). Persons with low EST scores can be screened out as being unlikely to pass the AFQT standard. Persons with high EST scores can be encouraged to apply by describing available incentives such as bonuses and enlistment guarantees.

A new EST has been developed because the Marine Corps felt that the previous EST had become obsolete [1]. The development had two stages: In the first stage, two overlength forms (containing about 50 percent more test items than would be needed in the final forms) were constructed. In the second stage, data on overlength forms were used to select items for the final forms.

The AFQT now consists of the Word Knowledge (WK), Paragraph Comprehension (PC), Arithmetic Reasoning (AR), and Mathematics Knowledge (MK) subtests of the Armed Services Vocational Aptitude Battery (ASVAB). For optimum prediction of AFQT scores, content of the EST should resemble that of the AFQT as much as practicable. PC was excluded because it takes three times as long per item as WK does, while measuring almost the same construct. The author therefore decided that the verbal part of the new EST would consist of 35 WK items (the same number as in the ASVAB), and the mathematics part would contain 30 AR and MK items (the same number as in AR). The ratio of AR and MK items was not preset; the numbers of these items were to depend on the results of the item selection procedure, in which AR and MK would be treated as measuring the same trait.

With permission from the Joint Service Selection and Classification Working Group, CNA used items from discontinued forms of the ASVAB and the AFQT. These forms were ASVAB 5X, 6X, 7X, 6E, 7E, and AFQT7A. The overlength forms were to contain 55 verbal and 45 math items so that at least a third of the items would be deleted on the final forms. The goal was to predict AFQT scores as accurately as possible, emphasizing AFQT percentile ranks of 31 and 50, which are the lower-end points of AFQT Categories IIIB and IIIA [1].

Development of the overlength forms has been described in an earlier CNA publication [2]. This research memorandum describes the selection of items for the final forms from those in the overlength forms, and the calculation of expectancy tables for predicting AFQT performance from the EST score.

MARINE CORPS EST

Overlength forms were printed by Headquarters, Marine Corps (HQMC), in May 1987 and distributed to USMC recruiters for experimental use during a limited period. HQMC provided CNA with applicants' answer sheets in January 1988. Item responses were entered into computers independently at HQMC and CNA. The two data sets were then compared to find typing errors. The corrected data file was sent to the Defense Manpower Data Center (DMDC) for operational ASVAB scores. DMDC used social security numbers (SSNs) to match EST and ASVAB records.

In accordance with the original USMC request to CNA [1], which emphasized the AFQT range between the 21st and 65th percentiles, only the applicants with AFQT scores in this range were used for selecting items for the final forms. Each item was correlated with the AFQT sum of standard scores. In each EST form, verbal items with the highest 35 correlations and math items with the highest 30 correlations were chosen for the final forms. This concluded the item selection phase. Then, to predict AFQT from EST, AFQT scores were regressed on total scores for the selected EST items. These regressions, performed independently for the two forms, provided the expected AFQT score at each EST score. The Marine Corps EST and the tables of predicted AFQT scores were printed and distributed by HQMC in July 1988.

Detailed results of these analyses are not provided in this document because the Marine Corps version has been superseded by the joint service version. Some details were given in the author's briefing to the Defense Advisory Committee (DAC) on Military Personnel Testing in October 1988 [3]. As work on the Marine Corps version was nearing completion, other services expressed interest in using the new EST. The Navy and the Air Force provided data on overlength forms, which were then added to the Marine Corps data. This document reports the analyses of the total joint service data.

DATA QUALITY

As before, DMDC matched EST records with operational records to add ASVAB scores to the data. The size of the matched sample was 1,281 for Form A and 1,109 for Form B. One important concern was whether recruiters had refrained from administering the ASVAB to applicants who had low scores on the overlength EST. Recruiters were instructed to administer the ASVAB to everyone who took the overlength EST during data collection; however, the extent of compliance with this instruction was unknown. The EST item responses were therefore scored and matched and unmatched groups were compared on their total EST scores.

The results of this comparison are presented in table 1. The means tend to be higher in the matched group, especially in the Navy and the Air Force. This trend may indicate that recruiters did in fact reject some applicants on the basis of their overlength EST scores, even though no tables for interpreting the scores had been provided. On the other hand, it may be that those who had low scores on the EST also tend to be careless in writing their social security numbers. Because of this ambiguity, and also because the alternative was to use Marine Corps data only, it was decided to use the Navy and Air Force samples.

Table 1. Sample sizes, means, and standard deviations of overlength EST scores

Service	Matched group N	Mean	S.D.	Unmatched group N	Mean	S.D.
			Form A			
USMC	580	71.4	18.7	398	71.5	21.5
A.F.	320	72.2	20.5	238	66.2	22.9
Navy	381	66.6	23.3	240	61.9	25.3
			Form B			
USMC	558	71.2	18.8	325	69.4	21.4
A.F.	363	73.4	18.8	212	68.8	23.0
Navy	188	67.9	21.4	137	62.8	27.4

Another concern related to data quality is the motivation of examinees. Because the recruiters knew the EST score had no effect on the application, the applicants also probably knew this. Therefore, the motivation of examinees might have varied between those who chose to answer the items carefully and those who did not. The ASVAB scores were operational, i.e., they were used to make selection and classification decisions about the applicants. Consequently, it was assumed that applicants were fully motivated while taking the ASVAB, and that ASVAB scores could therefore be used to evaluate motivation on the EST.

The AFQT sum of standard scores is given by

$$SSS = 2\ S\_VE + S\_AR + S\_MK \quad , \qquad (1)$$

where "S_" indicates a standard score. SSS was computed for each applicant who had taken the ASVAB. Some applicants might have been careless enough to score relatively lower on the EST than on the AFQT. These applicants appear as outliers on a scatterplot of the total overlength EST score against SSS. The scatterplot for Navy applicants is shown in figure 1. The outliers at lower right are clearly separated from most of the sample.

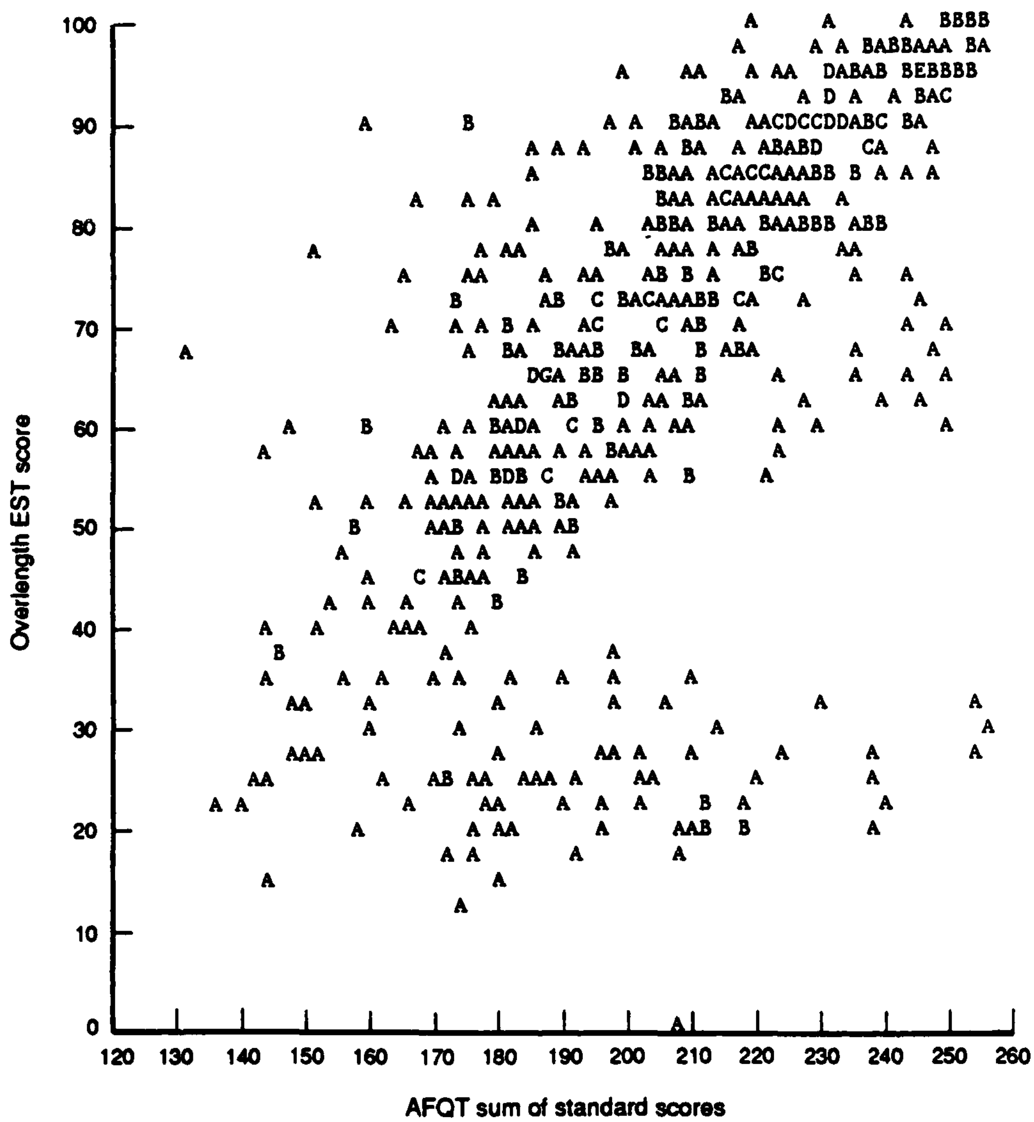

Figure 1. Plot of total overlength EST score against AFQT SSS, Navy data only

To identify outliers quantitatively, a cubic regression of EST score on SSS was fitted with forms A and B combined. The regression curves for the Marine Corps and Air Force were quite close at the 50th percentile of SSS, while the curve for Navy was four points lower. This indicated lower average motivation among Navy applicants. The Marine Corps and Air Force samples were therefore combined and a quadratic regression was fitted. The standard error of estimate was 14.6. Those who scored more than 40 points below their predicted overlength EST score were excluded from further analyses. Examinees were also excluded if they scored below the chance level of 25. The remaining sample size was 1,205 for Form A and 1,067 for Form B. The total sample contained 15.5 percent women and 20.6 percent blacks.

ITEM SELECTION

Five different indicators of item quality were compared by cross-validation. Two used item-SSS correlation in the sample and in a subsample of examinees scoring between the 21st and 65th percentiles of AFQT. Two procedures used a graphical procedure in Lord and Novick [4], again using the entire sample and the subsample. The fifth indicator used logistic regression of the item on SSS. To compare these measures, the total sample for each form was split randomly, with 50-50 probabilities, into a selection sample and a validation sample. Data from the selection sample were used to select items for the final EST forms, and those from the validation sample were used to evaluate the final forms. The best method was that which yielded the smallest residual variance on regressing AFQT on EST in the validation sample.

Results of these comparisons, presented to the DAC in February 1989, indicated that the best index for item selection was item-SSS correlation in the subsample [5]. This index was therefore used to develop the joint service EST forms. However, it turned out later that the results of the cross-validation were sensitive to minor changes in data editing rules, and also changed when a different pair of random selection and validation samples were created from the same total sample. It appears that true differences among tests constructed by the five methods are not large enough to be detected reliably with the available sample sizes. For that reason, detailed results of the comparison are not reported here.

DIFFERENTIAL ITEM FUNCTIONING

During test construction, items that suffer from differential item functioning (DIF), i.e., items that are harder for some groups of examinees than for others, must be eliminated. A good indicator of DIF is the Mantel-Haenszel (MH) statistic [6], which compares the difficulty of an item in a reference group with that in a focal group. Usually, these groups are a majority and a minority defined on the basis of race, ethnicity, or gender. Individuals in the two groups are matched on subtest scores, and at a given subtest score, the proportions of correct answers within the two groups are compared. When there is no DIF and the sample size is large, the MH statistic has a chi-square distribution with one degree of freedom ([6], p. 8). The desirable features of the MH statistic are simplicity and support in statistical theory. This statistic also provides a measure of effect size, i.e., of the extent to which the item functions differently in the two groups, but this measure was not needed in the present study.

The comparisons of interest were between blacks and whites and between women and men. (Hispanics could not be identified reliably from the ethnicity code available.) The data at a given subtest score cannot be used unless at least one examinee from each subgroup gets that score. Consequently, a part of the sample is unavailable for DIF analysis. For each person, unanswered items were excluded from the calculations. The sample size available for DIF analysis therefore varied somewhat from one item to another. The minimum DIF sample sizes for the two forms were 226 and 194 among blacks, and 177 and 153 among women.

Exclusion of items requires a decision rule. An item was eliminated if its MH statistic exceeded 6.63, which is the 99th percentile of the chi-square distribution, with one degree of freedom. In the total of 200 items in the two overlength forms, 12 were rejected in the MH analyses by race, and 14 were rejected by gender. The rejected items were either easier or harder equally often for the minority groups.

FINAL ITEM SELECTION

Item selection for the final forms used the item-SSS correlation in the subsample of applicants with AFQT percentiles between 21 and 65 (inclusive). Omitted items were treated as missing data. In each subtest, items were arranged in descending order by correlation. Items were selected in order from the top, excluding those rejected by the MH statistic. The number of items in the final forms was 35 for verbal and 30 for mathematics. The total number with acceptable correlations, but unacceptable MH statistics, was 6 in Form A and 16 in Form B. Items selected for use in the final forms are listed in the appendix. Summary statistics of the scores on these forms are given in table 2. Reliabilities were not calculated because they may capitalize on chance, being based on the same data that were used to select the items from the overlength forms.

Table 2. Descriptive statistics on final forms[a]

	Mean scores			Standard deviations		
Form	Verbal	Math	Total	Verbal	Math	Total
A	26.9	20.5	47.4	6.8	6.7	11.9
B	27.0	19.5	46.5	6.5	6.9	11.9

a. With Form A and B samples combined, AFQT SSS had a mean of 207.0 and a standard deviation of 25.3.

REGRESSION IN SUBGROUPS

Even after items with large DIF have been eliminated, regression of SSS on EST may not be the same in all subgroups. Equality of linear regressions across subgroups was therefore tested, using the GLM procedure in SAS [7]. This procedure allows independent statistical tests for the intercept and the slope of regression. To properly evaluate the subgroup differences in intercept, the mean of the minority was subtracted from the EST score in each regression analysis. Consequently, the difference between the intercepts equals the difference between predictions of the majority and minority regressions for the average member of the minority group.

Differences between men and women were not statistically significant at the .05 level. On Forms A and B, the F ratios for intercepts were 0.7 and 2.3; those for slope were 0.2 and 3.1. On the other hand, differences between blacks and whites were significant. F ratios were 22.3 and 58.4 for intercept, 5.4 and 2.8 for slope.

The regression lines were lower for blacks than for whites, the difference in intercepts being 5.0 SSS points with Form A, 10.2 with Form B. In other words, at any given EST score on Form A, the mean AFQT SSS was five points lower among blacks than among whites; for Form B the gap between regression lines was 10 points. While such differences are undesirable, they have been observed frequently [8]. Regression among whites was used to develop expectancy tables. This overpredicts the AFQT scores of blacks, and thus makes it easier for low-scoring blacks to appear eligible for ASVAB testing.

EXPECTANCY TABLES

Calculation of expectancy tables requires a regression equation to predict SSS from EST. In the white subgroup, the correlation of EST with SSS was .737 with Form A and .708 with Form B. Stepwise regression showed that a quadratic term was necessary, but a cubic term was not.

The multiple correlation was .739 for Form A and .708 for Form B. The standard errors of estimate were 15.6 and 16.7. The residual values, i.e., actual SSS minus the predicted values, were stored for further analyses.

The residuals were squared, and regressed on the first and second powers of the EST score. The results showed that variance of the residuals was not related to EST score, and could therefore be treated as constant. For computing expectancy tables, residuals were treated as being normally distributed, with standard deviation equal to the standard error of estimate.

The probability of obtaining or exceeding the 21st percentile of AFQT was computed as follows: For example, with EST = 25 on Form A, the expected value SSS calculated from the quadratic regression is 175.9. The boundary between the 20th and 21st percentiles of SSS is 165.5 [9]; therefore, the corresponding standard normal score is

$$z = (165.5 - 175.9) / 15.6 = -0.667 \quad .$$

The standard normal probability of exceeding this z score is .747, which after rounding is reported as 75 percent in the table. Such calculations were performed separately for Forms A and B, at EST scores 21 to 65 and AFQT percentiles of 21, 31, 50, and 65. The results are in the appendix.

Similar tables were computed for the Marine Corps' and Army's General Technical and Air Force's General composites, and then provided to the services. The joint service EST forms, along with the expectancy tables for AFQT, were printed and distributed in February 1989.

REFERENCES

[1] Memorandum from the Deputy Chief of Staff for Manpower (MPI-20), *Construction of an Enlistment Screening Test (EST) for Use with the Armed Services Vocational Aptitude Battery (ASVAB)*, 15 Oct 1986

[2] CNA Research Memorandum 90-119, *Development of Overlength Forms of a New Enlistment Screening Test*, by D. R. Divgi, in process

[3] Briefing to the Defense Advisory Committee on Military Personnel Testing, *Development of a New Enlistment Screening Test*, by D. R. Divgi, 31 Oct 1988

[4] Frederic M. Lord and Melvin R. Novick. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison Wesley, 1968

[5] CNA Memorandum 89-0246, *Joint Service Enlistment Screening Test: Item Selection and Expectancy Tables*, by D. R. Divgi, Feb 1989

[6] Paul W. Holland and Dorothy T. Thayer. *Differential Item Functioning and the Mantel-Haenszel Procedure*. RR-86-31. Princeton, NJ: Educational Testing Service, 1986

[7] SAS Institute, Inc. *Statistical Analysis System*, Version 5.03. Cary, NC: SAS Institute, Inc., 1986

[8] Arthur R. Jensen. *Bias in Mental Testing*. New York: Free Press, 1980

[9] Conversion Tables: *Armed Services Vocational Aptitude Battery Forms 8-17*. Washington, DC: Department of Defense, Jan 1989

APPENDIX A

LISTS OF ITEMS IN FINAL FORMS AND EXPECTANCY TABLES

Table A-1. Joint service enlistment screening test items in verbal parts of final forms

Form A			Form B		
Final item	Overlength item	Key	Final item	Overlength item	Key
1	2	B	1	1	B
2	1	A	2	2	A
3	13	B	3	5	D
4	7	A	4	7	B
5	5	A	5	12	C
6	18	C	6	4	C
7	4	C	7	9	C
8	10	B	8	15	B
9	27	B	9	19	A
10	6	A	10	25	B
11	25	C	11	20	D
12	26	B	12	18	C
13	17	C	13	28	D
14	46	B	14	41	C
15	22	A	15	36	C
16	20	C	16	37	D
17	30	C	17	34	D
18	47	C	18	43	A
19	42	B	19	31	D
20	24	D	20	32	B
21	38	D	21	46	B
22	33	C	22	30	A
23	50	B	23	24	B
24	28	B	24	44	B
25	43	B	25	47	D
26	31	D	26	40	B
27	37	D	27	38	D
28	52	A	28	54	A

Table A-1. (Continued)

Form A			Form B		
Final item	Overlength item	Key	Final item	Overlength item	Key
29	40	A	29	53	A
30	49	C	30	39	C
31	35	D	31	51	C
32	55	D	32	49	A
33	32	D	33	50	D
34	51	D	34	42	A
35	53	A	35	55	C

Table A-2. Joint service enlistment screening test items in mathematics parts of final forms

Form A			Form B		
Final item	Overlength item	Key	Final item	Overlength item	Key
1	3	C	1	2	A
2	10	B	2	3	D
3	23	B	3	4	C
4	4	A	4	6	B
5	8	A	5	11	A
6	7	D	6	7	A
7	6	D	7	22	C
8	11	D	8	17	C
9	12	C	9	21	A
10	19	B	10	9	B
11	17	B	11	10	A
12	14	D	12	29	D
13	22	D	13	14	A
14	21	C	14	32	C
15	32	C	15	35	C
16	27	C	16	15	D
17	16	C	17	34	D
18	24	B	18	26	A
19	30	A	19	24	D
20	31	C	20	40	A

Table A-2. (Continued)

Form A			Form B		
Final item	Overlength item	Key	Final item	Overlength item	Key
21	25	B	21	20	C
22	33	B	22	31	D
23	28	D	23	41	C
24	44	A	24	39	C
25	40	D	25	44	A
26	34	C	26	42	D
27	37	D	27	36	D
28	41	D	28	43	D
29	42	A	29	37	A
30	38	D	30	45	D

Table A-3. Expectancy table to predict AFQT from total EST score, Form A

EST score	Percent chance of AFQT percentile being at least 21	31	50	65
21	65	26	2	0
22	67	28	2	0
23	70	31	3	0
24	72	33	3	0
25	75	36	4	0
26	77	39	4	0
27	79	42	5	0
28	82	45	6	1
29	84	48	7	1
30	86	51	8	1
31	87	55	10	1
32	89	58	11	1
33	91	61	13	2
34	92	65	15	2
35	93	68	17	3
36	94	71	19	3
37	95	74	22	4
38	96	77	25	5
39	97	79	28	6
40	97	82	31	7

Table A-3. (Continued)

EST score	Percent chance of AFQT percentile being at least 21	31	50	65
41	98	84	34	9
42	98	87	38	10
43	99	89	42	12
44	99	90	46	14
45	99	92	50	17
46	99	93	54	19
47	100	95	58	22
48	100	96	62	25
49	100	97	66	29
50	100	97	70	33
51	100	98	73	37
52	100	98	77	41
53	100	99	80	45
54	100	99	83	50
55	100	99	86	54
56	100	100	88	59
57	100	100	90	63
58	100	100	92	67
59	100	100	94	71
60	100	100	95	75
61	100	100	96	79
62	100	100	97	82
63	100	100	98	85
64	100	100	98	88
65	100	100	99	90

Table A-4. Expectancy table to predict AFQT from total EST score, Form B

EST	Percent chance of AFQT percentile being at least			
score	21	31	50	65
21	65	29	3	0
22	68	31	4	0
23	71	34	4	0
24	74	37	5	1
25	76	40	6	1
26	79	43	7	1
27	81	47	8	1
28	83	50	9	1
29	85	53	11	2
30	87	56	12	2
31	89	60	14	2
32	90	63	16	3
33	92	66	18	4
34	93	69	21	4
35	94	72	23	5
36	95	75	26	6
37	96	78	29	7
38	96	80	32	9
39	97	83	35	10
40	98	85	39	12
41	98	87	42	14
42	98	89	46	16
43	99	90	49	18
44	99	92	53	20
45	99	93	56	23
46	99	94	60	26
47	100	95	64	29
48	100	96	67	32
49	100	97	70	36
50	100	97	74	39
51	100	98	77	43
52	100	98	79	47
53	100	99	82	51
54	100	99	84	55
55	100	99	87	58

Table A-4. (Continued)

EST score	Percent chance of AFQT percentile being at least 21	31	50	65
56	100	99	89	62
57	100	100	90	66
58	100	100	92	69
59	100	100	93	73
60	100	100	95	76
61	100	100	96	79
62	100	100	96	82
63	100	100	97	84
64	100	100	98	87
65	100	100	98	89