AD-A05	3 302 SIFIED	APPLIED SCIENCE ASSOCIATES INC VALENCIA PA JET ENGINE MECHANIC-AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TE-ETC(U) DEC 77 J K HAWLEY, C J MULLINS, J WEEKS AFHRL-TR-77-73 NL							/			
	OF AD A053302							lander (
	EMI	- Manazaran					Provinceptor					
					enterina interina generation interina generation interina interinterina interina int		nenssaer Thirtig Nensser Nensser		- 12722 - 1272 - 1272 - 1272 - 1272 - 1272		1112	
					END DATE FILMED 6-78 DDC							
	e Siran		-									
1.												/



NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Applied Science Associates, Inc., Box 158, Valencia, Pennsylvania 16059, under contract F33615-76-C-0057, project 7719, with Personnel Research Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Mr. Joseph Weeks, Manpower Development and Evaluation Branch, was the contract monitor.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

LELAND D. BROKAW, Technical Director Personnel Research Division

DAN D. FULGHAM, Colonel, USAF Commander

	DT DOCUMENTATION DAS	E I	READ INSTRUCTIONS	1
19 REPU	RT DUCUMENTATION PAG		BEFORE COMPLETING FORM	
FHRI TR.77.73	2. 60	SVI ACCESSION NO.	S. RECIPIENT S CATALOG NUMBER	
		6	TYPE OF REPORT & PLOD COVERED	
ET ENGINE MECHA	NIC-AESC 426X2:	(7)	DFinal rest.	
XPERIMENTAL JOI	BPERFORMANCE TESTS .	7	13	
			6. PERFORMING ORG. REPORT NUMBER	
AUTHOR(.)			8. CONTRACT OR GRANT NUMBER(s)	
ohn K! Hawley		(is)	Francis an adden Them	
oseph/Weeks			F33615-76-C0057	
PERFORMING ORGANI	ZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK	
Applied Science Asso	ciates, Inc.	/	AREA & WORK UNIT MAREAS	
lox 158 Valencia Pennsylvania	16059	(IC	62703F	
CONTROL LING OFFIC		10		
IQ Air Force Human	Resources Laboratory (AFSC)	(1)	Dec. 177 12 73	.7
rooks Air Force Base	, Texas 78235	Ċ	13. NUMBER OF AGE	Ľ
MONITOPING ACENC	Y NAME & ADDRESSIL different for	Controlling Office)	15. SECURITY CLASS. (of this report)	
ersonnel Research Div	vision			
ir Force Human Reso	Durces Laboratory		Unclassified	
brooks All Force base	, Texas 78233		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
DISTRIBUTION STATE	EMENT (of this Report)			
DISTRIBUTION STATE	EMENT (of the abstract entered in Bi	ock 20, if different from	m Report)	•
DISTRIBUTION STATE	EMENT (of the abstract entered in Bl	ock 20, il dillerent froi	m Report)	•
DISTRIBUTION STATE	EMENT (of the abstract entered in Bl	ock 20, if different from	m Report)	
DISTRIBUTION STATE	EMENT (of the abstract entered in Blue abstract entered in Blue abstract entered in Blue abstract entered in Bl	ock 20, if different from	m Report)	
DISTRIBUTION STATE	EMENT (of the abstract entered in Blue bottom and the abstract entered in Blue bottom	ock 20, il dillerent fro	m Report)	
DISTRIBUTION STATE	EMENT (of the abstract entered in Blue bottom and the abstract entered in Blue bottom and the bo	ock 20, il dilferent fro	m Report)	
DISTRIBUTION STATE	EMENT (of the abstract entered in Blu DTES on reverse side if necessary and iden itam analyzia	ntify by block number)	tion technical school grades	
DISTRIBUTION STATE SUPPLEMENTARY NO KEY WORDS (Continue SVAB oefficient alpha	OTES on reverse side if necessary and iden item analysis item difficulty	ntily by block number) multiple correla phi-coefficient	tion technical school grades validity	
DISTRIBUTION STATE SUPPLEMENTARY NO KEY WORDS (Continue SVAB pefficient alpha prrelation	on reverse side if necessary and iden item analysis item difficulty item variance	ntily by block number) multiple correla phi-coefficient reliability	tion technical school grades validity	
DISTRIBUTION STATE SUPPLEMENTARY NC KEY WORDS (Continue SVAB pefficient alpha prrelation riterion test ross-validation	OTES on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20	ntify by block number) multiple correla phi-coefficient reliability skill level stepwise regress	m Report) tion technical school grades validity	
DISTRIBUTION STATE SUPPLEMENTARY NO KEY WORDS (Continue SVAB oefficient alpha orrelation riterion test ross-validation	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20	ntily by block number) multiple correla phi-coefficient reliability skill level stepwise regress	n Report) tion technical school grades validity	
DISTRIBUTION STATE SUPPLEMENTARY NC KEY WORDS (Continue SVAB officient alpha orrelation riterion test ross-validation Historically, the	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side if necessary and iden Air Force has used technical sch	ntify by block number) multiple correla phi-coefficient reliability skill level stepwise regress tify by block number) nool grades (TSGs)	tion technical school grades validity	
DISTRIBUTION STATE SUPPL EMENTARY NC SUPPL EMENTARY NC SVAB oefficient alpha orrelation riterion test ross-validation Historically, the he current study was	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side If necessary and iden Air Force has used technical sch to develop a job-related criterio	ntily by block number) multiple correla phi-coefficient reliability skill level stepwise regress ntily by block number) nool grades (TSGs) n metric against wh	tion technical school grades validity ion to validate aptitude tests. The purpose of the to validate aptitude measures. Along	
SUPPLEMENTARY NO SUPPLEMENTARY NO SUPPLEMENTARY NO SVAB officient alpha orrelation riterion test ross-validation MISTRACT (Continue Historically, the he current study was his line, three Criteri eveloped, ANOVA re	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side If necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean	ntily by block number) multiple correla phi-coefficient reliability skill level stepwise regress ntily by block number) nool grades (TSGs) n metric against wh d 7 skill level of P. Criterion Test score	tion technical school grades validity ion to validate aptitude tests. The purpose of ich to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were s were significantly different across the	
DISTRIBUTION STATE SUPPLEMENTARY NC SUPPLEMENTARY NC SVAB oefficient alpha orrelation riterion test ross-validation MSTRACT (Continue Historically, the he current study was his line, three Criteri eveloped. ANOVA re elevant skill levels. Sto	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side if necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean 0 epwise regression results indicated	ntily by block number) multiple correla phi-coefficient reliability skill level stepwise regress tily by block number) nool grades (TSGs) n metric against wh d 7 skill level of P. Criterion Test score i that Armed Servic	tion technical school grades validity ion to validate aptitude tests. The purpose of the to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were as were significantly different across the es Vocational Aptitude Battery (ASVAB)	
DISTRIBUTION STATE SUPPLEMENTARY NO SUPPLEMENTARY NO SVAB pefficient alpha porrelation riterion test ross-validation Historically, the he current study was his line, three Criteri eveloped. ANOVA ro elevant skill levels. Stu formation subscales	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side II necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean 0 epwise regression results indicated were most predictive of Criterio	ntily by block number) multiple correla phi-coefficient reliability skill level stepwise regress multiple stepwise regress multiple correla phi-coefficient reliability skill level stepwise regress multiple correla phi-coefficient reliability skill level stepwise regress multiple correla stepwise regress stepwise regress multiple correla stepwise regress stepwise regress stepwis	tion technical school grades validity ion to validate aptitude tests. The purpose of the to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were es were significantly different across the es Vocational Aptitude Battery (ASVAB) e for experienced mechanics. Considering a subeales were meet related to Critarion	
SUPPLEMENTARY NO SUPPLEMENTARY NO SUPPLEMENTARY NO SVAB Defficient alpha Dorrelation riterion test ross-validation MISTRACT (Continue Historically, the ne current study was his line, three Criteri eveloped. ANOVA ro elevant skill levels. Sta formation subscales nly basic airmen, the est performance. Wh	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side If necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean of epwise regression results indicated were most predictive of Criterio ASVAB information measures an en TSGs were regressed on ASV	ntilly by block number) multiple correla phi-coefficient reliability skill level stepwise regress tilly by block number) nool grades (TSGs) n metric against wh id 7 skill level of PJ Criterion Test score i that Armed Servic on Test performance d general knowledg 'AB scores, general	tion technical school grades validity ion to validate aptitude tests. The purpose of tich to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were s were significantly different across the es Vocational Aptitude Battery (ASVAB) e for experienced mechanics. Considering e subscales were most related to Criterion knowledge scales were most consistently	
DISTRIBUTION STATE SUPPLEMENTARY NO SUPPLEMENTARY NO SVAB Defficient alpha Dirrelation riterion test ross-validation BSTRACT (Continue Historically, the ne current study was his line, three Criteri eveloped. ANOVA re levant skill levels. Sta formation subscales nly basic airmen, the est performance. Wh	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side If necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean 0 epwise regression results indicated were most predictive of Criteric ASVAB information measures an en TSGs were regressed on ASV	ntify by block number) multiple correla phi-coefficient reliability skill level stepwise regress tify by block number) nool grades (TSGs) n metric against wh d 7 skill level of P. Criterion Test score i that Armed Servic on Test performanc d general knowledg 'AB scores, general	tion technical school grades validity ion to validate aptitude tests. The purpose of hich to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were as were significantly different across the es Vocational Aptitude Battery (ASVAB) e for experienced mechanics. Considering e subscales were most related to Criterion knowledge scales were most consistently	-70
SUPPLEMENTARY NO SUPPLEMENTARY NO SUPPLEMENTARY NO SVAB Defficient alpha Direlation fiterion test ross-validation WestRACT (Continue Historically, the ne current study was his line, three Criteri eveloped. ANOVA ro levant skill levels. Sta formation subscales aly basic airmen, the est performance. Wh	on reverse side if necessary and iden item analysis item difficulty item variance Job Performance Test KR20 on reverse side If necessary and iden Air Force has used technical sch to develop a job-related criterio on Tests relevant to the 3, 5, an esults demonstrated that mean of epwise regression results indicated were most predictive of Criterio ASVAB information measures an en TSGs were regressed on ASV DITION OF 1 NOV 65 IS OBSOLETE	ntilly by block number) multiple correla phi-coefficient reliability skill level stepwise regress tilly by block number) nool grades (TSGs) n metric against wh nd 7 skill level of P. Criterion Test scorre l that Armed Servic on Test performance d general knowledg 'AB scores, general	n Report) tion technical school grades validity to validate aptitude tests. The purpose of tich to validate aptitude measures. Along AFSC 426X2-jet engine mechanic-were as were significantly different across the es Vocational Aptitude Battery (ASVAB) e for experienced mechanics. Considering e subscales were most related to Criterion knowledge scales were most consistently Unclassified	-70

Uncl	assified	
_		

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20 Continued:

Predictive of technical school performance. The regression of TSGs on Criterion Test scores indicated that only Test 1 was generally related to technical school performance.

. . . .

NTIS DDC	White Section
UNANNOUNCE JUSTIFICATION	D CI
DISTRIBUTION	AVAILABILITY CODES
A	nithe sector of period of the sector of sector of sector

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

TABLE OF CONTENTS

	Page				
INTRODUCTION	1				
RESEARCH METHODOLOGY	4				
Task Analysis and Item Pool Development Data Collection	4 6				
RESULTS	9				
Item Selection Discriminability and Reliability of the	9				
Final Versions of the Criterion Tests	17				
VALIDITY STUDIES	25				
The Relationship of the ASVAB and the Criterion Tests The ASVAB Versus Technical School Grades The Relationship Between the Criterion Tests and TSGs Cross-Validation					
CONCLUSIONS	41				
Summary Discussion	41 42				
APPENDIX 1 ITEM ANALYSIS OF TEST 1 FOR 3-LEVEL AND 1-LEVEL PERSONNEL	43				
APPENDIX 2 ITEM ANALYSIS OF TEST 2 FOR 5-LEVEL AND 3-LEVEL PERSONNEL	51				
APPENDIX 3 ITEM ANALYSIS OF TEST 3 FOR 7-LEVEL AND 5-LEVEL PERSONNEL	59				

i

LIST OF TABLES

Title

No.

and the second of the second o

Page

. 1	Task Domains of AFSC 426X2	4
2	Planned Frequencies of Subjects Within PAFSC 426X2	7
3	Obtained Frequencies of Subjects Within PAFSC 426X2	'
	Across Skill Levels at Selected Air Force Bases	7
4	Scores on the Three Criterion Tests by Skill Level	9
5	Partial ANOVA Results for the Three 150-Item	
	Criterion Tests	10
6	Items Retained in Criterion Test 1 With	
	Corresponding Difficulty Levels	14
7	Items Retained in Criterion Test 2 With	
	Corresponding Difficulty Levels	15
8	Items Retained in Criterion Test 3 With	
	Corresponding Difficulty Levels	16
9	Means and Standard Deviations of Total Scores on	
	Tests 1, 2, 3, and Pooled Test Score by Skill Level	17
10	Univariate and Multivariate Analysis of Variance Results	
	on the Three Criterion Test Scores and Their Sum	18
11	Correlations Between Test Scores and Associated	
	Skill Level for All Subjects	19
12	Reliability Coefficients on Tests 1, 2, and 3	
10	Classified by Skill Level	19
13	Item-Total Score Correlations for Final 100 Items	22
1/	Un lest 1	22
14	On Test 2	22
15	Item-Total Score Correlations for Final 100 Items	25
13	On Test 3	24
16	ASVAR Means and Standard Deviations by Skill Level	24
17	Correlations of ASVAB Scales Pooled Over Skill Levels	27
18	Correlations of ASVAB Scales with Criterion Test	
10	Scores by Skill Level	28
19	Summary of Stepwise Regression Analyses Involving	
	Criterion Test Scores and ASVAB Scales	29
20	Means and Standard Deviations of Technical School	
	Grades by Skill Level	34
21	Correlations of ASVAB Scales with Technical School	
	Grades by Skill Level	35
22	Summary of Stepwise Regression Results Using Technical	
	School Grades as the Criterion and ASVAB Scales as	
	Predictors by Skill Level	36
23	Correlations Between Criterion Test Scores and	
	Technical School Grades by Skill Level	37

LIST OF TABLES (continued)

<u>No</u> .	Title	Page
24	Stepwise Regression Results Using Technical School Grades as Criteria and Criterion Test Scores	
	as Predictors	38
25	Correlations Between Actual and Predicted Scores in	
	the Cross-Validation Subsample by Skill Level	39

iii

JET ENGINE MECHANIC-AFSC 426X2: EXPERIMENTAL JOB PERFORMANCE TESTS

INTRODUCTION

Historically, the Air Force has used technical school grades (TSGs) as a criterion for validating aptitude measures. However, TSGs have not proven to be a completely acceptable criterion. One of the primary reasons for the unacceptability of TSGs is the weak relationship between school grades and later job performance.

The purpose of this study was to develop a job-related criterion metric against which aptitude tests could be validated. In line with this objective, it was desired that the new criterion metric be more indicative of on-the-job performance than are TSGs. A constraint, imposed by the high developmental and administrative costs of job sample tests, was that the new criterion metric be a paper-and-pencil measure of job performance.

The use of paper-and-pencil measures usually entails lower fidelity in simulating job tasks than do actual job performance items. This decrement usually results from the fact that paper-and-pencil tests do not cover a sufficient number of job-relevant tasks. An approach to the problem of task coverage is to concentrate upon comprehensively identifying job components through task analysis. A second problem with paper-and-pencil items is that they often do not cover the complete range of skills within a particular task category. To insure fidelity, it is necessary that all task categories be adequately sampled, and that the complete range of skills within categories be covered. Each item on the paper-and-pencil measure then maps to a specific task category and skill level within the task category.

Once the relevant job skills have been identified and suitable paper-and-pencil measures developed, a link must still be established between job knowledge, as assessed on the criterion measure, and actual job performance. The current study makes the assumption that career progression is a direct measure of job performance. That is, the selection processes of the Air Force tend to promote job incumbents who successfully meet their job requirements. Similarly, when an incumbent ceases to satisfactorily perform his duties, promotion stops or slows down.

Following the assumption that career progression reflects job performance, the current study makes the further assumption that, to the extent test performance is associated with career level, the Criterion Tests (CTs) may be considered as indices of actual job proficiency.

Employing an indirect measure of job proficiency such as career progression does, however, entails a certain amount of risk. A primary element of risk is that there may be less relationship between job proficiency and career progression than initially assumed. The factors which account for career progression may only indirectly involve job knowledge. The Criterion Test items may, indeed, reflect job knowledge, but fail to discriminate skill groups. However, if Criterion Tests differentiate various skill groups, they then may be said to at least be a measure of "job success," with job success being defined as long-term job incumbency.

Given the foregoing objectives and assumptions, the Criterion Test development effort proceeds as follows. First, the Air Force Occupational Survey (AFOS) is used to identify the content of Air Force Specialty Code (AFSC) 426X2, Jet Engine Mechanic (formerly 432X0). The task responsibilities of 3, 5, and 7 skill level personnel in this AFSC are separately identified. Finally, a large pool of potential paper-and-pencil test items (70 percent more than planned) are designed in such a way as to discriminate 1 and 3, 3 and 5, and 5 and 7 skill level job incumbents. The ability of the CTs to discriminate career progression levels is the link between test performance and job proficiency. An initial editing and screening of the item pool to remove defective items results in tryout versions of the Criterion Tests.

The next step in the process is to administer the CTs to a large group of subjects representative of each skill level in the target AFSC. Results from this administration are used to select the "best" CTs. Best, in this context, refers to CTs made up of those items having the highest discriminating power, the most homogeneity, and the most acceptable distribution of responses to the distractors.

After defective items have been removed, and the final versions of the CTs assembled, the study proceeds with an investigation of test discriminability and reliability. As previously discussed, the ability of the Criterion Test scores to discriminate career levels links the CT results with job proficiency. It is also desirable to demonstrate that the responses of subjects to the CT items are consistent.

The last phase of the study involves the relationships of the Criterion Tests and the Armed Services Vocational Aptitude Battery (ASVAB), the standard Air Force aptitude measure used in the selection and classification of enlisted personnel. ASVAB-Criterion Test relationships are of interest since selection validity hinges upon being able to predict test performance, and, thus, indirectly, job proficiency, from ASVAB scores. The relationships between TSGs and the ASVAB scales permit a comparison of the present results with previous validity studies. ASVAB-TSG associations, in comparison with ASVAB-Criterion Test correlations, also give some indication of how changing the criterion for selection will possibly affect the composition of the pool of jet engine mechanics. Finally, should Criterion Test results be associated with career progression, the relationship of TSGs with the Criterion Tests will give an indication of how technical school performance relates to eventual job proficiency. Each of these issues is discussed in further detail in the next two sections. Research Methodology covers the development and evaluation of individual test items. The discriminability and reliability of the tests is addressed in the Results section. The Results section and the Validity Studies section address the relationships among the Criterion Tests, the ASVAB, and technical school grades.

RESEARCH METHODOLOGY

Task Analysis and Item Pool Development

The first step in developing a pool of paper-and-pencil job performance items was to systematically identify the separate task components of AFSC 426X2 using the AFOS. Various job components were then grouped into broad categories of activities called task domains. The task domains identified for each skill level are presented in Table 1.

TABLE 1 TASK DOMAINS OF ASFC 426X2

Level-3

- 1. Safety
- 2. Tools and Equipment
- 3. Inspections
- 4. Technical Orders, Forms, Maintenance Levels
- 5. Security
- 6. Shipping, Storage, Preservation
- 7. Jet Engine Theory
- 8. Jet Engine Operation and Maintenance

Level-5

- 1. Safety
- 2. Tools and Equipment
- 3. Inspections
- 4. Technical Orders, Forms, Maintenance Levels
- 5. Security
- 6. Shipping, Storage, Preservation
- 7. Jet Engine Theory
- 8. Jet Engine Operation and Maintenance
- 9. Training
- 10. Work Performance
- 11. Non-Work Personnel Issues
- 12. Supply

TABLE 1 (continued)

Level-7

- 1. Safety
- 2. Tools and Equipment
- 3. Inspections
- 4. Technical Orders, Forms, Maintenance Levels
- 5. Security
- 6. Shipping, Storage, Preservation
- 7. Jet Engine Theory
- 8. Jet Engine Operation and Maintenance
- 9. Work Performance
- 10. Non-Work Personnel Issues
- 11. Supply
- 12. Evaluation
- 13. Planning
- 14. Supervision
- 15. Training

As may be noted, the task domains for each of the three skill levels have some, but not complete, overlap. New domain elements are added as the skill level of job incumbents increases. For example, "Training" is a domain element for level 5, but not level 3. The fact that a particular domain element is on more than one list does not mean, however, that test items representing that element are the same across different skill levels. Clearly, the activities involved in a job element like "Safety" are different for a level 3 mechanic and a level 7 noncommissioned officer (NCO).

Following identification of the relevant task domains, job charts, or systematic matrices of skill components which make up the job of jet engine mechanic, were prepared. The job matrices were compiled separately for each skill level. Specifically, the job matrices identified:

- 1. The tasks done by most job incumbents at that skill level.
- 2. The tools and equipment used.
- The knowledge, skills and judgments needed to perform the task.
- Tasks generally <u>not</u> performed by mechanics at the next lower skill level, or, if performed at the lower level, done in a way requiring different skills, knowledge, or judgments.

Using the information contained in the job analysis matrices, a large pool of potential test items was then developed. Guidelines for item development were as follows:

- Items should measure job performance on tasks done by most mechanics at a particular skill level (e.g., 3, 5, and 7).
- Items should discriminate between the performance of mechanics at one skill level and the performance of those at the next lowest skill level, e.g., 1 versus 3, 3 versus 5, and 5 versus 7.

The unrefined item pool was next submitted to the Air Force for final approval. Test items were then edited and screened to eliminate obvious defectives. The pre-screening process resulted in a pool of 450 items (150 in each test for inclusion in the tryout versions of the Criterion Tests).

It should be noted at this point that the job performance Criterion Tests developed in the current study do not have the same objective as the Specialty Knowledge Tests (SKTs) used in the Weighted Airman Promotion System (WAPS). The SKTs are developed on the basis of training documentation and represent training standards of performance. On the other hand, the Criterion Tests were designed to reflect on-the-job performance--that is, job performance as it actually occurs in the field.

Data Collection

Following Air Force approval of the final sets of test items, and the preparation and printing of the testing materials, data collection began at 12 Air Force bases. The preliminary data collection plan is shown in Table 2. More subjects than actually required were scheduled at each base. Over-scheduling was done to serve as a precaution against collecting unusable data, or failing to obtain the expected number of subjects at some bases. The numbers planned for each base were a function of the actual number of each skill level reported to be available. Testing on each group of subjects was scheduled to be completed during one working day. To control for possible fatigue and order effects, the order of administration of the various tests was counterbalanced.

Air Force Base	3	5	7	Totals
Seymour Johnson	30	25	35	90
McGuire	25	40	40	105
Dover	35	85	42	162
Little Rock	38	20	32	90
Barksdale	28	18	26	72
Columbus	28	5	28	61
Nellis	25	35	12	72
Travis	80	80	40	200
George	26	9	37	72
Castle	28	70	64	162
Mountain Home	28	28	34	90
Holloman	48	15	45	108
Totals:	419	430	435	1284

TABLE 2PLANNED FREQUENCIES OF SUBJECTS WITHINPAFSC 426X2 ACROSS SKILL LEVELS ATSELECTED AIR FORCE BASES

As expected, the actual number of subjects obtained was lower than the number sought. The actual numbers of subjects obtained in each skill level at each Air Force base are presented in Table 3.

TABLE 3OBTAINED FREQUENCIES OF SUBJECTS WITHIN
PAFSC 426X2 ACROSS SKILL LEVELS AT
SELECTED AIR FORCE BASES

	S			
Air Force Base	3	5	7	Totals
Seymour Johnson	0	26	9	35
McGuire	21	20	7	48
Dover	55	72	50	177
Little Rock	29	35	38	102
Barksdale	22	6	21	49
Columbus	9	1	17	27
Nellis	6	31	18	55
Travis	4	46	19	69
George	5	15	15	35
Castle	22	34	30	86
Mountain Home	9	30	29	68
Holloman	12	20	30	62
Totals:	194	336	283	813

7

During approximately the same time period that job performance data was being collected from active duty jet engine mechanics, the Air Force was also administering each Criterion Test and the ASVAB to 422 basic airmen (level 1 personnel). Results from 419 level 1 subjects were eventually usable. The complete test data file thus contains scores from 1232 subjects--813 experienced mechanics and 419 basic airmen.

RESULTS

Item Selection

Following the administration of the Criterion Tests and the ASVAB, individual results were assembled and prepared for analysis. The first step in the data analysis was to obtain each subject's score on each of the Criterion Tests and the nine ASVAB scales. The Criterion Test scores were corrected for guessing using the standard formula used with multiplechoice tests. The resulting means, standard deviations, 25th, and 75th percentile scores for each group on the three Criterion Tests are given in Table 4.

TABLE 4 SCORES ON THE THREE CRITERION TESTS BY SKILL LEVEL

Criterion Test	Skill Level	Mean	Standard Deviation	25th Percentile	75th Percentile
1	1	9.41	8.91	3	14
1	3	75.15	21.86	63	90
1	5	72.32	22.25	60	88
1	7	93.57	16.03	86	103
2	1	7.46	8.38	0	12
2	3	21.48	14.04	12	31
2	5	28.13	15.21	18	38
2	7	54.78	17.48	43	66
3	1	3.57	6.94	-2	7
3	3	10.35	7.86	4	15
3	5	16.51	11.69	8	23
3	7	38.00	14.19	27	47

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Referring to Table 4, the mean test scores on all three Criterion Tests generally increase as a function of skill level. One exception is CT 1, where level 5 personnel do not score as high on the average as level 3 personnel. The quartile scores for each group indicate that the distributions of test scores are nearly symmetric about individual group means. An analysis of variance (ANOVA) performed on the three test scores across skill levels indicates that the group means are significantly different. Partial ANOVA results are presented in Table 5.

TABLE 5PARTIAL ANOVA RESULTS FOR THE THREE150-ITEM CRITERION TESTS

Test	<u>F-Statistic</u>	Degrees of Freedom	<u>P</u>
1	1654.22	3, 1228	<.0001
2	679.83	3, 1228	<.0001
3	632.11	3, 1228	<.0001

(Total N = 1, 232)

The next step in the test development effort was to reduce the number of items in each Criterion Test by selecting the 100 "best" items in each test. "Best," in this context, refers to those items which most closely meet the following criteria:

- The item discriminates between the appropriate skill levels. For example, good items in CT 1-the criterion for skill level 3--should be answered correctly more often by level 3 personnel than by basic airmen (level 1 personnel).
- 2. An item should be correlated with the total score. Items having a high correlation with the total score more consistently measure the job proficiency aspects on which the whole test discriminates skill level groups. Job proficiency is considered to be a unidimensional, global construct.
- 3. The distribution of answers to the item distractors should be satisfactory. This last check will identify items in which one or more response alternatives are rarely selected by the testees. Additionally, tabulating answer distributions will identify items where some incorrect response alternatives are marked more often by high scorers than by low scorers. Items having these characteristics are defective and should be rejected.

In applying the above selection rules, primary importance was given to the discriminating power of items. The rationale for this choice is based upon the fact that the overall purpose of the tests is to serve as criteria for their associated skill level. Thus, tests that do not discriminate between career levels are not suited for use as criteria. If the items do not discriminate career levels, then the link between the tests and job proficiency is broken. Selection rules 2 and 3 were included to lend support to the primary criterion for selection. These rules were to be used in the event that item discrimination alone did not provide sufficient information for deciding whether to retain or reject an item.

Items on the three Criterion Tests were next submitted to an item analysis and the following statistics obtained:

- 1. The index of discrimination (phi).
- 2. The rank of the discrimination index.
- 3. The item difficulty for the group relevant to the test.
- 4. The item variance for the group relevant to the test.
- 5. The item-total score correlation for the group relevant to the test.
- 6. The rank of the item-total score correlation.
- 7. The item difficulty for the skill group immediately lower than the criterion group. This is denoted the "non-relevant 'group.'"
- 8. The item variance for the non-relevant group.
- 9. The item-total score correlation for the nonrelevant group.
- The average rank of the discrimination index and the item-total score correlation for the criterion group.

The complete lists of item statistics for each of the tests are presented in Appendices 1, 2, and 3. Items are listed in order of their discriminating power, as indicated by phi. Also included in the last column of each appendix is a table entry indicating whether the item was selected for inclusion in the final versions of the Criterion Tests. A "Yes" or a "No" indicates this choice.

Referring to Appendix 1, which contains the item analysis statistics for CT 1, the indices of discrimination (phi coefficients) computed across items for skill levels 1 and 3 range from 0.64 to 0.17. The corresponding item difficulties indicate that 60-90 percent of level 3 personnel generally answer the items correctly. These difficulty levels indicate a desirable situation, since CT 1 is designed to differentiate level 3 from level 1 personnel. The corresponding item-total score correlations for the retained items on CT 1 range from 0.49 to 0.14. None of the retained items in CT 1 has a negative correlation with the total score.

The item analysis statistics on the second test (see Appendix 2) indicate that CT 2 does a relatively poor job of discriminating level 5 from level 3 personnel. The discrimination coefficients for the two relevant groups (skill levels 3 and 5) range from 0.21 to -0.23. For retained items, the range is from 0.21 to .01. Given the sample size involved (N = 530), a value of at least 0.085 is required for a correlation coefficient to be considered significantly different from zero at the p=.05 level of significance. Using the p=.05 criterion, only 31 test items on CT 2 yield a discrimination index which is individually significantly different from zero.

The reason for the low discriminating power of items in CT 2 is apparent when considering the item difficulties for the two relevant groups. The problem is that most items on CT 2 are generally too difficult for level 5 personnel. In most cases, level 5 personnel answer the items correctly more often than level 3 personnel. However, the magnitude of the difference is not large enough to reliably differentiate the two groups. There is too much overlap in the response distributions. The level of response overlap is evident from the quartile scores presented in Table 4. Here, it is noted that more than 25 percent of level 3 personnel score higher on CT 2 than the mean score of the level 5 personnel.

CT 3 presents a situation midway between the relatively good characteristics of CT 1 and the poor characteristics of CT 2. On CT 3, the discrimination coefficients for retained items range from 0.48 to .01. Again, given the sample size involved (N = 619), a correlation of approximately 0.08 is required for significance at the p=.05 level. Using that criterion, 85 items on CT 3 individually yield significant indices of discrimination. Looking at the corresponding item difficulty statistics for items with $\phi < .08$, it is again the case that items on CT 3 having low discriminating power suffer from the same problem as the majority of items on CT 2: the items are generally too difficult for the criterion skill level personnel.

In summary, Criterion Tests 2 and 3 do a less satisfactory job of differentiating their relevant skill groups than does CT 1. On both CTs 2 and 3, the major problem concerns items that are too difficult for the criterion skill groups. This result may indicate that differentiating within skill levels of jet engine mechanics is much more difficult than simply differentiating mechanics from non-mechanics. One of the difficulties encountered in the item development phase of the study was deciding specifically which task elements separated level 3 jobs from level 5 jobs, and level 5 jobs from level 7 jobs. There were no clearly defined criteria. In most cases, job content differences across skill levels were a matter of the degree of the skills involved. This overlap in the job content of the 3, 5, and 7 level skill groups is reflected in their respective performance on the Criterion Tests.

To recap the item selection phase of the study, Tables 6, 7, and 8 present sequential lists of all retained Criterion Test items and the corresponding difficulty indices for all skill levels.

TABLE 6 ITEMS RETAINED IN CRITERION TEST 1 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		SI	ci11	Leve	1	Pooled			SI	ci11	Leve	e1	Pooled
	Item	1	3	5	7	3-5-7		Item	1	3	5	7	3-5-7
1	2	.24	.79	.62	.74	.70	51	81	.26	.93	.94	. 98	.95
2	3	. 27	.77	.69	.91	.79	52	82	.43	.89	.88	.96	.91
3	6	.15	.45	.43	.77	.55	53	83	.42	.81	.85	.92	.86
4	8	.24	.87	.82	.91	.86	54	87	.27	.83	.88	.97	.90
5	9	.21	.85	.87	.95	.90	55	89	.24	.84	.80	.89	.84
6	10	.19	.81	.84	.89	.85	56	90	.23	.86	.89	.97	.91
7	14	.18	.49	.54	.72	. 59	57	91	.44	.93	.93	.99	.95
8	15	.20	.71	.61	.92	.74	58	92	.23	.61	.62	.86	.70
9	17	.27	.86	.90	.87	.88	59	93	.27	. 59	.52	.51	.53
10	18	.54	.95	.92	.97	.95	60	94	.30	.65	.70	.89	.76
11	21	.52	.91	.85	.95	.90	61	95	.25	.91	.90	.98	.93
12	23	.14	.64	.66	.88	.73	62	96	.54	.86	.87	.93	.89
13	26	.41	.60	.61	.93	.72	63	97	. 29	.65	.70	.87	.75
14	27	.16	.54	.64	.78	.66	64	98	.21	.64	.54	.65	.60
15	28	.40	. 88	.91	.87	.89	65	99	. 31	.89	.82	.94	.88
16	31	.14	.36	.43	.83	.55	66	101	.19	.52	.40	.44	.44
17	32	.26	.80	.80	.85	.82	67	102	.41	.87	.79	.86	.83
18	34	.33	.91	.90	.98	.93	68	104	.36	.79	.79	.88	.82
19	35	.18	.79	.80	.86	.82	69	105	.22	.41	.48	.80	.57
20	36	.16	.74	. 79	.83	.79	70	106	.27	.68	.63	.72	.68
21	37	.18	. 66	.70	.92	.77	71	108	.18	.54	.36	.45	44
22	40	. 62	.95	.93	.96	.95	72	110	.19	.47	.31	. 52	.42
23	41	. 25	.85	. 87	.95	.90	73	111	. 26	.75	. 69	.79	74
24	42	. 30	.81	.84	.93	.87	74	112	.23	.72	.68	.81	74
25	43	. 38	. 81	. 80	. 88	.83	75	114	.41	.60	.73	.87	.75
26	44	. 32	. 70	.64	.73	. 69	76	115	. 62	.94	.87	.98	92
27	46	. 36	.75	.75	. 88	. 80	77	116	.17	.70	.68	. 57	65
28	47	. 11	. 29	. 19	.82	. 43	78	117	. 34	.77	.84	.89	84
29	48	. 37	. 89	.91	.95	.92	79	119	. 26	.54	.48	. 57	53
30	49	. 36	. 71	. 68	.78	.72	80	120	.23	.74	.71	.92	79
31	52	.11	.61	. 54	. 59	. 58	81	121	. 21	. 53	. 33	.28	36
32	54	. 41	.93	. 87	.97	.92	82	123	. 26	.70	. 61	.77	69
33	55	. 24	.72	.75	.84	. 78	83	124	.23	. 58	. 62	.83	69
34	58	.33	.74	.65	.79	.72	84	126	. 21	.81	.80	.90	84
35	59	. 39	.89	.82	.90	.86	85	128	. 48	.80	.78	.96	85
36	61	.24	. 66	.77	.94	. 80	86	129	.11	. 60	.65	.85	71
37	62	.13	.71	.66	.78	.00	87	131	. 17	.63	.39	.52	49
38	63	59	96	. 80	98	94	88	132	. 30	.75	.67	84	75
39	64	30	.81	.83	.95	.94	89	134	. 22	. 59	. 38	. 22	38
40	65	34	.76	.78	.80	78	90	135	. 25	. 47	.43	. 49	46
41	66	42	70	60	56	61	91	136	10	46	52	56	52
42	68	35	76	78	. 50	.01	92	137	18	.40	52		60
43	69	25	79	.70	05	.01	93	138	.27	56	62	.05	63
44	70	38	.84	. 78	. 75	. 78	94	140	.14	. 38	. 38	. 48	42
45	71	20	.04	.10		. 90	05	141	.17	76	. 30	.40	79
46	72	. 29	.00	.07		92	06	142	15	66	59	64	62
47	74	31	. 70	64	60	68	07	145	20	. 55	. 10	.62	55
47	75	. 31	. 72	.04	.09	.00	0.9	146	11	56	.49	.02	66
40	76	. 20	. /1	.02	.03	.05	00	149	24	. 10	.00	.01	.53
50	77	. 21	.00	.42	.42	.40	100	150	33	.40	. 50		
20	//	. 21	.00	.09	. 90	.92	100	100		.00	. 11	.13	./0

TABLE 7 ITEMS RETAINED IN CRITERION TEST 2 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		SI	ci11	Leve	el	Pooled				SI	kill	Leve	1	Pooled
	Item	1	3	5	7	3-5-7			Item	1	3	5	7	3-5-7
1	2	.13	.30	.36	.48	.39		51	75	.19	.21	.24	.34	.27
2	3	.11	.19	.31	.44	.32		52	76	.43	.80	.81	•95	.86
3	4	.23	.28	.31	.30	.30		53	77	.31	.34	.36	.49	.40
4	5	.25	. 39	.46	.66	.51		54	78	.24	.30	.43	•73	.50
5	7	.17	.08	.12	.12	.11		55	80	.36	.57	.76	.94	.78
6	9	. 29	.53	.67	.92	.72		56	82	.57	.66	.71	•89	.76
7	12	.35	.43	.51	.58	.51		57	83	.52	.59	.67	.70	.66
8	13	. 39	. 39	.44	.73	.53		58	84	.15	.18	•22	• 36	.26
9	14	. 27	.22	.29	.72	.42		59	85	.28	.17	•20	•41	•27
10	16	.20	.12	.19	.65	.33		60	86	.18	.58	.60	.72	•64
11	18	.16	.30	.37	.66	.46		61	88	.47	.55	.59	.74	.63
12	19	. 43	. 57	.59	.82	.67		62	89	.41	.60	.63	.78	.67
13	20	. 29	.18	.21	.40	.27		63	90	.32	.55	.56	.88	.67
14	22	. 36	.55	.67	.89	.72		64	92	.29	.27	.37	.27	.31
15	24	. 30	.27	.37	.67	.45		65	93	.30	.30	.35	.70	.46
16	25	. 38	.55	.64	.76	.66		66	99	.28	.40	.54	.80	.60
17	26	. 19	. 27	.29	.49	.36		67	100	.26	.35	.36	.66	.46
18	27	.12	.11	.21	. 39	.25		68	101	.16	.17	.18	.36	.24
19	29	. 32	. 56	.59	.59	.58		69	103	.26	.35	.42	.62	.47
20	30	.44	.61	.62	.73	.66		70	104	.27	.27	.40	.80	.51
21	31	.18	.61	.67	.91	.74		71	105	.37	.38	.42	.59	.47
22	32	. 23	.06	.08	.08	.08		72	106	.26	.30	.35	.37	.34
23	33	.45	.65	.72	.92	.77		73	107	.25	.26	.33	.26	.29
24	34	. 22	. 30	.38	.62	.44		74	108	.25	.32	.35	.45	.38
25	35	.18	.12	.14	.14	.14		75	109	.33	.38	.43	.62	.48
26	37	. 58	.76	.81	.92	.84		76	111	.30	.26	.30	.44	.34
27	38	.15	.45	.50	.67	.55		77	112	.38	.45	.52	.69	.56
28	39	. 47	.52	.62	.82	.66		78	114	.41	.64	.64	.71	.66
29	42	. 31	. 29	.33	.50	.38		79	116	.31	.49	.60	•70	.61
30	43	. 38	.63	.74	.88	.76		80	117	.34	.32	.35	.49	.39
31	45	. 19	. 09	.13	.27	.17		81	118	.24	.43	.45	.52	.47
32	46	. 11	. 19	.23	.63	.36		82	119	.23	.20	.21	.16	.19
33	47	. 29	. 20	.22	.43	.29		83	121	.38	.31	.40	.78	.51
34	48	. 29	. 36	.43	.66	.49		84	123	.17	.25	.41	.81	.51
35	50	. 44	.75	.82	.95	.85		85	124	.33	.55	.65	.75	.66
36	51	. 29	. 46	.55	.71	.58		86	127	.43	.31	.36	.32	.34
37	52	. 20	. 19	.22	.34	.26		87	129	.28	.37	.49	.58	.49
38	55	. 20	. 38	.43	.56	.46		88	130	.17	.24	.33	.57	.39
39	57	. 33	. 25	.34	.73	.46		89	131	.27	.20	.34	.59	.40
40	59	. 49	. 60	.63	.71	.65		90	133	.20	.17	.22	.26	.22
41	60	. 42	. 48	.61	.70	.61		91	134	.27	.59	.66	.76	.68
42	62	. 25	. 49	.63	.90	.69		92	135	.39	.51	.56	.93	.68
43	64	. 26	. 30	.40	.53	.42		93	136	.14	.19	.19	.42	.27
44	65	. 27	. 25	.44	.42	.39		94	138	.15	.29	.48	.75	.53
45	66	. 20	. 15	.18	.33	.23		95	139	.38	.28	.32	.34	.32
46	67	. 52	. 69	.68	.89	.76		96	142	.41	.41	.48	.52	.48
47	68	. 24	. 16	.22	.18	.19		97	144	.15	.17	.21	.24	.21
48	69	. 32	. 42	.46	.79	.56		98	145	.22	.23	.35	.62	.42
49	71	. 23	. 24	.29	.48	.35		99	147	.12	.24	.36	.40	.34
50	74	. 26	. 40	.44	.24	.36	1	100	149	.43	.36	.42	.73	.51

TABLE 8 ITEMS RETAINED IN CRITERION TEST 3 WITH CORRESPONDING DIFFICULTY LEVELS

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

		SI	kill	Leve	el	Pooled				SI	<i11< th=""><th>Leve</th><th>21</th><th>Pooled</th></i11<>	Leve	21	Pooled
	Item	1	3	5	7	3-5-7			Item	1	3	5	7	3-5-7
1	2	.27	.40	.57	.86	.63	5	1	68	.25	.61	.60	.80	.67
2	4	.35	.38	.34	.49	.40	5	2	69	.26	.23	.28	.42	.31
3	6	.29	.51	.49	.60	.53	5	3	72	.21	.19	.21	.45	.29
4	7	.28	.31	.36	.36	.35	5	4	73	.32	.38	.50	.71	.55
5	8	. 34	.32	.39	.59	.44	5	5	75	.32	.29	.32	.52	.38
6	9	.21	.22	.22	.28	.24	5	6	77	.29	.15	.15	.24	.18
7	10	.16	.15	.08	.20	.14	5	7	79	.35	.41	.39	.48	.43
8	11	.31	.20	.26	.47	.32	5	8	80	.20	.24	.30	.58	.38
9	12	.18	.22	.15	.27	.21	5	9	81	.26	.43	.63	.89	.67
10	13	.10	.09	.11	.24	.15	6	0	82	.26	.34	.37	.43	.38
11	14	.28	.33	.43	.65	.48	6	1	83	.28	.23	.32	.49	.36
12	15	.15	.16	.13	.17	.15	6	2	85	.13	.23	.18	.25	.22
13	16	.25	.30	.23	.33	.28	6	3	86	.23	.27	.31	.46	.35
14	17	.23	.26	.34	.70	.44	6	4	87	.24	.21	.36	.59	.40
15	18	.24	.20	.24	.70	.39	6	5	88	.16	.11	.12	.16	.13
16	19	.13	.18	.18	.48	.28	6	6	90	.17	.18	.28	.66	.39
17	20	.25	.52	.73	.94	.75	6	7	91	.27	.38	.45	.67	.51
18	21	.24	.27	.37	.59	.42	6	8	92	.16	.24	.21	.34	.26
19	22	.38	.27	.24	.46	.32	6	9	93	.29	.20	.31	.69	.41
20	24	.30	.44	.58	.65	.57	7	0	101	.25	.19	.26	.72	.40
21	25	.22	.24	.20	.66	.37	7	1	102	.53	.43	.49	.57	.50
22	27	.67	.79	.86	.97	.88	7	2	103	.20	.19	.27	.55	.35
23	28	.30	.48	.48	.61	.52	7	3	104	.23	.25	.39	.55	.41
24	29	.37	.53	.71	.89	.73	7	4	105	.31	.40	.49	.49	.47
25	31	.17	.19	.16	.44	.26	7	5	: 27	.17	.25	.22	.30	.26
26	32	.16	.32	.47	.54	.46	7	6	108	.41	.46	.57	.65	.57
27	33	.39	.48	.48	.48	.48	7	7	109	.18	.24	.25	.40	.30
28	34	.23	.25	.27	.36	.30	7	8	113	.30	.28	.43	.71	.49
29	35	.24	.26	.24	.65	.39	7	9	115	.27	.26	.26	.50	.34
30	39	.31	.34	.42	.72	.51	8	0	117	.16	.14	.17	.29	.21
31	40	.16	.11	.12	.41	.22	8	1	119	.28	.32	.34	.55	.41
32	41	.22	.34	.34	.51	.40	8	2	120	.25	.29	.27	.29	.28
33	43	.23	.38	.45	.72	.52	8	3	121	.42	.31	.33	.45	.37
34	44	.27	.20	.36	.67	.43	8	4	122	.19	.24	.39	.78	.49
35	45	.16	.45	.60	.93	.68	8	5	123	.45	.50	.64	.82	.67
36	46	.18	.15	.09	.13	.12	8	6	126	.32	.36	.55	.77	.58
37	47	.25	.15	.14	.45	.25	8	7	127	.27	.23	.26	.41	.31
38	49	.19	.17	.24	.35	.26	8	8	129	.15	.34	.43	.64	.48
39	51	.50	.57	.67	.81	.69	8	9	131	.33	.30	.31	.49	.37
40	52	.20	.32	.29	.39	.33	9	0	132	.27	.34	.36	.40	.37
41	53	.23	.15	.17	.38	.24	9	1	135	.32	.35	.42	.47	.42
42	54	.15	.15	.12	.20	.16	9	2	136	.66	.76	.81	.92	.84
43	55	.58	.66	.80	.89	.80	9	3	137	.13	.33	.23	.30	.28
44	56	.39	.53	.57	.71	,61	9	4	138	.26	.23	.34	.43	.35
45	57	.21	.12	.12	.52	.26	9	5	140	.29	. 27	.26	.35	.29
46	59	.31	.43	.54	.66	.55	9	6	143	.52	.60	.66	.75	.68
47	62	.28	.39	.52	.78	.58	9	7	144	.20	.12	.12	.27	.17
48	63	.26	.41	.37	.58	.46	9	8	145	.16	.18	.17	.30	.22
49	64	.29	.38	.52	.68	.54	9	9	146	. 38	. 29	.40	.60	.44
50	66	.42	.53	.63	.76	.65	10	0	149	. 31	.52	.54	.64	.57

16

Discriminability and Reliability of the Final Versions of the Criterion Tests

Following selection of the 100 best items in each of the three Criterion Tests, the tests were again scored using only the retained items. The means and standard deviations for each test, plus the total of the test scores, classified by skill level, are presented in Table 9. The test mean scores shown in Table 9 now have a maximum possible value of 100 rather than 150. The new test scores are also corrected for guessing.

TABLE 9 MEANS AND STANDARD DEVIATIONS OF TOTAL SCORES ON TESTS 1, 2, 3, AND POOLED TEST SCORE BY SKILL LEVEL (Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

			Test Score		
Skill Level		<u>1</u>	2	3	Total
1	М	3.98	5.34	3.13	12.45
	SD	7.48	6.56	5.78	13.40
3	м	61.95	15.35	8.54	85.84
	SD	17.89	11.70	7.18	29.56
5	м	59.06	24.00	17.78	100.84
	SD	17.84	13.10	11.10	35.23
7	м	73.46	45.62	38.57	157.65
	SD	12.27	13.65	14.24	34.07
Pooled 3-5-7	м	64.76	29.46	21.57	115.79
	SD	17.36	17.86	17.15	45.62

As was the case before final item selection, the total scores in Table 9 generally increase as a function of increasing skill level. This result holds over all three Criterion Tests. One exception is, again, CT 1 where level 3 personnel score slightly higher than level 5 personnel (61.95 versus 59.06). A multivariate analysis of variance (MANOVA) performed on the three individual test scores indicates that the group mean profiles are significantly different across skill levels (p < .0001). A corresponding univariate analysis of variance on the sum of the test scores was also highly significant (F=1547.71, with 3 and 1228 DF, p < .0001). The analysis of variance results are summarized in Table 10. TABLE 10. UNIVARIATE AND MULTIVARIATE ANALYSIS OF VARIANCE RESULTS ON THE THREE CRITERION TEST SCORES AND THEIR SUM

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

	Mult	ivariate				Univarié	ate				
				T		T ₂	2	T ₂		Tt	
Effect	Fa	df	러	۲	ର୍ଘ	E4	러	F1	ର୍ଘ	떠	ଧ
Main ^b	3018.67	6	.000	1844.29	.0001	750.86	.0001	749.86	.000	1547.71	.0001
3 vs 1	877.42	3,1226	.0001	2336.35	.0001	105.54	.000	38.96	.0001	888.78	.0001
5 vs 3	48.08	3,1226	.000	5.41	.019	73.18	.0001	48.05	.0001	22.02	.000
7 vs 5	316.90	3,1226	.0001	167.14	.000	570.66	.000	871.96	.000	683.89	.000

^aThe test statistic for the multivariate test of the main effect is an asymptotic chi-square developed from a likelihood-ratio criterion. The multivariate group contrasts are tested with an exact F developed from the same likelihood-ratio criterion.

^bThe tests of the main effect have 3 degrees of freedom, and the error terms have 1228 degrees of freedom in all cases.

1.8

The analysis of variance results indicate that the test scores do discriminate between personnel at each of the skill levels. Correlation coefficients computed between skill levels (e.g., 1, 3, 5, and 7) and scores on each of the three Criterion Tests and their sum are shown in Table 11. The correlations lend support to the analysis of variance results, and also give an idea of the degree of association between individual test performance and corresponding skill level.

TABLE 11	CORRELATIONS BETWEEN TEST SCORES
	AND ASSOCIATED SKILL LEVEL FOR
	ALL SUBJECTS

(Total	N = 1,232;	1 = 419;	3 = 194; 5 =	= 336; 7 = 283)
	<u>T</u> 1	<u>T</u> 2	<u>T</u> 3	<u> </u>
	.851	.770	.723	.876

One of the primary objectives of the Criterion Test development effort has thus been achieved: the test results disciminate very well among skill levels. Individual Criterion Test Performance is also highly associated with skill level.

Table 12 presents reliability coefficients for each of the three Criterion Tests categorized by skill level.

TABLE	12	RELIABILITY	COEFFICI	ENTS ON	TESTS	1,	2,
		AND 3 CLASSI	FIED BY	SKILL 1	EVEL		

(Total N =	1,232; 1 = 419;	; 3 = 194; 5 =	= 336; 7 = 283)
Skill		Test	
Level	1	2	3
1	.67	.55	.44
3	.95	.86	.63
5	. 95	.88	.85
7	.92	.91	. 91
Pooled 3-5-7	.95	.94	. 94

The reliability coefficients shown in Table 12 are the KR20 version of Cronbach's coefficient Alpha:

(1)
$$r_{tt} = \frac{I}{I-1} \begin{bmatrix} 1 & \frac{\Sigma \sigma_i^2}{\sigma_x^2} \\ 1 & \frac{i=1}{\sigma_x^2} \end{bmatrix}$$

In Equation 1, I is the total number of items in the test,

 σ_i^2 is the variance of each individual item i,

 $\sigma_i^2 = P_i(1-P_i)$ where P_i is the proportion of respondents who answer item i correctly = difficulty of item i,

and σ_x^2 is the variance of the distribution of total scores.

Due to time and scheduling constraints, the current situation permitted only one administration of each of the Criterion Tests. One test administration permits reliability assessment using either a split-half procedure or coefficient alpha (Kuder-Richardson reliability). Split-half reliability estimation was ruled out because of the difficulties involved in splitting the Criterion Tests into equivalent or "parallel" halves. In order to construct parallel halves, each test would have had to be divided into two parts, with each part being equivalent in terms of content, item difficulty, and so forth. In all likelihood, this could have been accomplished; however, a significant increase in development costs would have resulted. Additional costs would also have resulted from the practical necessity that the parallelism of the tests be empirically demonstrated.

Coefficient alpha, or KN20, provides information about the consistency of individual performance from item to item within a single test. The statistic thus provides an estimate of the internal equivalence of the tests. Coefficient alpha has been shown to be a lower bound estimate of the true reliability coefficient of a test. That is, coefficient alpha will always be lower than the true reliability coefficient, and lower than a reliability estimate obtained using another procedure such as split-half or parallel forms. Thus, should coefficient alpha provide an "acceptable" reliability estimate, it may then be said that the true reliability of the test is also "acceptable."

Given the assumption that job proficiency is a unidimensional construct, assessing intra-subject response consistency using coefficient alpha is acceptable. However, if job proficiency is, in reality, a complex or multidimensional construct, then stressing internal consistency using coefficient alpha may not be appropriate. The scope of the present study does not include assessing the factorial nature of the Criterion Tests, or investigating the underlying structure of the concept of job proficiency. The KR20 reliability coefficients are included only as descriptive indices of the degree of intra-subject response consistency across the various items in the tests.

Considering the reliability coefficients presented in Table 12, several patterns are apparent which are worth noting. First, each of the tests provides high reliability/response consistency estimates for its criterion group (e.g., T_1 for level 3 personnel, etc.). Second, as expected, response consistency decreases as the overall difficulty level of the tests increases.

As a final word on reliability/consistency, and also to summarize the descriptive phase of the study, Tables 13, 14, and 15 present item-total score correlations for each of the items on the three Criterion Tests. The correlations are also presented by skill level. The results shown in Tables 13, 14, and 15 extend the information supplied by the KR20 reliability/consistency estimates. That is, CT 1, which yields the highest reliability estimates, also has the highest and most consistent item-total score correlations. This is particularly true for level 3 and level 5 personnel. Item-total score correlations on CTs 2 and 3 are generally lower and less consistent than those of CT 1. This latter situation is also reflected in the lower reliability estimates for CTs 2 and 3.

TABLE 13 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 1

Skill Level Pooled Skill Level Pooled $\frac{1}{.09} \quad \frac{3}{.43} \quad \frac{5}{.22} \quad \frac{7}{.17} \quad \frac{3-5-7}{.26}$ Item 2 1 .30 2 3 .09 .25 .24 .24 3 .05 .20 .31 .23 .35 6 54 87 .11 .30 .32 .23 4 .33 8 .05 .38 .42 .35 .40 55 89 .10 .42 .34 .26 5 9 .10 .18 .37 .18 .35 .30 6 10 -. 03 . 27 . 19 . 14 .21 56 90 -.06 .38 .38 .14 .37 .27 7 14 .16 .25 .20 .26 57 91 .12 .26 .40 .04 .33 8 58 92 .04 .37 .33 .36 15 . 20 . 27 . 24 . 27 .34 .40 9 17 .20 .25 .27 .19 .21 59 93 .04 .22 .19 .11 .15 10 18 .19 .16 .25 .15 .23 60 94 -.01 .33 .35 .28 .38 .33 21 .19 .30 .35 .16 11 61 95 -. 07 . 48 . 41 . 01 .39 .43 12 23 .14 .37 .41 .30 62 96 .23 .49 .35 .30 .39 .39 13 26 .08 .39 .28 .26 63 97 .18 .38 .30 .30 .36 .30 14 27 .08 .14 .30 .31 64 98 .13 .28 .33 .40 .34 65 99 -. 05 . 43 . 32 . 37 28 .05 .25 .28 .15 15 .19 .39 31 .07 .29 .20 .30 .36 66 101 .13 .28 .21 .21 16 .21 32 -. 03 . 22 . 20-. 02 67 102 .05 .19 .26 .21 17 .16 .23 34 . 20 . 39 . 41 . 10 .37 . 28 . 44 . 40 . 27 18 68 104 .39 19 35 .08 .17 .35 .15 .26 69 105 .06 .36 .40 .40 .46 36 -. 03 . 19 . 07 . 26 70 106 .09 .32 .36 .31 20 .16 .34 . 39 37 .10 .37 .30 .34 71 108 .13 .21 .18 .18 21 .19 22 40 .13 .18 .40 .23 .30 72 110 .10 .36 .34 .24 .34 41 . 05 . 32 . 39 . 34 . 38 23 73 111 .17 .47 .40 .36 .41 . 31 24 42 . 07 . 21 . 37 . 21 74 112 . 19 . 39 . 47 . 33 .43 25 43 .14 .32 .33 .25 .32 75 114 . 28 . 44 . 43 . 37 .45 26 44 .17 .29 .21 .20 .24 76 115 .01 .23 .34 .16 . 32 27 46 .03 .26 .23 .14 .26 77 116 . 22 . 35 . 32 . 16 .22 28 47 .08 .25 .08 .32 78 117 -. 21 . 23 . 36 . 13 . 36 .28 48 .09 .12 .30 .16 .22 79 119 . 27 . 46 . 36 . 38 29 .39 30 49 .17 .29 .23-.05 .20 80 120 . 24 . 42 . 47 . 48 .50 52 .11 .34 .19 .11 .21 31 81 121 . 21 . 38 . 17 . 12 .16 32 54 . 16 . 16 . 29 . 35 . 29 82 123 . 25 . 38 . 42 . 51 •45 33 55 .08 .18 .34 .29 .30 83 124 . 22 . 47 . 49 . 48 .52 58.15.27.31.1759.12.38.39.27 . 29 34 84 126 .15 .39 .45 .52 .46 . 36 35 85 128 - 13 . 41 . 37 . 21 .40 36 61 . 16 . 28 . 29 . 17 . 32 86 129 .04 .42 .45 .50 .49 .10.41.28.22 87 131 37 62 .14 .25 .30 .19 .28 .29 88 132 .32 .53 .45 .46 38 63 .14 . 21 . 23 . 29 . 26 .50 89 134 . 22 . 26 . 25 . 19 39 64 .10.42.38.32.41 .14 . 42 . 29 . 30 . 35 90 135 40 .30 65 . 25 . 38 . 31 . 26 . 30 91 136 . 20 . 43 . 41 . 37 . 20 . 22 . 22 . 10 . 15 .38 41 66 92 137 . 22 . 32 . 37 . 41 .44 42 68 .16.40.32.26.35 93 138 .38 .34 .44 .43 43 .41 69 . 16 . 43 . 25 . 13 . 32 . 25 . 26 . 32 . 43 94 140 44 70 . 13 . 43 . 24 . 15 . 22 .34 95 141 . 24 . 34 . 46 . 52 45 71 . 12 . 24 . 33 . 13 . 29 .43 96 142 . 24 . 39 . 44 . 37 46 72 .14 .44 .40 .32 .41 .39 47 74 -. 01 . 46 . 25 . 28 . 30 97 145 .30 .34 .29 .50 .37 75 .08 . 35 . 29 . 13 . 24 48 98 146 .15 .29 .30 .47 .38 76 .03 .19 .19 .18 .15 99 149 . 24 . 30 . 29 . 20 49 .28 50 77 .13 .26 .35 .31 .35 100 150 .38 .41 .43 .43

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

.40

TABLE 14 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 2 (Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

	Skill Level Pooled		Skill Level Pooled
	Item 1 3 5 7 3-5-7		Item 1 3 5 7 3-5-7
1	2 .03 .28 .18 .24 .26	51	75 .0510 .14 .13 .15
2	3 .05 .04 .18 .23 .26	52	76 . 20 . 35 . 42 . 44 . 41
3	4 .0406 .13 .03 .05	53	77 .07 .11 .07 .09 .15
4	5 .12 .14 .28 .19 .31	54	78 .08 .16 .30 .50 .47
5	7 .18 .02031302	55	80 .12 .26 .34 .44 .44
6	9 .08 .28 .33 .34 .44	56	82 . 20 . 29 . 29 . 41 . 38
7	12 .11 .20 .21 .18 .22	57	83 . 17 . 29 . 28 . 18 . 23
8	13 .13 .25 .12 .18 .32	58	84 . 06 . 10 . 16 . 14 . 22
9	14 .06 .27 .20 .26 .46	59	85 .07 .16 .04 .31 .28
10	160000 .16 .35 .47	60	86 .04 .38 .26 .20 .28
11	18 .11 .22 .22 .31 .39	61	88 . 24 . 33 . 29 . 18 . 30
12	19 . 20 . 24 . 31 . 31 . 37	62	89 . 18 . 41 . 36 . 30 . 36
13	20 .14 .16 .16 .26 .29	63	90 . 24 . 34 . 38 . 48 . 49
14	22 . 17 . 32 . 33 . 30 . 42	64	92 .11 .07 .04 .0001
15	24 .03 .37 .19 .22 .39	65	93 .06 .31 .22 .31 .42
16	25 . 23 . 29 . 36 . 26 . 34	66	99 .12 .23 .34 .49 .46
17	26 .11 .31 .07 .25 .28	67	100 . 19 . 17 . 18 . 30 . 35
18	27 .01 .10 .13 .27 .30	68	101 . 02 12 . 13 . 29 . 25
19	29 .12 .13 .21 .27 .16	69	103 .06 .07 .23 .33 .32
20	30 .14 .27 .17 .26 .24	70	104 .11 .22 .36 .53 .55
21	31 .06 .26 .23 .22 .36	71	105 10 11 .09 .04 .18
22	32 .0815010101	72	106 15 22 15 04 12
23	33 . 22 . 32 . 37 . 38 . 42	73	107 - 02 - 21 - 01 - 16 - 05
24	34 .06 .05 .14 .28 .31	74	108 08 14 .12 .18 .18
25	35 14 07-04 16 05	75	109 13 29 26 37 36
26	37 . 20 32 36 27 34	76	111 14 30 19 27 28
27	38 08 30 30 31 34	77	112 20 32 .28 .28 .34
28	39 16 32 22 28 36	78	114 12 20 .27 .10 .19
29	42 08 19 24 19 27	79	116 02 22 .37 .44 .36
30	43 .19 43 39 28 41	80	117 15 12 .16 .29 .25
31	45 - 02 - 01 09 20 23	81	118 09 19 .25 .31 .24
32	46 - 05- 03 16 25 38	82	119 19 07 .02 .11 .00
33	47 .04 .05 .14 .15 .24	83	121 09 13 14 15 37
34	48 14 13 15 27 30	84	123 03 24 .35 .52 .56
35	50 35 27 30 20 38	85	124 11 17 21 38 29
36	$50 \cdot 55 \cdot 27 \cdot 59 \cdot 29 \cdot 50$	86	127 11 09 04-17 - 03
37	52 11 15 08 15 10	87	129 08 21 22 30 28
38	55 = 04 13 18 27 25	88	130 08 29 45 45 48
30	57 10 07 12 27 39	89	131 14 02 .27 .42 .41
40	59 23 36 29 28 28	90	133 11 02 01 29 14
40	60 12 24 35 28 32	01	134 23 36 38 46 38
41	62 08 35 31 24 44	02	135 04 26 25 22 42
43	64 10 17 20 28 28	02	136 03 15 22 40 37
45	65 02 25 17 10 17	95	138 03 28 37 44 40
44	66 - 09 03- 05 10 14	05	130 14 06 01- 17 - 00
45	67 18 31 28 25 34	95	142 20 15 16 06 11
40	68 05-02 09-01 02	07	144 02 = 02 11 21 12
47	69 23 30 37 39 47	08	145 05 17 22 28 38
40		00	147 06 14 17 24 21
50	74 .09 .29 .19 .19 .24	100	149 07 18 21 12 34
		100	

TABLE 15 ITEM-TOTAL SCORE CORRELATIONS FOR FINAL 100 ITEMS ON TEST 3 (Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

Skill Level Pooled Skill Level Pooled Item 1 3-5-7 7 <u>Item 1 3</u> 3 5 5 7 3-5-7 51 68 .15 .39 .31 .27 1 .15 .20 .27 .29 .42 .33 2 2 .17 .16 .17 .12 .19 52 69 .10-.06 .20 .19 .23 3 6 .22 .19 .22 .31 .24 53 72 -.01 .05 .25 .28 .35 4 7 -.01 .11 .18 .14 .12 54 73 .12 .08 .20 .35 .35 .29 8 .16 .09 .16 .30 5 .30 55 75 .15 .21 .24 .19 9 .09-.04-.01 .14 56 77 .07 .07 .09 .05 .13 .09 6 .12 .25 .31 .23 .17 .07 .20 .41 .09 .17 .25 .28 10 .05-.02-.03 .03 57 79 .23 7 .09 80 8 11 .11 .16 .26 .22 .33 58 .39 .15 .10 .03 .10 .15 59 81 .41 9 12 10 13 .04-.17 .10 .16 60 82 .10 .22 .24 .19 .19 .20 11 14 .26 .29 .32 .35 .39 61 83 .12-.07 .20 .23 .27 85 .02-.02-.06 .14 .13 .13-.05 .07 12 15 .05 62 .06 .17 .08 .20 .07-.02 86 .08 .13-.04 .13 13 16 .08 63 .11 .07 .29 .27 64 87 .04 .19 .19 .15 17 .32 14 .43 65 88 .09-.14 .01 .08 12 .16 .37 .23 18 .06 15 .50 66 90 .01-.05 .33 .39 .49 19 .03-.03 .18 .17 .33 16 20 .11 .10 .26 .25 17 .38 67 91 .04 .02 .14 .37 .32 .11 .03 .23 .35 .36 18 21 68 92 .08 .02 .10 .27 .20 .33 22 .11 .15 .21 .38 19 69 93 .06 .14 .25 .35 .48 .13 .21 .10 .25 .22 20 24 70 101 .09 .12 .25 .46 .54 21 25 .08-.02 .24 .33 71 102 .15 .23 .20 .19 .46 .21 22 27 .18 .34 .31 .21 .32 72 103 -. 03-. 07 .18 .22 .34 23 28 .31

.14

.17

. 28

.23

.43

.34

.22

.19

.07

.13

.50

.39

.37

.25

.37

.22

.12

.18

.29

.06

.21

.16

.28

.27

.17

.27

.21

100 149 .16 .10 .20 .25

.20 .21 .16 .17 73 104 .01 .19 .30 .17 .20 74 105 .15 .01 .09 .27 .39 .05 .20 .35 .30 .07 .01 .16 .30 75 107 .03 .10 .13 .24 .34 76 108 . 19 . 22 . 27 . 33 .12 .07 .15 .12 .18 .17 .19 .13 .14 77 109 .09 .10 .12 .24 .10 78 113 . 14 .29 .38 .22 ·11 ·12 ·14 ·11 .16 .06 .08 .31 .34 79 115 . 09 .10 .20 .36 .46 .10 .09 .21 .32 .39 80 117 . 07 .21 .10 .19 .06-.05 .20 .33 .39 81 119 -. 01 .14 .11-.01 .18 .07 .23 .31 .28 82 120 . 12 .07 -.03 .24 .10 .13 .20 .15 .32 83 121 .07 .08 .09 .05 .14 .18 .36 .46 .51 84 122 .15 .16 .26 .40 .05 .21 .31 .27 85 123 . 17 .23 .30 .40 .45 86 126 .03 .27 .10 .02 .00 .12 .05 .25 .21 .11 .28 .25 .38 87 127 -. 00 .16 .05-.06 .10 .31 .36 . 11 .10 .05-.01 .08 .19 88 129 .19 .11 .16 89 131 .05 .09 .21 .14 .13 .25 .25 90 132 .01 .11 .01 .23 52 -.04-.02 .09 .05 .10 .19 .15 91 135 . 26 .22 ·01-·11 ·11 ·33 .31 54 -.04 .10 .16 .04 92 136 . 17 .33 .34 .16 .13 .04 .21 .06 .03 .17 .34 .35 .37 .36 93 137 .13 .18 .20 .23 .25 94 138 . 16 .12 .10 .22 .12 .09 .15 .17 .01-.00 .24 .53 .54 95 140 96 143 .30 .34 .22 .39 .06 .11 .17 .15 .22

 97
 144
 .05-.10
 .21
 .28

 98
 145
 .03-.02
 .03
 .17

 99
 146
 .05
 .12
 .12
 .15

24

25 31

26 32

27 33

28

29 35

30 39

31 40

32

33 43

34 44

35

36 46

37

38 49

39

40

41

42

43 55

44 56

45 57

46 59

47 62

48 63

49 64

50 66

29

34

41

45

47

51

53

.14 .18 .28 .31

.15 .15 .19 .25

.08 .14 .22 .46

.15 .30 .40 .28

24

.40

.26

.36

.35

VALIDITY STUDIES

The intent of the present study was to develop a series of paper-andpencil measures of job performance against which aptitude tests such as the ASVAB could be validated. The validity study will, thus, primarily address the associations between the ASVAB scales and the three Criterion Tests. Specifically, validity assessment is developed in four phases:

- 1. The relationship between the ASVAB and the Criterion Tests.
- 2. The relationship between technical school grades (the previous criterion) and the ASVAB.
- 3. The relationship between technical school grades and the Criterion Tests.
- 4. Cross-validation studies.

Validation studies of this type usually concern the regression of an outcome measure on a test score. Phase 1 thus embodies the main features of the present validation effort. Phases 2 and 3 are included to assess the relationship of the previous criterion with the newly developed one. The last phase--cross-validation--is included to get an estimate of the degree of shrinkage in predictability involved with using weights derived from one sample to predict job performance in another independent group of subjects.

The Relationship of the ASVAB and the Criterion Tests

The means and standard deviations of each of the skill level groups on the nine ASVAB scales are presented in Table 16. A multivariate analysis of variance (MANOVA) performed on the ASVAB scores indicates that the skill level groups have significantly different mean profiles (asymptotic chisquare = 458.0319 with 27 degrees of freedom, p <.0001). The corresponding univariate tests indicate that two ASVAB subscales--Coding Speed and Automotive Information--are primarily responsible for this significant difference. From the mean scores presented in Table 16, it is noted that the basic airmen score higher than the older mechanics on Coding Speed, while scoring lower on Automotive Information. Coding Speed is a clerical skill heavily dependent upon quickness and mental agility. Thus, age logically could be a factor in these scores. The fact that the older, more skilled mechanics score higher than basic airmen on Automotive Information is also not surprising, given the mechanical exposure of these subjects. TABLE 16 ASVAB MEANS AND STANDARD DEVIATIONS BY SKILL LEVEL

			(Total N =	1,232; 1 =	419; 3 = 194	; 5 = 336; 7	= 283)			
					ASVAB Scale					
Skil Leve		(1) Coding Speed	(2) Word <u>Knowledge</u>	(3) Arith. <u>Reasoning</u>	(4) Tool <u>Knowledge</u>	(5) Space <u>Perception</u>	(6) Mech. Comp.	(7) Shop Info.	(8) Auto. <u>Info</u> .	(9) Elec. Info.
Ч	M SD	50.64 15.76	20.16 3.74	14.94 5.36	14.57 4.63	14.23 5.38	14.58 3.86	14.69 3.98	13.91 5.65	15.03
m	M SD	46.52 10.98	20.40 3.28	14.16 4.87	15.37 5.56	13.41 5.23	14.37 4.13	15.37 4.37	15.85 5.53	15.59 4.43
2	M CS	42.95 16.65	18.92 5.45	13.36 5.52	14.91 6.64	12.13 6.18	13.15 5.45	14.52 5.97	16.29 6.40	15.24
-	N QS	42.90 13.40	20.77 4.28	15.07 5.27	17.93 6.79	11.63 5.38	15.05 4.83	17.84 5.03	19.95 4.92	18.64 4.17

26

The correlation matrix for the nine ASVAB scales, pooled over skill levels, is presented in Table 17. Most of the correlations between ASVAB scales are in the low-to-moderate range, e.g., 30's to 40's. One exception to this general result involves the Information Scales (Shop, Automotive, and Electronic) where the correlations are noticeably higher. However, given the experience of the subjects involved, prior expectation is that the Information Scales will be highly correlated. The correlation matrix for level 1 personnel considered separately shows approximately the same pattern, except that the Information Scales are not as highly correlated.

TABLE 17 CORRELATIONS OF ASVAB SCALES POOLED OVER SKILL LEVELS

(Total N = 813; 3 = 194; 5 = 336; 7 = 283)

	<u>1</u>	2	3	4	5	<u>6</u>	7	<u>8</u>	9
1	1.00								
2	0.39	1.00							
3	0.39	0.50	1.00						
4	0.10	0.23	0.17	1.00					~.
5	0.32	0.36	0.44	0.39	1.00				
6	0.32	0.50	0.46	0.43	0.53	1.00			
7	0.22	0.49	0.34	0.61	0.37	0.60	1.00		
8	0.13	0.39	0.29	0.57	0.26	0.53	0.70	1.00	
9	0.18	0.49	0.38	0.49	0.31	0.53	0.68	0.65	1.00

Table 18 shows the correlations between individual ASVAB scales and the three Criterion Test scores plus the total test score. Individual correlation results are also presented by skill level.

None of the ASVAB scales is highly correlated with any of the Criterion Tests within any skill level. The individual ASVAB scales that correlate most highly with Criterion Test performance are the Information Scales (Shop, Automotive, and Electronic). Arithmetic Reasoning also has a slightly higher than average correlation with the Criterion Tests.

In order to more completely assess the relationship between the ASVAB scales and the Criterion Tests, a stepwise regression procedure was used to select the best sets of ASVAB predictors for the test scores. Stepwise regression is used to achieve a compromise between the increased precision gained by including as many predictors as possible, and the practical necessity of reducing complexity and costs by including only as many variables as necessary. A stepwise regression procedure enters into regression only those variables with an independent contribution to prediction that is statistically significant. After the first step in the process, the contribution of each variable already in regression is again

TABLE 18 CORRELATIONS OF ASVAB SCALES WITH CRITERION TEST SCORES BY SKILL LEVEL

Skill					ASVAB	Scale			121	
Level	Test	1	2	3	4	5	6	7	8	9
	1	.171	.273	. 209	. 247	.174	.250	.291	.303	.341
	2	.081	.216	.240	. 218	.196	.178	.255	. 225	.334
T	3	.158	.211	.179	.095	.146	.131	.089	.079	.133
	Т	.203	.349	.311	. 286	.256	.283	.325	.313	.411
	1	. 231	. 371	. 283	.214	. 257	. 390	. 360	. 428	. 399
	2	.010	.231	. 201	.149	. 264	. 275	.114	. 281	. 275
3	3	.147	.170	.236	.112	. 282	.190	.110	.200	.255
	Т	.179	.357	.308	.216	. 329	. 391	. 289	.418	.413
	1	.254	. 293	.336	.116	.245	.228	.219	. 299	.320
5	2	.063	.159	.234	.064	.111	.134	.169	.211	.255
-	3	.174	. 240	.247	.095	.142	.127	.226	.168	.263
	Т	.207	.283	.335	.112	.210	.205	.245	. 283	.340
	1	.245	.113	.208	.159	.190	.170	.091	.118	.165
-	2	.197	.175	. 229	.147	.193	.200	.178	. 216	.235
/	3	.217	.181	. 269	.068	.152	.092	.092	.136	.155
	Т	.258	.187	. 279	.145	. 209	.180	.143	.186	.218
	1	.213	. 291	. 312	.211	.190	. 281	. 291	. 365	. 384
	2	.024	. 189	. 235	.210	.052	. 204	.275	. 362	. 379
3-5-7	3	.077	. 214	.250	.201	.038	.165	. 280	. 323	. 363
	Т	.119	. 265	. 305	.238	.107	.249	.323	.402	.431

(Total N = 1,232; 1 = 419; 3 = 194; 5 = 336; 7 = 283)

evaluated to see if it is needed, given that the new variable is now in regression. The process usually results in a regression equation containing those variables with the largest individual contribution to prediction.

Stepwise regression results using job performance test scores as criteria and the ASVAB scales as predictors are presented in Table 19. The results in Table 19 are given separately for pooled 3, 5, and 7 level personnel and for basic airmen (level 1). Basic airmen are separated from experienced jet mechanics for analysis purposes because of the possibility that the experience factor (e.g., time in service, attendance at technical school, job incumbency, and so forth) would moderate the relationship between the Criterion Tests and the ASVAB.

Separate regression solutions are also presented for the total sample and a "computing subsample." The results involving the computing subsample are to be used in the later cross-validation exercise. In all analyses, the three most predictive ASVAB scales are entered into regression.
TABLE 19 SUMMARY OF STEPWISE REGRESSION ANALYSES INVOLVING CRITERION TEST SCORES AND ASVAB SCALES

Test 1--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 3, and 8 Percentage of Variance Explained $(R^2) = 18.8\%$ **Regression Coefficients:** B0 = 37.07B9 = 0.712B3 = 0.525B8 = 0.487ANALYSIS OF VARIANCE P MS F Source DF SS 812 244763.00 Total 45998.09 62.41 .0001 Regression 3 15332.70 198764.91 809 245.69 Residual

Test 1--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

Variables in Regression--9, 3, and 1 Percentage of Variance Explained = 18.0% **Regression Coefficients:** B0 = 38.40B9 = 1.045B3 = 0.385B1 = 0.096ANALYSIS OF VARIANCE F P DF MS Source SS 401 108406.50 Total Regression 3 19552.29 6517.43 29.19 .0001 Residual 398 88854.21 223.25

Test 1--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 2, and 8 Percentage of Variance Explained: 17.1% **Regression Coefficients:** B0 = -11.878B9 =0.338 B2 =0.383 B8 =0.219 ANALYSIS OF VARIANCE Source DF SS MS F P Tetal 418 23369.85 Fegression 1334.58 28.60 .0001 3 4003.74 Residual 415 19366.11 46.67

TABLE 19 (continued)

Test 2--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 1 Percentage of Variance Explained: 17.4% Regression Coefficients: B0 = 8.120B9 = 0.967B8 = 0.632B1 = -0.129ANALYSIS OF VARIANCE Source DF SS MS F Total 812 259029.75 Regression 3 45023.32 15007.77 56.73 .0001 Residual 809 214006.42 264.53

Test 2--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

P

Variables in Regression--9, 5, and 8 Percentage of Variance Explained: 18.4% Regression Coefficients: B0 = 8.884B9 = 1.144B5 = -0.538B8 = 0.509ANALYSIS OF VARIANCE Source DF SS MS F P Tota1 401 124156.47 Regression 22793.47 3 7597.82 29.83 .0001 Residual 398 101363,00 254.68

Test 2--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 3, and 7 Percentage of Variance Explained: 14.1% Regression Coefficients: B0 = -5.073B9 = 0.342B3 = 0.187B7 = 0.169ANALYSIS OF VARIANCE Source DF SS MS F P Tota1 418 17987.88 Regression 3 2527.86 842.62 22.62 .0001 15460.02 **Residual** 415 37.25

TABLE 19 (continued)

Test 3--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 6 Percentage of Variance Explained: 15.3% Regression Coefficients: B0 = 0.317B9 = 1.096B8 = 0.560B6 = -0.469ANALYSIS OF VARIANCE Source DF SS F MS Total 812 238692.72 Regression 36527.64 3 12175.88 48.72 .0001 Residual 809 202165.08 249.90

Test 3--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

P

Variables in Regression--9, 3, and 5 Percentage of Variance Explained: 18.5% Regression Coefficients: B0 = -0.096B9 = 1.240B3 = 0.666B5 = -0.608

ANALYSIS OF VARIANCE

Source	DF	SS	MS	<u>F</u>	P
Total	401	122331.89			
Regression	3	22633.00	7544.33	30.12	.0001
Residual	398	99698.88	250.50		

Test 3--Basic Airmen (Level-1; N = 419)

Variables :	in Regres	sion2, 1, a	nd 4		
Percentage	of Varia	nce Explained	: 6.8%		
Regression	Coeffici	ents:			
BO = -	-6.724				
B2 =	0.294				
B1 =	0.046				
B4 =	0.108				
		ANALYSIS OF	VARIANCE		
Source	DF	SS	MS	F	P
Total	418	13965.04			
Regression	3	954.08	318.03	10.14	.0001
Residua1	415	13010.96	31.35		

TABLE 19 (continued)

Total Test Score--Pooled 3, 5, and 7-Level Personnel (Total Sample; N = 813)

Variables in Regression--9, 8, and 3 Percentage of Variance Explained: 22.0% Regression Coefficients: B0 = 37.211B9 = 2.371B8 = 1.390B3 = 1.073ANALYSIS OF VARIANCE Source DF SS MS F Total 812 1689969.00 371977.88 123992.63 Regression 3 76.11 Residual 809 1317991.12 1629.16

Total Test Score--Pooled 3, 5, and 7-Level Personnel (Computing Subsample; N = 402)

P

.0001

Variables in Regression--9, 3, and 5 Percentage of Variance Explained: 22.6% Regression Coefficients: BO = 49.011 B9 = 3.713 B3 = 1.610 B5 = -1.262 ANALYSIS OF VARIANCE

Source	DF	SS	MS	F	P
Total	401	815801.50			
Regression	3	184368.03	61456.01	38.74	.0001
Residual	398	631433.44	1586.52		

Total Test Score--Basic Airmen (Level-1; N = 419)

Variables in Regression--9, 2, and 4 Percentage of Variance Explained: 24.7% Regression Coefficients: B0 = -25.709B9 =0.786 B2 =0.987 B4 =0.442 ANALYSIS OF VARIANCE Source DF SS MS F P Tota1 418 75095.65 Regression 3 18573.60 6191.20 45.46 .0001 **Residual** 415 56522.05 136.20

Table 19 presents a great deal of regression information. However, there are several patterns which appear to cut across the various analyses. The most obvious result is that all of the regression runs are highly significant. Next, considering the analyses involving the pooled 3, 5, and 7 level personnel, the most obvious result is that ASVAB scale 9--Electronic Information--is entered first in all the runs. The Electronics Information scale contains questions concerning knowledge of electricity and electronics. This scale thus appears to represent a high order experiential factor. ASVAB scale 8--Automotive Information--is also included in the final set of predictors in all analyses. Automotive Information does not, however, have the consistently high relationship with Criterion Test performance of Scale 9. The overall predictive efficiency of the various regression analyses as indexed by the squared multiple correlation (e.g., the percentage of variance explained) ranges from 15.3% for CT 3, up to 22.0% for the total test score. These squared multiple correlations are comparable to those usually obtained in studies involving the regression of aptitude or intelligence measures on achievement test performance.

The corresponding regression analyses involving the basic airmen indicate a similar pattern. In all cases except CT 3, Electronics Information is again the best single predictor of Criterion Test performance. The most potent predictor after Electronics Information is Word Knowledge. On CT 3, which primarily addresses the procedural and administrative aspects of job performance, Coding Speed and Word Knowledge are the best predictors of Criterion Test performance. However, the relationship between the ASVAB and CT 3 for basic airmen is substantially lower than is the case with the other tests.

In summary, the regression results involving experienced jet engine mechanics indicate that the ASVAB information scales, particulary Electronics Information and Automotive Information, are the most consistent predictors of Criterion Test performance. For basic airmen, Electronics Information is also a significant predictor; however, other ASVAB scales involving academic experience (e.g., Word Knowledge and Arithmetic Reasoning) are more important than is the case with the experienced mechanics. Taken together, the regression results suggest that previous mechanical or electronic experience plus general educational level, as indexed by verbal and arithmetic skills, are the best predictors of Criterion Test performance.

The above results are not surprising given the fact that many of the experienced mechanics taking the tests have several years of direct shop and mechanical experience. On-the-job experiences for experienced mechanics should result in related scores on the ASVAB information scales and the Criterion Tests. Additionally, verbal and arithmetic skills are usually a factor in test performance.

The ASVAB Versus Technical School Grades

The second phase of the validation study involves the relationship between technical school grades and the ASVAB. Phase 2, along with Phase 3, is included in the validation study for the following reasons:

- To permit a comparison of relationships in the present sample with previously established relationships.
- To obtain some information concerning the extent to which changing the criterion for selection will affect the composition of the pool of jet engine mechanics.

As previously indicated, the Air Force has traditionally used technical school grades (TSGs) as a criterion measure for the validation of aptitude tests. Changing the criterion measure could have the effect of selecting on different aptitude dimensions, thus changing the composition of the pool of basic airmen assigned to jet engine mechanic courses. In order to determine whether or not using the job knowledge tests as criteria will change the dimensions of selection, the ASVAB scales are regressed on TSGs, and the results compared with the previous finding concerning the relationships between the Criterion Tests and the ASVAB.

Table 20 presents the means and standard deviations of technical school grades separated by skill level. Since basic airmen have not been to the technical school, only experienced jet engine mechanics are included.

	S	Skill Level						
	3	<u>5</u>	<u>7</u>					
Mean	88.24	86.88	86.83					
Standard Deviation	5.98	6.18	5.18					
N	167 (86%)	275 (82%)	89 (31%)					

TABLE 20 MEANS AND STANDARD DEVIATIONS OF TECHNICAL SCHOOL GRADES BY SKILL LEVEL

The numbers of observations in each of the skill groups displayed in Table 20 are lower than the numbers indicated in previous analyses. This decrement resulted from the fact that a number of subjects did not match the Technical Training Course Grade File, or matched the file but did not have a course number of 3ABR426X2 or 3ABR432X0. Several subjects matched the Technical Training Course Grade File and attended the course, but did not have a valid grade. The extent of subject loss, for one reason or another, is greatest for the 7 level personnel. Only 31% of the 7 level subjects have valid TSGs. The corresponding capture rates for level 3 and level 5 personnel are 86% and 82%, respectively.

The mean TSGs do not show much difference across skill levels. Mean TSG scores range from 88.24 to 86.83. An analysis of variance performed on the TSG mean scores across skill levels resulted in an F of 2.76, with 2 and 528 degrees of freedom; this value is not statistically significant (p = .0625). An additional observation concerning the TSG scores is the relatively small size of the standard deviations. The range of standard deviations is from 5.18 to 6.18. Such restricted ranges of TSG scores could have the effect of depressing the relationships of TSGs with the ASVAB scales and Criterion Tests.

The correlations of the ASVAB scales with ISGs, separated by skill level, are given in Table 21. None of the ASVAB subscales individually has a strong relationship with TSGs. The strongest associations, which hold across skill levels, are ASVAB scales 2, 3, 6, and 9--Word Knowledge, Arithmetic Reasoning, Mechanical Comprehension, and Electronics Information. The stepwise regression results summarized in Table 22 extend and clarify the correlation results.

TABLE 21 CORRELATIONS OF ASVAB SCALES WITH TECHNICAL SCHOOL GRADES BY SKILL LEVEL

		ASVAB Scale														
	3	.129	.407	<u>3</u> .414	$.1\frac{4}{121}$. <u>5</u> . 294	. <u>6</u> . <u>34</u> 7	.237	. 228	<u>9</u> . 335						
Skill Level	5	.308	. 339	.404	.092	. 227	.277	.224	. 283	. 333						
20102	7	.045	.082	.148	. 171	.141	.437	.270	. 286	.224						
Pooled		.227	.321	. 369	.109	.240	.317	.227	.241	. 298						

(101a1 n - 331, 3 - 107, 3 - 273, 7 - 07)	(Т	otal	N	=	531	; 3	=	167	; 5	=	275	5;	7	=	89	1
---	----	------	---	---	-----	-----	---	-----	-----	---	-----	----	---	---	----	---

TABLE 22SUMMARY OF STEPWISE REGRESSION RESULTS
USING TECHNICAL SCHOOL GRADES AS THE
CRITERION AND ASVAB SCALES AS PREDICTORS
BY SKILL LEVEL

		Sca	les	s I	nc	lude	d in	Re	gres	Percentage of Variation Explain						
			To	ota	1		Com	put	ing		Total		Computing			
			Sample N		Subsample			N	Sample	N	Subsample	N				
Skill Level	3		3,	2,	9	167	3,	9,	1	96	28.1%	167	31.0%	96		
	5		3,	9,	1	275	1,	3,	9	144	23.5%	275	27.2%	144		
	7		6,	8,	2	89	6,	9,	3	42	20.5%	89	34.7%	42		
Pooled			3,	9,	1	531	3,	1,	6	282	19.6%	531	25.1%	282		

Considering Table 22, the variables entered into regression are similar for the 3 and 5 level personnel. ASVAB scales 3, 9, and either of scales 1 or 2, are selected in all cases. For level 7 personnel, the situation is different. Here, the regression results are less consistent, except for the fact that Mechanical Comprehension is always included in the set of predictors.

The results for the pooled sample represent a blending of the two previous conditions. In the pooled sample, Coding Speed and Word Knowledge are included in both regression analyses. However, Electronics Information is included in the runs involving the total sample, while Mechanical Comprehension is entered into regression in the runs involving the computing subsample.

The stepwise analyses summarized in Table 22 have somewhat higher levels of precision than is the case for the regression results involving the Criterion Tests and ASVAB scales. In the regressions involving the ASVAB and TSGs, the squared multiple correlations (SMCs) range from 19.6% to 34.7%. Additionally, the ASVAB scales most predictive of the dependent measure are different in the two situations. With the jobrelated Criterion Tests, ASVAB information measures are the best predictors. On the other hand, general academic skills such as Word Knowledge and Arithmetic Reasoning are most indicative of technical school performance. These latter scales are also related to Criterion Test performance, but not to the same extent as with TSGs. It thus appears that the two criteria--the job performance Criterion Tests and TSGs--are related to slightly different aspects of individual aptitude. Using the TSGs as a criterion would appear to favor better schooled, more academically proficient job candidates; using the Criterion Tests would apparently favor candidates having previous experience with, or interest in, mechanical/electrical tasks.

The Relationship Between the Criterion Tests and TSGs

In order to further clarify the relationship between the two criterion measures, the third phase of the validation study concerns the direct association between the Criterion Tests and technical school grades. Table 23 presents the correlations between the three Criterion Tests scores, their total, and TSGs.

TABLE 23 CORRELATIONS BETWEEN CRITERION TEST SCORES AND TECHNICAL SCHOOL GRADES BY SKILL LEVEL

	C	riterion Tes		
Level <u>N</u>		2	<u>3</u>	Total Score
167	.478	. 423	.279	. 570
275	.608	.381	.244	. 525
89	.416	. 277	.181	.346
531	. 528	. 255	.134	.408
	<u>N</u> 167 275 89 531	$ \underbrace{ \begin{array}{c} N \\ 1 \\ 167 \\ 275 \\ 89 \\ 416 \\ 531 \\ 528 \\ \end{array} } \underbrace{ \begin{array}{c} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\$		$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$

The correlations presented in Table 23 indicate several results worthy of mention. First, the correlations between TSGs and the Criterion Test scores are higher for CT 1 than the other tests. This result is not surprising, since CT 1 primarily addresses the job content of level 3 personnel. The material taught in technical school is most closely related to level 3 job skills. The job content items in CTs 2 and 3 do not relate as directly to the material presented in the technical school, thus their correlation with technical school performance should not be as high.

A second observation, concerning the relationship between Criterion Test performance and TSGs, is that, with one exception, the correlations decline with increasing skill level. This suggests that job incumbency reduces the relationship between school performance and the behaviors measured by the Criterion Tests. A variety of factors such as selection and on-the-job experience could account for this result.

The stepwise regression results summarized in Table 24 support and extend the previous correlation results. Table 24 presents only those variables which meet an entry requirement of having a partial-F statistic which is significant at the p = .05 level. If no variable is indicated, then no Criterion Test met the entry requirement.

TABLE 24 STEPWISE REGRESSION RESULTS USING TECHNICAL SCHOOL GRADES AS CRITERIA AND CRITERION TEST SCORES AS PREDICTORS

Skill Level	iatan.	1	Vari Reg	ables : ression	in 1	Dad ki	Percentage of Variation Explained							
Sample	Tot	<u>a1</u>	<u>N</u>	Comput	ting	<u>N</u>	% Total	<u>N</u>	% Computing	<u>N</u>				
3	3 1, 2 5 1		167	1, 3		96	28.2	167	37.9	96				
5			275	1		144	36.9	275	31.6	144				
7	1 8		89	10 230 <u>4</u>		42	17.3	89	-	42				
Pooled	1,	3	531	1,	3	282	29.2	531	27.8	282				

(Total N = 531; 3 = 167; 5 = 275; 7 = 89)

Not surprisingly, the regression results again indicate that scores on CT 1 are most predictive of technical school performance. CT 1 enters regression first in all cases. For skill levels 5 and 7, CT 1 is the only score significantly associated with technical school performance. The squared multiple correlations also indicate that the precision of prediction is higher for lower skill level mechanics than for the more experienced personnel.

Cross-Validation

The last phase of the validation study involves cross-validation of the regression results. In a restricted sense, cross-validation involves computing regression parameters in one sample of subjects and then assessing the predictive efficiency of these weights in a similar but independent group of subjects.

In order to accomplish the above objective, the total sample of subjects was, first, randomly divided into two subsamples denoted the "computing sample" and the "cross-validation sample." Regression parameters were derived using the computing sample, and then used to generate predicted scores in the cross-validation sample. In each case, weights were obtained for the three most predictive ASVAB scales. Before being used to generate predicted scores, the raw regression coefficients were converted to integer weights by multiplying by 10 and then rounding to the nearest integer. The final step involved computing the correlations between actual and predicted scores using the subjects in the crossvalidation sample.

Table 25 presents the correlations computed between the actual and predicted scores for selected situations. Since only certain Criterion Tests are relevant for specific skill groups, e.g., CT 1 for skill level 3 and so forth, not all sets of results are of interest. Only those

IN	
SCORES	
AND PREDICTED	SKILL LEVEL
CORRELATIONS BETWEEN ACTUAL	DSS-VALIDATION SUBSAMPLE BY
TABLE 25	THE CR

e Lo Vo	Comp Sample 2 R ²	49 .369	93.216	36.070	95.226		
IN Total	Cross-Val Sample r r	.386 .14	.306 .09	.293 .08	.442 .19		
KEDICTED SCORES LEVEL T3	Val Comp ple Sample r2 R2 N	N/A 98	N/A 173	.054 .084 140	.137 .185 411		
WEEN ACTUAL AND PI UBSAMPLE BY SKILL iable T2	coss-ValCompCross-SampleSampleSamrr2R2r	N/A	193 .037 .156	N/A .232	404 .163 .184 .370		
TABLE 25 CORRELATIONS BET THE CROSS-VALIDATION S Var	$\begin{array}{c cccc} Cross-Val & Comp & Cross-Val & Cross-V$.205 .042 .310 71 .372 .138 .416	.455 .207 .272 131 N/A .	.212.045.347 47 N/A	.364 .132 .251 249 .420 .176 .180 .		
	skill Level Correlation:	3	5	7	Pooled		

correlations relevant to a particular group are presented. For purposes of comparison and evaluation, the squared multiple correlations from the corresponding regression analysis involving the computing subsample are presented with the squared cross-validation correlation coefficients.

As is typically the case in studies of this nature, the squared correlations computed from the cross-validation sample show the shrinkage associated with applying regression weights in a second sample. The actual magnitude of the cross-validation correlations usually reflects the squared multiple correlations obtained in the computing subsample. However, in most cases, the squared correlations are so low that they must be considered practically not significant.

CONCLUSIONS

Summary

The purpose of the present study was to develop a set of job-related Criterion Tests against which to validate aptitude measures. Toward that objective, three 150-item job performance Criterion Tests were developed and administered to 1232 personnel ranging from basic airmen to 7-skill level senior jet engine mechanics. By selecting those items which best differentiated skill levels, and were most homogeneous in terms of their correlation with the total test score, the Criterion Tests were reduced to a final form of 100 items. Analysis of variance and correlation results indicated that scores on the Criterion Tests were highly associated with actual skill levels. Reliability assessment using coefficient alpha also demonstrated that subject responses were sufficiently consistent.

In terms of the associations between the Criterion Test scores and other measures, the results of the study indicated that ASVAB information subscales were most predictive of Criterion Test performance for experienced mechanics. Information measures were also related to test performance for basic airmen, but general knowledge subscales, e.g., Word Knowledge, played a more prominent role for these subjects.

When technical school grades were regressed on the ASVAB, academically oriented scales, such as Word Knowledge and Arithmetic Reasoning, were found to be most consistently predictive of technical school performance. A possible implication of this result is that changing the standard for selection as a jet engine mechanic from TSGs to the Criterion Tests may tend to select in favor of potential candidates with previous mechanical/electrical experience.

As a final step in comparing the two criterion measures, TSGs were regressed on the Criterion Tests. The results indicated that only CT 1 was consistently predictive of technical school performance. This result is not surprising, since CT 1 covers material similar to that presented in the technical school course of study.

The final phase of the study involved cross-validation. Here, regression results from a randomly selected half of the subjects were used to generate predicted scores for the other half. The derived correlations between the actual and predicted scores demonstrated the shrinkage typically found in regression cross-validation studies. Given the low percentages of variance accounted for in the initial regression analysis, the corresponding correlations in the cross-validation phase indicated a very low relationship between actual and predicted scores.

Discussion

Considered as a whole, the analytical results indicate a statistically significant relationship between Criterion Test performance and the aptitude measures. The degree of the relationship is such, however, that the practical significance of the relationship is marginal. That is, ASVAB scores cannot be used to reliably predict future job knowledge, as measured by the Criterion Tests. The Criterion Test scores do, however, appear to be quite indicative of the actual career level of jet engine mechanics.

The rather weak relationship between the Criterion Tests and the ASVAB found in the present study is typical of results obtained in other similar investigations. When aptitude measures are used to predict job success, the relationships are usually attenuated by a number of "moderator" variables. These moderator variables, such as age, IQ, geographical background, and so forth, serve to modify the relationship between the predictor variable and the criterion measures. One solution to this problem involves computing separate regression equations for major subgroups in the population. This could be done for low IQ and high IQ personnel, for urban versus rural personnel, and so forth. Another possibility, which possibly would involve less "capitalization on chance," is to include the moderator variables as predictors in the regression equation. This would be done by collecting a broad range of biographic data indicative of the background, relevant experiences, interests, and personality of the subjects, in addition to aptitude information. The biographical variables are then used, separately or in a combined form, in an attempt to account for variation that normally becomes part of residual error.

The present Criterion Test development study ideally lends itself to such a procedure. Not only are a sufficiently large number of subjects involved, but also a considerable amount of relevant biographical information has already been obtained on the subjects. However, due to the contractual limits of the study, the biographic variables were not included in the regression analysis. The consideration of biographical moderator variables thus remains an area for future study.

The current study also makes the assumption that career progression is a valid measure of job proficiency. If Criterion Test performance is associated with career progression, it is then assumed that the Criterion Tests may also be considered a measure of job proficiency. However, in reality, the validity of the first assumption is open to criticism. It might well be argued that career progression is really a function of endurance. That is, personnel who stay in the Air Force will get promoted. In that case, actual job proficiency may only marginally be a factor in promotion. In order to strengthen the theoretical bases of the present study, in the event the Criterion Tests are used to select personnel, it would be necessary to more completely specify the relationship between Criterion Test performance and actual job proficiency. APPENDIX 1 ITEM ANALYSIS OF TEST 1 FOR 3-LEVEL AND 1-LEVEL PERSONNEL

1041 - 410. 3 -612. 1 (Total N

	Include	Yes																				
	1*1	53.0	44.0	2.5	11.5	63.5	53.5	34.5	32.0	36.5	21.0	64.0	25.5	19.5	22.0	48.0	32.0	26.0	57.5	13.0	17.0	
	rit(0)	002	.068	063	.003	.004	.079	.068	.151	055	010	093	.061	.194	.048	.086	.093	.121	.205	.067	009	
	s ² (0)	.19	60.	.18	.17	.14	.17	.19	.13	.15	.18	.13	.17	.22	.10	.20	.19	.11	.20	.19	.21	
= T34)	Diff(0)	.25	.10	.24	.22	.17	.22	.25	.16	.19	.24	.16	.21	.32	.11	.29	.26	.12	.27	.25	.30	
C :4TH	Rank	105	86	2	19	122	101	62	56	64	32	117	39	26	30	81	48	35	67	2	14	
= T ¹ CTO	$r_{it(3)}$.195	.235	.470	.399	.153	.205	.277	.297	.274	.369	.169	.346	.383	.370	.244	.325	.357	.220	.443	.414	
- N TPIO	s ² (3)	90.	.21	.08	.11	.16	.13	.10	.18	.15	.12	.19	.15	.08	.23	.10	.13	.22	.12	.14	.10	
	Diff(3)	66.	.71	16.	.87	.79	.85	.88	.76	.81	.86	.74	.81	.91	.64	.88	.85	.66	.86	.84	.89	
	\$(0,3)	.64	.62	.62	.61	.60	.59	.59	.59	.59	.58	.57	.57	.56	.55	.55	.55	.55	.55	.55	.55	
	Item	81	62	95	80	35	6	17	141	10	06	36	126	34	23	11	41	142	17	89	66	
	Rank	1	2	e	4	2	9	2	80	6	10	11	12	13	14	15	16	17	18	19	20	

 $\frac{*}{\mathbb{R}}$ is the average rank of $\phi(0,3)$ and ^rit(3).

Include	Yes																			
12	20.5	39.5	33.0	35.5	33.5	8.41	21.0	51.5	64.0	78.5	20.0	54.5	45.5	29.5	31.0	85.5	30.0	55.5	75.0	38.0
r _{it(0)}	026	.050	.003	.167	.000	.078	.149	.008	003	.200	.094	.033	.151	.188	.082	.165	.045	.098	.003	.113
s ² (0)	60.	.20	.10	.14	.15	.18	.19	.20	.20	.24	.21	60.	.16	.18	.15	.24	.14	.20	.18	.18
Diff(0)	.10	.27	.11	.17	.18	.24	.25	.28	.27	.41	.29	.10	.21	.24	.18	.39	.16	.27	.23	.23
Rank	20	57	43	47	42	2	15	75	66	127	6	77	58	25	27	135	23	73	111	36
rit(3)	.396	.291	.342	.331	.344	.412	.413	.256	.214	.139	.429	.247	.290	.384	.382	.106	.387	.258	.179	.356
s ² (3)	.24	.14	.24	.21	.21	.17	.17	.15	.16	.06	.15	.25	.21	.19	.22	.10	.23	.18	.20	.20
Diff(3)	.60	.83	.61	.70	.71	.79	.79	.81	.80	.93	.81	.56	.71	.74	.66	.89	.63	11.	.72	.72
\$(0,3)	.54	.53	.52	.52	.52	.51	.51	.50	.50	.50	.49	.49	.49	.48	.47	.47	.47	.47	.47	.47
Item	129	87	52	116	75	2	69	42	32	54	64	146	15	120	37	48	131	Э	55	112
Rank	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

Include	Yes																			
R*	65.0	24.0	38.0	80.0	67.5	35.0	44.0	28.0	25.0	81.0	62.5	35.0	37.0	84.0	57.0	87.0	70.5	35.5	51.5	54.5
r _{it(0)}	.198	.129	.194	072	.118	.145	.138	.171	.220	.228	.081	.220	.165	.174	.014	.087	.025	.197	.088	.157
s ² (0)	.25	.20	.24	.16	.24	.07	.18	.25	.21	.25	.17	.23	.24	.25	.19	.13	.13	.22	.23	.19
Diff(0)	.44	.27	.40	.20	.41	.08	.24	.43	.30	.44	.21	.35	.39	.43	.26	.15	.15	.33	.38	.25
Rank	89	9	33	116	06	24	41	80	I	112	74	18	21	114	59	118	84	13	44	49
$r_{it(3)}$.234	.462	.363	.171	.229	.386	.344	.436	.489	.178	.258	.404	.393	.176	.290	.168	.236	.417	.342	.324
s ² (3)	.07	.19	.10	.22	.10	.25	.21	60.	.19	.10	.23	.16	.14	.11	.22	.25	.25	.18	.15	.22
Diff(3)	.93	.75	.89	.66	.88	.46	.70	.90	.75	.89	.64	.79	.84	.87	.68	.54	.54	.76	.81	.66
<u> (0,3)</u>	.46	.46	.46	.45	.44	.44	.44	.44	.43	.42	.42	.42	.42	.41	.41	.40	.40	.40	.40	.40
Item	16	111	59	76	28	136	123	72	132	82	98	104	70	102	106	27	108	65	43	61
Rank	41	42	43	44	45	95	47	48	49	50	51	52	53	54	55	56	57	58	59	60

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

Include	Yes																				
1 <u>%</u>	92.0	49.5	81.5	65.5	48.0	82.0	42.0	75.5	36.5	92.5	38.0	76.5	61.5	76.5	68.0	66.0	64.5	90.5	86.0	76.0	
r _{it(0)}	.264	.020	061	.174	.024	.088	.181	.186	.040	.198	.140	.221	.143	.078	.044	.040	.221	.228	.001	.192	
s ² (0)	.25	.17	.23	.25	.19	.17	.23	.23	.22	.24	.17	.25	.22	.23	.15	.21	.16	.23	.13	.23	
Diff(0)	.56	.21	.35	.51	.26	.21	.36	.34	.32	.61	.22	.43	.32	.38	.18	.29	.20	.64	.15	.36	
Rank	123	37	100	67	31	98	17	83	4	115	5	85	50	62	61	46	52	103	93	54	
rit(3)	.152	.351	.208	.271	.370	.218	.407	.240	.464	.175	.463	.235	.324	.245	.280	.338	.317	.203	.223	.304	
s ² (3)	.04	.24	.18	.08	.23	.24	.18	.19	.20	.03	.24	.15	.21	.19	.25	.23	.25	.04	.25	.21	
Diff(3)	.95	.61	.77	16.	.65	.59	.76	.74	.72	.96	.58	.81	.70	.75	.52	.65	.55	.95	.45	.71	
¢(0,3)	.40	.39	.39	.39	.38	.38	.37	.37	.37	• 36	.36	.35	.35	.35	.35	.35	.34	.33	.33	.32	
Item	18	92	117	21	26	134	68	58	74	63	124	83	44	46	101	94	145	40	9	64	
Rank	61	62	63	64	65	99	67	68	69	70	71	72	73	74	75	76	17	78	62	80	

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

-
P
ne
ā
H
'n
0
9
-
×
G
E
E
E.
A

Rank	Item	<u>∳(0,3)</u>	Diff(3)	s ² (3)	$r_{it(3)}$	Rank	Diff(0)	s ² (0)	r _{it(0)}	12	Include
81	150	.32	.66	.22	.361	34	.32	.22	.284	57.5	Yes
82	14	.32	.49	.25	.247	78	.18	.15	.088	80.0	Yes
83	121	.32	.53	.25	.393	22	.21	.17	.168	52.5	Yes
84	115	.31	.94	.06	.235	88	.65	.23	.213	86.0	Yes
85	63	.31	.59	.24	.224	92	.27	.20	.005	88.5	Yes
86	128	.30	.80	.16	.421	12	.49	.25	.028	49.0	Yes
87	22	.29	.35	.23	.118	131	.11	.10	046	109.0	No
88	110	.29	.47	.25	.338	45	.19	.15	013	66.5	Yes
89	96	.29	.86	.12	.467	ß	.56	.25	.263	46.0	Yes
90	140	.28	.38	.24	.244	80	.13	.12	.121	85.0	Yes
16	73	.28	.45	.25	.195	106	.19	.15	.053	98.5	No
92	138	.28	.56	.25	.323	51	.27	.20	.287	71.5	Yes
93	125	.27	.32	.22	.190	107	.10	60.	.049	100.0	No
94	107	.27	.45	.25	.118	132	.19	.16	.046	113.0	No
95	31	.26	.36	.23	.269	69	.13	.11	.042	82.0	Yes
96	119	.26	.54	.25	.422	10	.27	.20	.218	53.0	Yes
16	149	.26	.48	.25	.262	71	.23	.18	.121	84.0	Yes
98	47	.26	.29	.21	.285	60	60.	.08	089	78.0	Yes
66	99	.25	.70	.21	.222	96	.42	.24	191.	97.5	Yes

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

Include	No	No	Yes	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	
1 <u>%</u>	112.5	105.0	71.0	0.111	95.5	85.0	100.5	108.5	125.5	121.0	69.5	123.5	87.5	113.0	126.5	111.5	125.0	72.5	95.0	0.011	
r _{it(0)}	.189	.008	.128	.019	.118	.338	.253	040	160.	.259	001	.008	.160	016	.083	.085	.052	.143	.116	.175	
s ² (0)	.22	.22	.15	.19	.24	.18	.24	.13	.22	.25	.16	60.	.17	.25	.22	.17	.20	.24	.16	.18	
Diff(0)	.66	.32	.19	.25	.41	.24	.42	.15	.34	.54	.20	.10	.22	.51	.32	.22	.28	.40	.20	.24	
Rank	125	109	40	129	87	65	95	110	143	133	29	136	63	113	139	108	134	28	72	119	
$r_{it(3)}$.147	.188	.345	.125	.235	.272	.222	.179	.038	011.	.372	160.	.275	.177	.075	.189	.109	.377	.261	.167	
s ² (3)	60.	.25	.24	.25	.22	.25	.22	23	.24	.18	.24	.19	.24	.20	.25	.24	.25	.24	.23	.24	
Diff(3)	06.	.57	.41	.48	.66	.47	.67	.35	.58	11.	.41	.26	.42	.72	.53	.42	.48	.60	.38	.42	
<u>♦(0,3)</u>	.25	.23	.23	.23	.23	.23	.23	.23	.22	.22	.22	.21	.20	.20	.20	.20	.19	.19	.19	.18	
Item	113	13	137	62	85	135	38	78	1	50	105	1	5	127	56	67	109	26	148	133	
Rank	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

Include	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
ı*	92.0	123.5	94.0	67.0	114.0	108.0	136.0	127.5	83.0	92.0	116.0	134.5	136.5	104.5	127.5	137.5	140.0	137.0	129.0	110.5	
r _{it(0)}	.004	.148	.062	.218	.174	.215	.057	003	.229	.025	.014	.225	052	.055	.275	.072	022	.115	.040	.129	
s ² (0)	.20	.12	.15	.24	.24	.24	.21	.24	.22	.22	.20	.24	60.	.25	.23	.23	.14	.13	.25	.23	
Diff(0)	.28	.14	.19	.42	.59	.58	.31	.39	.68	.34	.27	.39	.10	.45	.35	.38	.17	.15	.57	.36	
Rank	68	126	99	11	104	91	146	128	38	55	102	138	141	76	121	140	144	137	120	82	
rit(3)	.270	.144	.271	.422	.201	.226	076	.135	.348	.298	.204	.076	.048	.254	.164	.066	037	.085	.166	.243	
s ² (3)	.25	.20	.23	.24	.18	.20	.25	.25	.16	.25	.24	.25	.14	.25	.24	.24	.16	.14	.24	.23	
Diff(3)	.46	.28	.35	.60	.76	.72	.45	.53	.79	97.	.39	.51	.17	.55	.41	.43	.20	.18	.60	.38	
\$(0,3)	.18	.17	.17	.17	.16	.14	.14	.12	.12	.12	.12	.10	.10	60.	.05	.05	.04	.03	.03	.02	
Item	100	130	86	114	122	24	16	20	4	57	84	33	139	25	51	39	103	144	143	60	
Rank	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and ^rit(3).

IX 1 (continued)

×	4	
-	٠	
E	Ì	
2	1	
ä	1	
ρ	1	
•		

R* Include	96.5 No	35.5 No	45.0 No	18.5 No	07.0 No	43.5 No	45.5 No	35.5 No	47.5 No	49.0 No	50.0 No
rit(0)	.240	044 1	006 1	.223 1	.079 1	.073 1	.040	.152 1	.060 1	.033 1	.121 1
s ² (0)	.23	.24	.16	.25	.25	.24	.21	.21	.14	.19	.24
Diff(0)	.63	.58	.20	.54	.51	.39	.30	.30	.17	.26	.58
Rank	53	130	148	64	70	142	145	124	147	149	150
r _{it(3)}	.314	.120	098	.223	.264	.039	044	.148	078	139	176
s ² (3)	.23	.24	.15	.25	.25	.23	.17	.17	.08	.08	.15
Diff(3)	.65	.59	.19	.53	.49	.36	.22	.22	60.	60.	.19
¢(0,3)	10.	.01	01	01	02	04	08	08	11	19	36
Item	п	147	29	118	45	19	30	88	53	80	12
Rank	140	141	142	143	144	145	146	147	148	149	150

 $\frac{*}{R}$ is the average rank of $\phi(0,3)$ and $^{\Gamma}it(3)$.

APPENDIX 2 ITEM ANALYSIS OF TEST 2 FOR 5-LEVEL and 3-LEVEL PERSONNEL

				(Total	N = 530;	3 = 194	4; 5 = 336	~			
Rank	Item	\$(3,5)	Diff(5)	s ² (5)	^r it(5)	Rank	Diff(3)	s ² (3)	$r_{it(3)}$	12	Include
1	80	.21	.77	.18	.403	4	.57	.24	.235	2.5	Yes
2	138	.20	.49	.25	.334	15	.28	.20	.191	8.5	Yes
e	65	.19	.43	.25	.197	63	.24	.18	.220	33.0	Yes
4	131	.16	.33	.22	.264	41	.19	.15	027	22.5	Yes
2	123	.15	.42	.24	.290	29	.27	.20	.195	17.0	Yes
9	104	.15	.40	.24	.368	10	.26	.19	.180	8.0	Yes
2	66	.14	.55	.25	.301	23	.40	.24	.208	15.0	Yes
80	6	.14	.69	.21	.318	19	.55	.25	.283	13.5	Yes
6	3	.14	.31	.21	.211	56	.19	.15	.039	32.5	Yes
10	43	.14	.75	.19	.384	5	.62	.24	.399	7.5	Yes
п	145	.13	.35	.23	.187	70	.23	.18	.149	40.5	Yes
12	129	.13	.50	.25	.188	69	.36	.23	.187	40.5	Yes
13	62	.13	.63	.23	.298	24	.49	.25	.365	18.5	Yes
14	60	.13	.61	.24	.351	12	.48	.25	.254	13.0	Yes
15	22	.13	.68	.22	.330	16	.55	.25	.310	15.5	Yes
16	27	.12	.19	.15	.132	88	.10	60.	.049	47.0	Yes
17	124	.12	.67	.22	.159	80	.55	.25	.151	48.5	Yes
18	147	.12	.35	.23	.112	67	.23	.18	.077	57.5	Yes
19	78	.11	.43	.25	.328	17	.32	.22	.159	18.0	Yes
20	50	.11	.85	.13	.424	2	.76	.18	.371	11.0	Yes

51

 $\star \overline{R}$ is the average rank of $\phi(3,5)$ and $^{r}it(5).$

Include	Yes																			
R*	36.0	44.0	21.5	59.0	58.5	14.5	54.0	76.5	42.0	18.0	44.0	73.0	63.0	49.0	56.0	22.0	56.5	52.5	55.0	35.5
$r_{it(3)}$.279	.140	.263	.311	.032	.279	021	.037	.268	.338	.325	124	025	.225	.119	.314	.206	.172	.206	.114
s ² (3)	.25	.21	.25	.19	.18	.18	.10	.20	.16	.25	.25	•06	.21	.21	.23	.22	.25	.21	.24	.24
Diff(3)	.52	.29	.51	.26	.24	.23	.12	.27	.20	.56	.46	.07	.30	.30	.35	.66	.45	.31	.41	.39
Rank	51	99	20	74	92	e	81	125	55	9	57	114	93	64	11	80	76	67	71	31
$r_{it(5)}$.223	.190	.316	.181	.128	.421	.159	.033	.212	.384	.209	.063	.122	.197	.172	.383	.178	.189	.186	.287
s ² (5)	.23	.24	.24	.23	.22	.22	.15	.23	.20	.23	.25	.10	.24	.24	.25	.19	.25	.24	.25	.25
Diff(5)	.63	.40	.62	.36	.34	.33	.19	.36	.28	.65	.55	.12	.39	.38	.43	.74	.53	.38	.49	.46
<u>∳(3,5)</u>	.11	.11	п.	.10	.10	.10	60.	60.	60.	60.	60.	.08	.08	.08	.08	.08	.08	.07	.07	.07
Item	39	64	116	24	57	130	16	92	14	25	51	45	34	18	48	33	12	121	142	S
Rank	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

 $\frac{*}{\overline{R}}$ is the average rank of $\phi(3,5)$ and $^{\Gamma}it(5).$

Include	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes											
ı*]	40.0	25.5	62.5	61.0	40.0	48.0	56.0	67.5	46.0	80.0	95.0	39.0	50.5	94.5	75.0	67.5	54.5	98.0	45.5	
rit(3)	.332	.366	060.	131	.251	.285	.261	.186	.318	041	.166	.334	.061	.013	.228	.032	.135	006	.377	
s ² (3)	.25	.18	.13	.18	.22	.23	.21	.24	.24	.14	.20	.24	.23	.13	.21	.16	.20	.07	.25	
Diff(3)	.46	11.	.15	.23	.68	.38	.29	.39	.60	.17	.27	.61	.36	.16	.30	.20	.28	.08	.46	
Rank	39	6	82	78	35	50	65	87	43	110	139	26	48	135	95	79	52	138	32	
$r_{it(5)}$.266	.376	.157	.167	.279	.234	.196	.134	.252	.074	-,030	.296	.238	012	.113	.161	.221	026	.287	
S ² (5)	.25	.15	.16	.20	.20	.25	.23	.25	.22	.17	.22	.22	.24	.16	.23	.18	.22	60.	.25	
Diff(5)	.53	.82	.21	.28	.73	.44	.35	.45	.66	.22	.32	.66	.42	.20	.35	.24	.33	.10	.51	
<u> (3,5)</u>	.07	.06	•06	•06	.06	•06	.06	.06	.06	.06	.06	.06	.05	.05	.05	.05	.05	.04	.04	
Item	112	37	84	71	82	109	2	13	83	68	107	134	103	133	106	46	42	1	38	
Rank	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	

 $\frac{*}{\overline{R}}$ is the average rank of $\phi(3,5)$ and ^rit(5).

Include	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yer	
12	61.0	82.0	83.5	42.5	51.0	0.69	92.0	75.5	49.0	58.0	61.5	102.0	58.0	43.5	43.5	91.5	67.0	96.5	89.0	60.0	
$r_{it(3)}$.129	050	.137	.273	.347	.240	.106	.002	.403	.278	.282	.005	.225	.252	.329	.082	.110	053	103	033	
s ² (3)	.24	.14	.23	.25	.25	.24	.22	.14	.24	.21	.23	.13	.23	.25	.24	.15	.24	.20	.16	60.	
Diff(3)	.38	.17	.38	.51	.55	.39	.31	.18	.59	.31	.63	,16	.37	.56	.42	.18	.41	.28	.21	.10	
Rank	62	103	105	22	38	73	118	84	30	47	53	133	44	14	13	108	58	116	100	141	
rit(5)	.198	160.	.085	.311	.274	.181	.054	.136	.287	.238	.216	007	.252	.336	.343	.078	.205	.060	.101	059	
s ² (5)	.24	.16	.24	.25	.24	.24	.23	.16	.23	.23	.22	.15	.24	.24	.25	.17	.25	.22	.18	.11	
Diff(5)	.43	.21	.42	.55	.59	.43	.35	.21	.63	.35	.67	.19	.40	.60	.46	.21	.44	.32	.24	.12	
<u> (3,5)</u>	. 04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	•04	.04	.04	.04	.04	.04	.03	.03	.03	
Item	55	144	105	135	88	74	127	47	59	93	31	99	149	19	69	52	118	4	75	35	
Rank	60	61	62	63	64	65	99	67	68	69	70	11	72	73	74	75	76	11	78	61	

 $\frac{*}{R}$ is the average rank of $\phi(3,5)$ and ^rit(5).

Include	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	
ا %ا	101.0	70.5	44.5	90.5	100.5	95.5	103.5	111.5	106.5	64.5	82.5	93.5	46.5	96.5	106.5	122.5	113.0	71.0	119.0	66.0	
$r_{it(3)}$.109	.071	.373	.137	.218	.113	.081	181	.432	.404	.264	143	.419	.094	.074	160	.002	.301	011	.252	
s ² (3)	.14	.14	.25	.15	.19	.22	.20	.05	.24	.24	.19	.13	.15	.22	.22	.13	.15	.21	.22	.23	
<u>Diff(3)</u>	.17	.17	.54	.19	.26	.32	.27	.05	.61	•58	.26	.16	.81	.32	.34	.15	.19	.70	.32	.63	
Rank	122	60	7	98	117	106	121	136	25	40	75	96	1	66	119	150	130	45	140	33	
$r_{it(5)}$.041	.202	.384	111.	.057	.082	.043	015	.297	.265	.180	.113	.434	.105	.054	281	.010	.246	043	.284	
s ² (5)	.16	.16	.24	.17	.21	.23	.21	.06	.23	.24	.20	.14	.14	.22	.23	.14	.16	.21	.22	.23	
Diff(5)	.20	.20	.57	.21	.29	.35	.30	.07	.63	.61	.28	.18	.83	.34	.36	.16	.20	.71	.33	.64	
<u> (3,5)</u>	.03	.03	.03	.03	.03	.03	.03	.03	.03	.02	.02	.02	.02	.02	.02	.02	.02	.01	.01	.01	
Item	85	136	90	20	26	108	139	32	89	86	111	101	76	117	11	67	119	67	125	114	
Rank	80	81	82	83	84	85	86	87	88	89	60	16	92	93	94	95	96	26	98	66	

 $\frac{*}{\overline{R}}$ is the average rank of $\phi(3,5)$ and ^rit(5).

7	2
;	ž
1	Ξ
4	5
1	E
1	ວັ
-	-
•	V
-	4
1	-1

1	\square	
i	-	
1	4	
1	Ŧ	
í	۰.	
J	-	
1	- 4	
	-	
1	~	

Include	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
12	74.5	80.0	87.0	102.0	114.0	124.5	70.0	125.0	115.5	96.0	128.0	106.5	127.0	62.0	66.0	102.0	80.0	122.0	72.5	121.5	
rit(3)	.288	.169	.144	.152	006	015	.286	109	.085	.086	.078	.051	089	.236	.192	.142	.269	.162	.275	.187	
s ² (3)	.24	.24	.22	. 21	.15	.15	.15	.19	.17	.21	.19	.21	.14	.20	.25	.24	.24	.25	.12	.17	
Diff(3)	.62	.58	.34	.29	.19	.19	.82	.26	.21	.30	.25	.30	.18	.72	.54	.41	.42	.48	.86	.22	
Rank	49	59	72	101	124	144	34	143	123	83	146	102	142	11	18	89	54	127	27	126	
$r_{it(5)}$.234	.202	.184	960.	.034	085	.283	079	.039	.147	093	.094	070	.368	.321	.131	.214	.022	.294	.024	
s ² (5)	.23	.24	.23	.21	.16	.15	.15	.19	.17	.21	.18	.21	.14	.20	.25	.24	.24	.25	.12	.16	
Diff(5)	.63	.59	.34	.30	.19	.19	.82	.26	.21	.30	.24	.29	.17	.71	.53	.40	.41	.47	.85	.21	
\$(3,5)	.01	.01	.01	.00	00.	.00	00.	.00	.00	.00	00.	.00	01	01	01	01	01	01	01	01	
Item	30	29	100	54	128	62	49	53	28	120	94	63	98	81	95	146	140	61	40	23	
Rank	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	

 $\frac{*}{R}$ is the average rank of $\phi(3,5)$ and ^rit(5).

Include	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
1 <u>%</u> 1	102.5	133.0	75.0	82.5	116.5	119.0	108.5	118.5	121.5	109.5	118.5	131.5	140.5	85.0	134.0	78.0	91.0	0.66	103.0	85.0	
r <u>it(3)</u>	.040	.055	.471	.349	.094	.102	.080	081	.004	.058	.167	.093	129	.222	068	.345	.187	.172	.100	.003	
s ² (3)	.22	.21	.23	.25	.18	.24	.24	.20	.23	.18	.19	.07	.14	.24	.16	.25	.24	.25	.24	.14	
Diff(3)	.32	.30	.65	.47	.24	.41	.39	.27	.35	.23	.25	.08	.16	.39	.20	.52	.40	.54	.43	.16	
Rank	85	145	28	42	109	113	91	112	115	60	107	132	149	37	134	21	46	61	68	131	
$r_{it(5)}$.134	087	.291	.255	.075	.063	.130	.072	.062	.131	.079	007	131	.275	007	.315	.241	.201	.188	.008	
s ² (5)	.21	.21	.23	.25	.17	.24	.23	.18	.22	.16	.17	.06	.12	.23	.13	.25	.23	.25	.23	.11	
<u>Diff(5)</u>	.31	.29	.63	.45	.22	.39	.36	.24	.32	.21	.22	.06	.14	.35	.16	.47	.35	.49	.37	.12	
<u> </u>	02	02	02	02	02	03	03	03	03	03	03	03	04	04	04	05	05	05	05	06	
Item	137	1	58	73	11	17	143	126	41	70	110	9	132	87	148	56	113	122	96	10	
Rank	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	

57

 $\frac{*}{R}$ is the average rank of $\phi(3,5)$ and ^rit(5).

¢

Include	No	No	No									
1*	113.0	135.0	145.0	127.5	136.5	146.0	125.0	120.5	142.5	92.5	135.0	
rit(3)	.256	.035	115	127	.086	117	.107	.085	.051	.272	.153	
s ² (3)	.24	.20	.14	.15	.21	.14	.22	.21	.17	.24	.25	
Diff(3)	.39	.28	.17	.18	.31	.17	.32	.29	.22	.61	.46	
Rank	86	129	148	111	128	147	104	94	137	36	120	
$r_{it(5)}$.134	.014	129	.073	.016	104	.086	.120	017	.277	.053	
S ² (5)	.22	.17	.11	.11	.18	.10	.19	.17	.11	.24	.18	
<u>Diff(5)</u>	.33	.22	.12	.13	.24	.12	.25	.22	.12	.43	.24	
<u> (3,5)</u>	06	06	06	07	07	08	08	08	13	18	23	
Item	102	44	72	15	21	141	36	150	115	80	16	
Rank	140	141	142	143	144	145	146	147	148	149	150	

 $\frac{*}{R}$ is the average rank of $\phi(3,5)$ and ^rit(5).

T DEPCONNET i TOGT. ANATVETS OF

ŕ

		Include	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
		12	2.5	12.5	27.0	11.0	3.0	6.5	24.5	8.0	7.5	23.0	18.5	21.5	13.0	45.0	8.5	22.5	38.0	28.0	25.5	40.5	
LENJOURN		$r_{it(5)}$.235	.227	.355	.326	.243	.229	.277	.219	.260	.260	.082	.281	.208	.154	.321	.202	.391	.243	.127	.156	
ט-רבעבר	128-	s ² (5)	.19	.15	.18	.17	.10	.24	.24	.21	.20	.22	.12	.25	.10	.15	.23	.24	.24	.23	.14	.20	
LEVEL AND	= 283)	Diff(5)	.26	.19	.24	.22	.12	.39	.60	.30	.27	.33	.13	.57	.11	.18	.36	.42	.43	.63	.17	.27	
-/ VOJ	: 336; 7	Rank	4	23	51	18	ı	7	42	80	9	36	26	31	13	76	3	29	59	38	32	61	
C 1031 10	: 619; 5 =	$r_{it(7)}$.434	. 321	.228	.340	.529	.377	.245	.364	. 385	.258	. 305	.280	.350	.160	.447	.286	.199	.255	.280	.195	
CICITW	Cotal N =	s ² (7)	.19	.22	.21	.23	.25	.17	.06	.22	.23	.21	.25	.11	.24	.25	.22	.20	.20	.10	.25	.25	
A MALL C	(1	Diff(7)	.74	.66	.71	.65	.51	62.	.94	69.	.65	.71	.46	.87	.40	.48	.68	.73	.72	. 89	.43	.55	
AFFENDIA		\$(5,7)	.48	. 48	.47	. 43	.43	.40	.39	.38	.38	.38	.36	.34	.33	.32	.32	.31	.30	.30	.29	.29	
		Item	101	25	18	35	57	122	45	63	60	17	47	2	40	19	44	39	113	81	31	103	
		Rank	1	3	3	4	5	9	1	80	6	10	11	12	13	14	15	16	11	18	19	20	

 $\frac{*}{R}$ is the average rank of $\phi(5,7)$ and $^{r}it(7)$.

Include	Yes	Yes	Yes	Yes	Yes	Yes															
121	13.0	51.0	31.0	20.5	24.5	35.0	19.5	57.5	43.0	20.5	29.5	33.0	27.0	48.5	59.5	45.0	35.0	23.5	27.5	27.0	
$r_{it(5)}$.210	.199	.249	.205	.266	.251	.140	.166	.234	.171	.329	.284	.098	.225	• 00 •	.182	.123	.283	.189	.206	
s ² (5)	.21	.25	.16	.19	.25	.19	.25	.23	.25	.18	.20	.24	.14	.19	.24	.24	.24	.23	.25	.23	
D1ff(5)	.30	. 44	.20	.25	.53	.74	.44	.35	.56	.24	.72	.59	.17	.26	.40	.38	.39	.64	.51	.37	
Rank	5	80	39	17	24	44	12	87	57	11	28	34	21	63	84	54	33	6	16	14	
$r_{it(7)}$.412	.146	.251	.340	.311	.244	.352	.129	.202	.356	.291	.272	.329	.189	.134	.216	.272	.360	.342	.348	
s ² (7)	.24	.20	.25	.25	.17	.06	.22	.24	.17	.25	60.	.16	.23	.25	.24	.24	.24	.14	.20	.24	
Diff(7)	.58	.72	.46	.51	.78	.94	.68	.59	.78	.46	.90	. 80	.37	.47	.62	.59	.60	. 83	.71	.58	
\$(5,7)	.28	.28	.27	.27	.27	.26	.24	.24	.23	.23	.23	.23	.22	.22	.22	.22	.21	.21	.21	.21	
Item	80	43	72	115	62	20	16	87	126	22	29	68	53	11	146	63	80	123	73	21	
Rank	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	

 $\star \overline{R}$ is the average rank of $\phi(5,7)$ and ^rit(7).

	lude																				
10	Inc	Yes																			
i	R*	30.0	87.5	29.5	53.0	45.5	36.5	57.0	58.0	82.5	64.0	53.0	66.5	62.0	49.5	29.0	64.5	66.0	76.5	44.5	51.5
	it(5)	.320	.102	.243	.205	.195	.230	.116	.264	062	.106	.212	.024	.323	.147	.196	.290	.192	069	.139	.388
2	S ⁻ (5)	.24	.22	.25	.22	.10	.22	.22	.11	.07	60.	.22	.13	.15	.18	.25	.24	.20	.21	.19	.23
	Diff(5)	.43	.34	. 44	.32	.12	.33	.32	. 88	.07	.10	.32	.15	. 82	.24	.53	. 39	.28	.31	.26	.64
	Rank	19	133	15	62	97	27	67	68	116	78	55	81	11	45	Э	73	75	95	30	43
r.	it(7)	.337	022	.344	.189	.243	.299	.185	.183	.041	.156	.213	.144	.176	.243	.438	.171	.163	.106	.281	.244
2	S ⁻ (7)	.23	.25	.23	.25	.20	.25	.25	.02	.15	.18	.25	.21	.07	.24	.21	.25	.24	.25	.24	.18
	Diff(7)	.64	.54	. 65	.51	.27	.51	.49	.98	.19	.24	.49	.30	.93	.40	. 69	.55	.42	.45	.41	.77
	¢(5,7)	.21	.21	.21	.20	.19	.18	.18	.18	.18	.18	.18	.17	.17	.17	.16	.16	.15	.15	.15	.15
,	Item	14	119	129	75	144	41	131	27	10	13	83	145	136	109	64	104	69	86	127	99
	Rank	41	42	43	44	45	46	47	48	65	50	51	52	53	54	55	56	57	58	59	60

 $\star \overline{R}$ is the average rank of $\phi(5,7)$ and ^rit(7).

7	Ż
-	5
\$	
Ŧ	3
5	3
2	5
-	-
~	2

2	×	ï	
۲	-	۱	
f		1	
ć		i	
ŕ	١		
۴	1	1	
•	9	i.	

Include	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes												
I <u>*</u>	73.5	57.0	63.5	65.0	50.0	53.5	75.0	80.0	69.0	45.0	93.0	101.0	49.0	91.0	92.0	93.5	66.5	69.0	68.5	65.0	
$r_{it(5)}$.078	.138	.082	.116	.126	.203	166	.173	.046	.321	.092	.066	.195	.116	.069	.138	.195	.103	.187	.118	
s ² (5)	.13	.22	.14	.25	.17	.24	.23	.25	.18	.16	.22	.18	.25	.21	.12	.11	.25	.22	.25	.17	
Diff(5)	.16	. 68	.17	.48	.21	.58	.35	.55	.24	.81	.34	.24	.50	.29	.15	.12	.54	.34	.49	.22	
Rank	86	52	64	99	35	41	83	92	69	20	115	130	25	108	109	111	56	60	58	50	
rit(7)	.131	.221	.189	.186	.270	.245	.142	.118	.177	.332	.042	.001	.307	.060	.052	.049	.204	.196	.200	.230	
s ² (7)	.20	.16	.21	.24	.22	.21	.25	.22	.23	.10	.25	.22	.24	.24	.18	.16	.23	.25	.24	.21	
Diff(7)	.28	.81	.29	.62	.33	.70	.47	.67	.35	68.	.45	.34	.61	.40	.23	.20	.65	.44	.58	.30	
\$(5,7)	.14	.14	.14	.14	.13	.13	.13	.12	.12	.12	.12	.12	.11	.11	.11	.11	.10	.10	60.	60.	
Item	12	51	117	28	92	56	4	59	49	55	121	16	9	52	11	54	149	138	102	107	
Rank	61	62	63	64	65	99	67	68	69	70	11	72	73	74	75	76	11	78	61	80	

* \overline{R} is the average rank of $\phi(5,7)$ and $^{r}it(7)$.

	Include	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No
APPENDIX 3 (continued)	12	77.5	46.0	52.5	68.5	90.5	78.0	110.5	88.0	92.5	63.5	102.5	94.5	79.0	96.5	0.06	115.5	111.5	104.0	69.5	122.5
	rit(5)	.124	.220	.262	.279	.115	.264	.128	077	034	.084	.037	.150	.007	.023	.026	020	.179	.151	.051	.071
	s ² (5)	.19	.22	.25	.24	.20	.23	.23	.10	.15	.24	.18	.25	.17	.10	.08	.10	.24	.22	.23	.16
	Diff(5)	.26	.67	.57	.40	.27	.37	.36	.12	.18	.57	.23	.48	.21	.11	60.	.11	.58	.34	.35	.20
	Rank	74	10	22	53	96	70	134	88	100	37	114	67	65	66	85	135	126	110	40	145
	$r_{it(7)}$.164	.358	. 327	.219	.106	.176	024	.127	060.	.257	.042	.106	.188	.103	.133	024	.011	.050	.246	086
	s ² (7)	.23	.19	.23	.25	.23	.25	.25	.14	.18	.23	.21	.25	.20	.13	.11	.13	.23	.24	.24	.18
	Diff(7)	.34	.75	.65	.48	.35	.44	.43	.17	.24	.64	.29	.55	.27	.16	.12	.15	.64	.39	.40	.23
	\$(5,7)	60.	60.	.08	.08	.08	.07	.07	.07	.07	.07	.07	.07	.07	.07	.06	.06	.05	.05	.04	.04
	Item	140	143	108	62	34	82	141	15	85	24	137	32	6	88	46	110	114	Э	132	128
	Rank	81	82	83	84	85	86	87	88	89	06	16	92	93	94	95	96	16	98	66	100

 $\frac{*}{R}$ is the average rank of $\phi(5,7)$ and ^rit(7).

Include	No	Yes	No	No	No	Yes	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	No	No
1×1	111.5	92.0	103.5	111.0	114.5	76.5	125.0	78.0	114.0	94.5	91.5	108.5	102.0	104.0	108.5	114.0	117.0	83.5	112.5	110.5
$r_{it(5)}$.100	.181	.043	.094	125	.112	.137	038	.043	.201	.151	069	.160	.051	.029	012	.008	043	080	.043
s ² (5)	.16	.24	.13	.21	.14	.25	.23	.20	.21	.23	.25	.08	.21	.20	.20	.14	.14	.12	.10	.22
Diff(5)	.20	.43	.16	.29	.17	. 49	.63	.27	.31	.36	.47	60.	.31	.27	.28	.17	.17	.13	.11	.32
Rank	122	82	104	118	124	47	143	48	119	62	72	105	16	94	102	112	117	49	106	101
$r_{it(7)}$.020	.142	.078	.033	.015	.237	068	.233	.030	.150	.174	.075	.121	.109	.080	.048	.038	.232	.063	.086
s ² (7)	.18	.25	.15	.22	.15	.25	.23	.20	.22	.23	.25	.08	.21	.20	.20	.14	.14	.11	60.	.21
Diff(7)	.23	.46	.18	.32	.18	.51	.65	.28	.33	.37	.48	60.	.31	.27	.28	.17	.17	.13	.10	.31
<u>♦(5,7)</u>	.04	.03	.03	.03	.02	.02	.02	.02	10.	.01	.01	.00	.00	.00	.00	.00	01	01	02	02
Item	23	135	70	112	84	105	98	120	74	7	33	57	148	106	48	125	100	147	94	42
Rank	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
APPENDIX 3 (continued)

Include	No																			
R*	132.5	129.0	106.0	126.0	136.5	127.5	138.0	133.5	111.0	121.5	136.5	139.0	120.0	129.5	119.0	138.0	137.5	107.5	114.5	131.5
rit(5)	.040	.076	.156	.176	054	.022	043	156	077	015	.024	006	.075	.040	.225	.089	042	009	.132	.100
s ² (5)	.16	.24	.20	.25	.05	.24	.13	.04	.19	.15	.22	.21	.14	.15	.24	.12	.07	.12	.25	.25
Diff(5)	.20	.58	.29	.44	.05	.40	.16	.04	.26	.18	.34	.30	.17	.19	.60	.15	.07	.15	.44	. 49
Rank	144	136	89	128	147	129	149	139	93	113	142	146	107	125	103	140	138	17	06	123
$r_{it(7)}$	077	030	.126	.005	143	.005	180	041	.110	.047	065	121	.060	.014	.079	052	034	.156	.123	.016
$\frac{s^2(7)}{s^2(7)}$.15	.25	.19	.24	.04	.23	.12	.03	.17	.13	.21	.19	.11	.12	.25	60.	.03	.08	.23	.24
Diff(7)	.18	.56	.27	.41	.04	.37	.14	.03	.22	.15	.29	.25	.12	.14	.52	.10	.04	60.	.36	.41
\$(5,7)	02	02	02	03	03	03	03	03	04	-,05	05	06	07	07	08	08	08	08	08	09
Item	89	36	71	78	26	76	95	50	66	60	139	124	118	96	142	37	67	133	130	65
Rank	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140

 $\frac{*}{\overline{R}}$ is the average rank of $\phi(5,7)$ and ^rit(7).

65

APPENDIX 3 (continued)

Include	No									
1×1	131.0	145.0	135.0	140.5	138.5	122.0	148.5	134.0	144.5	140.5
$r_{it(5)}$.081	.007	074	.073	.128	012	048	038	009	065
S ² (5)	.22	.08	.11	.23	.25	.20	.19	.16	.18	.25
Diff(5)	.67	60.	.13	.35	.43	.28	.26	.20	.24	.47
Rank	121	148	127	137	132	98	150	120	141	131
$r_{it(7)}$.020	156	.005	034	009	.106	221	.022	057	003
s ² (7)	.24	.04	•06	.18	.21	.13	.12	60.	.08	.08
Diff(7)	.58	.04	.06	.24	.31	.16	.14	.10	60.	.08
¢(5,7)	09	10	12	12	13	14	15	15	20	42
Item	30	61	38	2	58	111	150	134	116	I
Rank	141	142	143	144	145	146	147	148	149	150

AU.S. GOVERNMENT PRINTING OFFICE: 1978-771-122/5

 $\frac{*}{R}$ is the average rank of $\phi(5,7)$ and ^rit(7).

66