1993

Examining differences between circular and rectangular item response answer sheets upon the armed services vocational aptitude battery subtests

Mark David Johnson
Loyola University Chicago

Examining Differences Between Circular and Rectangular Item Response Answer Sheets Upon the Armed Services Vocational Aptitude Battery Subtests

Thesis Submitted

to the Faculty of the School of Education

of Loyola University of Chicago

By

Mark David Johnson

In Partial Fulfillment of the Requirements

for the Degree

of

Master of Arts

January, 1993

Mark David Johnson

Loyola University of Chicago

EXAMINING DIFFERENCES BETWEEN CIRCULAR AND RECTANGULAR ITEM RESPONSE ANSWER SHEETS UPON THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY SUBTESTS

The intent of the present study was to examine for differences noted upon the Armed Services Vocational Aptitude Battery (ASVAB) subtests as a result of different answer sheet formats, including the variables of service recruit education level, ability level, and ethnicity. Two answer sheet formats were used: an answer sheet with a vertical rectangular response area for each item option of a subtest question; and an answer sheet with a circular response area for each item option of a subtest question. Multivariate analyses of variance demonstrated significant divergence between the two formats on two speeded subtests within the ASVAB; these results conformed with previous research. Additionally, other subtests demonstrated slightly less, but nonetheless significant, differences between the two answer sheet formats; these results were not previously demonstrated. Significant differences between the levels of recruit education level, ability and ethnicity variables were demonstrated for most of the ASVAB subtests, after examining appropriate post-hoc measures of significance and multiple regression analyses.

TABLE OF CONTENTS

Page

VITA

The author, Mark David Johnson, was born on June 14, 1963 in Joliet, Illinois.

He attended Augustana College of Rock Island, Illinois from September, 1980 through May, 1984, earning a Bachelor of Arts degree in the discipline of psychology. From September, 1984 to August, 1987, he worked at the Will County Regional Superintendent of Schools Office as a Computer Operator, Assistant Director of a state-funded project and an Administrative Aide to the Superintendent.

In November, 1987, Mr. Johnson accepted a position as a Personnel Psychologist within the Plans/Research/Programs Division of the Testing Directorate, Headquarters United States Military Entrance Processing Command and is presently employed there.

LIST OF TABLES

Tables | Page

LIST OF TABLES

Tables Page

INTRODUCTION

The Armed Services Vocational Aptitude Battery (ASVAB) is a standardized aptitude battery used by the Department of Defense (DoD) to measure the vocational abilities of recruits who enlist into the United States Armed Services. The services use the ASVAB to gauge the recruits' abilities in verbal, mathematical, scientific, technical and industrial knowledge to identify available job opportunities suitable for each recruit within their services. The test battery is shown in Table 1.

Table 1: Subtests, Number of Items, and Testing Times for the ASVAB.

Subtest (Abbreviations)	Number of Items	Time (Minutes)
General Science (GS)	25	11
Arithmetic Reasoning (AR)	30	36
Word Knowledge (WK)	35	11
Paragraph Comprehension (PC)	15	13
Numerical Operations (NO)	50	3
Coding Speed (CS)	84	7
Auto/Shop (AS)	25	11
Math Knowledge (MK)	25	24
Mechanical Comprehension (MC)	25	19

Table 1: Subtests, Number of Items, and Testing Times for the ASVAB (Cont'd).

Subtest (Abbreviations)	Number of Items	Time (Minutes)
Electronics Information (EI)	20	9

The Verbal composite (VE) is also considered a subtest; it is the sum of the raw number-right scores for Word Knowledge and Paragraph Comprehension (VE=WK+PC). Service specific composite scores are generated from subtest standard scores and also assist to provide job classification information for service recruits. Likewise, academic and vocational composite scores generated from the administration of the ASVAB through the DoD Student Testing Program provide vocational or career exploration information to high school students undecided about their futures after high school.

In 1992, the format of the current answer sheet used to administer the ASVAB will be changed to support replacement of the optical mark reader (OMR) system currently used to score ASVAB answer sheets. Presently, the ASVAB answer sheet format provides, for each question, item response options bordered by parallel vertical rectangular lines; this format is a rectangular response format. This type of answer sheet is obsolete for use with the new OMR systems to be acquired; therefore, a new answer sheet format providing item response options for each question represented by a closed circle, and compatible with current technology, has been developed for use with the ASVAB. This format is a circular response format.

According to prior research, score differences associated with use of a circular response format versus a rectangular response format are likely to occur. A study by Valentine and Cowan (1974) revealed that answer sheet format was a source of number-right score differences associated with performance upon the same test.

Sims and Maier (1983) investigated ASVAB subtest scores during a comparison analysis of scores obtained from the 1980 reference population used by the National Opinion Research Center (NORC) (McWilliams, 1980) to renorm the ASVAB and samples of male military applicants and recruits who operationally completed the ASVAB. When using general ability as a control variable, they showed the military applicants scored 3.01 raw score points higher on CS and 1.14 raw score points higher on NO than the 1980 reference population. CS and NO are speeded subtests; both have a large total of items to be answered in a short time period.

Earles, Guiliano, Ree and Valentine (1983) further examined the score differences noted by Sims and Maier and cited three possible rationales for their cause. First of all, the differences could reflect true aptitude differences between the norming and military (operational) groups. Secondly, differences could have arisen because answer sheets with different formats were used for norming and operational purposes. Lastly, differences in test administration procedures were considered a possible cause of the score differences.

Earles, et al. considered the second hypothesis and investigated whether or not the answer sheet format differences accounted for the score differences. They conducted a study using 512 male Air Force recruits as participants. Half the participants was randomly given the ASVAB test form and circular response answer sheets used in the 1980 NORC norming study. The other half of the sample was given the current operational, rectangular response answer sheets and a current operational ASVAB test form to complete. The directions for the NORC answer sheet specified that each item response area should be completely filled in, considered time consuming by Earles, et al. (1983). The operational answer sheets required filling in a vertical item response area, which could be completed more rapidly (Earles, et al., 1983). Also, differences were evident between the organization of subtest item response grids upon the two answer sheets. On the operational sheet, item response grids corresponded to the organization of subtest questions upon the operational ASVAB form; the NORC answer sheet had no correspondence between the layout of item response grids and the respective subtest item order within the ASVAB norming test form used.

The results of Earles, et al.'s study replicated those of Sims and Maier (1983). They found raw score differences of 3.61 points for NO and 1.48 points for CS; with the group using the operational, rectangular response answer sheet scoring higher than those using the NORC, circular response answer sheet. They concluded differences in answer sheet formats and ASVAB subtest item organization layouts were the probable causes of the score differences (Earles, et al., 1983).

In order to generalize the above results for the whole operational testing population (as previous studies used only military recruits as subjects), Ree and Wegner (1990) examined score differences for the NO and CS subtests. They analyzed results from a study examining two equivalent groups of 4,299 ASVAB test applicants. Both groups were randomly given abbreviated answer sheets displaying the subtest item response formats for NO and CS. One answer sheet displayed the NORC, circular response format for the two tests; the other sheet, the operational, rectangular response areas. Ree and Wegner (1990) confirmed use of the rectangular response answer sheet resulted in higher mean scores for NO and CS than use of the circular response sheet, with a 3.19 mean raw score difference on NO, and a 1.34 mean raw score difference on CS displayed. Furthermore, they stated answer sheet effects were the likely reason for differences between the NO and CS scores representative of both item response format groups in previous analyses.

Bloxom, McNulty, Branch, Waters, Barnes and Gribben (1990) analyzed the answer sheet effects cited above, using the operational, rectangular response answer sheet and the circular response sheet developed for use with the new OMR systems. They used randomly equivalent groups of active military recruits with each answer sheet ($N_{Rectangular\ Sheet}=3148$ and $N_{Circular\ Sheet}=3160$) and examined all 10 ASVAB subtests for differences. They confirmed significant differences with univariate T-score statistics between the answer sheet formats, with those using the rectangular response sheets scoring 2.88 raw score points higher on NO and 1.64 raw score points higher on CS than those using the circular response sheet. No significant differences were cited for the other eight subtests (defined as power subtests, as the total of items to be answered per time period is much lower than for NO and CS).

Based on the above studies, it is expected differences will occur between the two answer sheet formats on NO and CS. The power subtests, though not showing significant differences in prior studies, might show previously unconsidered differences under a different method of analysis.

The goal of this study is to objectively determine if significant differences exist among the independent variables of answer sheet format, ability, education level, and ethnicity as reflected by the ASVAB subtest scores through using multivariate analysis of variance (MANOVA). The ten subtests will be the dependent variables for the MANOVA procedures, and the two answer sheet formats will be randomly distributed to two recruit groups using an equivalent groups distribution method. Independent recruit criteria variables of recorded ability, education level at enlistment and ethnic background will be combined with the answer sheet format variable to explore for any effects unique to one of the variables or any possible interaction which might occur. Univariate measures of significance will be used to explore for significant differences particular to the independent variables and the amounts of significant variance associated with the subtests when each independent variable is considered.

DESIGN

This data used in this study was collected in the spring and summer of 1990. Approximately 6400 military recruits, completing the initial portion of basic training for active duty, were used in the study at Air Force, Army, Marine Corps and Navy Recruit Training Centers. For data collection using the circular and rectangular response answer sheet formats, the recruits were administered ASVAB 13c, an equivalent test form to the reference form used during the ASVAB renorming, on a non-operational basis (i.e., the scores would not be included in the recruit's permanent record for training or job requirements). Using an equivalent groups distribution scheme, the two answer sheet formats were randomly given to all test subjects, halving the total sample into two subgroups of about 3200 subjects each. The data collected was examined using MANOVA procedures for significant differences related to the independent variables as reflected in the subtest scores for each MANOVA procedure. In the three procedures, the answer sheet format variable was paired with the recruit ability, ethnicity and education level variables respectively.

METHOD

The subjects of the study were active duty recruits in early stages of basic training at the Recruit Training Centers (RTC) for the four armed services. Data was collected from these recruits during April, May and June, 1990. A breakdown of service representation, date of testing and answer sheet format associated with study participants is provided in Table 2 below.

Table 2: Applicant Breakdown by Testing Location, Date, and Answer Sheet Type.

Testing Location	Rectangular Answer Sheet	Circular Answer Sheet	Total
Army: Fort Jackson April 2 - May 25	1379	1375	2754
Navy: San Diego RTC April 2 - July 2	909	914	1823
Air Force: Lackland AFB April 2 - May 4	521	522	1043
Marine Corps: San Diego April 30 - May 11	393	392	785
Total	3202	3203	6405

The recruits were tested in groups based upon their availability on scheduled testing days.

Both samples of applicants were given two pencils, an

ASVAB test booklet, scratch paper and one of the two answer sheets. Subsequently, they were instructed to complete the answer sheet, providing the following information: name, date of test, social security number, ASVAB test form, sex, education level, service and component, test site and ethnic group. After finishing this task, the recruits completed the ASVAB subtests, following the standard procedure used for operationally administering the ASVAB.

Once the testing sessions were completed, the answer sheets were collected and sent weekly to a government contractor for scoring. The contractor personnel checked the sheets for abnormalities (i.e., stray marks) and prepared them for scoring on the appropriate OMR systems for the circular or rectangular response sheets. The answer sheets were then rescanned using two different OMR systems to examine for machine-related differences. All discrepancies between the two separate scoring runs for each answer sheet format on the appropriate OMR systems were examined and resolved before the data was used for analysis purposes.

Lastly, all individuals with ASVAB subtest scores falling out of the valid subtest score ranges were excluded (ie. range for GS is 1-25, any score not within that range was omitted) from the analysis. Also, since the item ranges for the subtests varied from 15 to 84 test items (see Table 1), the subtest scores for each applicant from each answer format group were transformed into standard Z-scores to provide a valid basis of comparison among the ten subtests (see Hays, 1981).

MANOVA procedures were used to examine the data collected based on the circular and rectangular answer sheet formats and the recruit criteria variables. The subtest scores were examined for any overall significant differences associated with combining the answer sheet format variable and each of the recruit criteria variables separately, and any interaction revealed as significant from these pairings. This method of analysis was chosen for two principal reasons. Use of MANOVA procedures allowed for control of excessive inflation of experimentwise Type I and Type II error rates and the notable decrease in the power of the analysis associated with examining multiple dependent variables through univariate analysis methods (Hasse & Ellis, 1987). Furthermore, many of the ASVAB subtests are intercorrelated, as revealed in Table 3. If univariate analysis were used to examine each subtest as a dependent variable in this experiment, one would assume all intercorrelations between the subtests equal zero. Information collected through MANOVA procedures, vice univariate analysis, allows the researcher to account for intercorrelations among the dependent variables and

Table 3: Subtest Pearson Correlations

	Subtests									
	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI
GS	1.00									
AR	.51	1.00								
WK	.63	.48	1.00							
PC	.45	.50	.55	1.00						
NO	.08	.32	.14	.26	1.00					
CS	.11	.31	.16	.31	.61	1.00				
AS	.52	.36	.37	.28	-.01	.02	1.00			
MK	.42	.68	.39	.43	.37	.35	.16	1.00		
MC	.57	.51	.43	.38	.07	.13	.65	.38	1.00	
EI	.63	.46	.50	.38	.03	.05	.64	.32	.63	1.00

examine for significant differences associated with each independent variable that might erroneously be omitted, leading to inaccurate conclusions about possible sources of significant differences within each independent variable.

Three recruit subject criteria, displayed in Table 4, were included with the answer sheet format variable to examine for possible differences upon the multivariate linear composite. Two-way MANOVA procedures were used examining education, ability and ethnic background respectively with the answer sheet format variable. Three categories of education level (non-high school graduates, high school graduates, and high school graduates with further higher education) were first combined with the answer sheet format variable. Secondly, ability as measured on the recruit's enlistment ASVAB scores was considered with the answer sheet format variable. This

Table 4: Recruit Subject Breakdowns for Education Level, Ability Level and Ethnic Background by Answer Sheet Format.

	Rectangular Answer Sheet		Circular Answer Sheet	
	Cases	%	Cases	%
Education Level				
Non-High School Graduate	416	13.4	397	12.8
High School Graduate	1823	58.6	1871	60.1
High School Graduate w/ Further Ed.	870	28.0	845	27.1
Total	3109	100.0	3113	100.0
Ability Level				
Low Ability	974	31.2	1025	32.7
Medium Ability	1158	37.1	1138	36.3
High Ability	988	31.7	970	31.0
Total	3120	100.0	3133	100.0
Ethnic Background				
American Indian	35	1.1	37	1.2
Hispanic	251	8.2	249	8.0
Asian	63	2.0	64	2.1
Black	667	21.7	674	21.6
White	2018	65.6	2052	65.9
Other Ethnic	43	1.4	38	1.2
Total	3077	100.0	3114	100.0

criterion was defined by the recruit's standard score (Z-scores) total achieved upon his enlistment ASVAB subtests divided by 10, the number of subtests; in other words, it was the mean Z-score for the recruit's enlistment ASVAB standard scores. The ability variable was divided into low, medium and high groups, using +/- one-half of the group's standard deviation about the mean to determine the appropriate group cut points. Six categories of ethnicity (American Indian, Hispanic, Asian, Black, White, and Other Ethnic) were used with the answer sheet format variable in the third procedure.

RESULTS

The first two-way MANOVA procedure examined the combination of the answer sheet format and recruit education level variables. The overall MANOVA hypothesis considered was whether the answer sheet formats (rectangular and circular item responses) and the three recruit education levels (non-high school graduates, high school graduates, and high school graduates with some college education) showed significant differences in their mean vectors as reflected by the ASVAB subtest scores, and if any evident significant interaction occurred between the variables. If the omnibus multivariate tests used showed significant differences were found, then follow-up tests would be run to see which subtests were associated with those differences identified by the overall MANOVA procedure. Assumptions to consider are that the groups are random samples from the available recruit population with the same variance; the ten ASVAB subtests have a multivariate normal distribution; and similar variance-covariance matrices exist for the two recruit groups within this portion of the study. For this analysis, the homogeneity of variance assumption was not fulfilled, regardless of the transformation of the subtest raw scores scores into Z-scores.

Three omnibus MANOVA significance tests were used to examine the overall hypothesis: Pillai's Trace, Hotelling's Trace, and Wilks's Lambda. If one of the tests should be regarded the highest, it would be Pillai's Trace, considered the most robust in significance for designs where the homogeneity of variance assumption is violated (Olson, 1976).

All three omnibus tests were significant at the $\alpha < .01$ level for the answer sheet format and the ability group variables; showing overall significant effects associated with both independent variables were present. However, significant interaction at the $\alpha \leq .05$ level between these two variables was not identified.

The next point to consider was the nature of significant differences particular to the mean vectors of each independent variable, as indicated by performance upon the individual subtests, and whether the levels within the independent variables showed significant departure from one another. All omnibus and follow-up significance test are shown in Tables 5a and 5b.

Univariate F-tests were initially used to identify which of the ten subtests showed significant differences in the mean vectors associated with the levels of each independent variable. For the recruit education level variable, each subtest except MC indicated significant differences among the three variable groups at the $\alpha < .01$ level. The WK and MK subtests both displayed very high F-scores, with CS, NO, AR, PC and GS showing lower, but still high F-scores, and AS and EI the lowest significant F-scores.

Table 5a: Omnibus Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Education Level Variables.

Omnibus Test	Recruit Education Level and Answer Sheet Format	Recruit Education Level	Answer Sheet Format
Pillai's Trace	.00487	.12566	.03027
F(Pillai's Trace)	1.51688	41.62076	19.37688
df(Pillai's Trace)	20,12416	20,12416	10,6207
Significance of F(Pillai's Trace)	.065	.000	.000
Hotelling's Trace	.00489	.14062	.03122
F(Hotelling's Trace)	1.51657	43.63549	19.37688
df(Hotelling's Trace)	20,12412	20,12412	10,6207
Significance of F(Hotelling's Trace)	.065	.000	.000
Wilks' Lambda	.99513	.87560	.96973
F(Wilks' Lambda)	1.51673	42.62751	19.37688
df(Wilks' Lambda)	20,12414	20,12414	10,6207
Significance of F(Wilks' Lambda)	.065	.000	.000

Stepdown F-tests (Roy and Bargman, 1958) were also used to examine for similar effects. All subtests except for EI reflected significant differences between the three variable levels at the $\alpha \leq .01$ level upon the overall multivariate effect being examined. The MK subtest

Table 5b: Follow-Up Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Education Level Variables.

Education Level

Subtest	Univariate F-Test	Stepdown F-Test
GS	50.09845**	50.09845**
AR	65.87332**	28.82839**
WK	102.38153**	39.01450**
PC	54.21206**	12.97421**
NO	73.26074**	40.43823**
CS	79.36613**	14.91862**
AS	12.19790**	66.73162**
MK	277.82388**	150.48399**
MC	1.16961	13.49627**
EI	11.21956**	.94798

Answer Sheet Format

Subtest	Univariate F-Test	Stepdown F-Test
GS	.02109	.02109
AR	1.47360	2.21552
WK	1.76887	2.33336
PC	.53666	.01185
NO	159.24446**	168.62511**
CS	18.40429**	11.84761**
AS	.65024	4.17498*
MK	1.20238	2.93410
MC	.36942	.71848
EI	.08233	.33615

* = Significant at $\alpha \leq .05$; ** = Significant at $\alpha \leq .01$.

demonstrated the highest F-score, with AS, GS, NO and WK all showing similar scores, and the CS, MC and PC subtests, the lowest significant F-scores. Significant departures among the education levels resulting from shared effects between intercorrelated subtests were also likely; the stepdown F-values compensated for shared effects and probably represented more appropriate levels of significant impact for each subtest than the univariate F-values.

To examine the sources of the significant differences between the education level variable's three groups, Scheffe's S test procedure was used to look at the variable's levels on each subtest. The significance level used for these procedures was $\alpha \leq .05$.

Significant differences between the non-high school graduate group and the recruits with some college education group were noted by the above contrasts on the GS, AR and WK subtests. All three groups showed significant departures from one another on the PC, NO, CS, AS and MK subtests. Yet, for the MC subtest, no significant differences were apparent among any of the three groups. Significant departures between the high school graduate group and the recruits with some college education group were noted for the EI subtest.

Likewise, significant differences for the two answer sheet format types associated with the ten subtests scores were investigated also. The univariate F-test procedures identified the NO and CS subtests as indicating significant differences, at the $\alpha < .01$ level, occurred between the circular and rectangular response formats. The NO subtest demonstrated a very large F-test value, with CS showing a much lower, but still highly significant one.

The stepdown F-tests also confirmed significant departure between the two formats for both subtests at the $\alpha < .01$ level and also for the AS subtest at the $\alpha \leq .05$ level. The NO F-value showed a slight increase (relatively speaking), and the CS F-value a slight decrease, from their respective univariate F-values. As these two subtests are highly intercorrelated, some of the significant effects demonstrated for NO and CS in the univariate F-tests are probably shared between the two subtests, which is accounted for in the stepwise F-tests.

Because order effects may influence the stepdown F-test procedure (Bray and Maxwell, 1982), an a priori ordering based on the operational test administration procedure was originally used for this procedure. Since significant differences associated with AS were unanticipated, based on the univariate F-test results, the subtest order for stepdown analysis purposes was adjusted to explore for possible significant differences associated with other subtests. NO and CS, because of significant differences between the answer sheet formats associated with their scores as demonstrated in the univariate F-tests, were respectively placed in the first two positions for the analysis, since they have already been confirmed

Table 5c: Alternate Stepdown Procedure #1, Subtest Order of NO, CS, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
NO	189.34123	.000
CS	21.77105	.000
GS	2.48880	.115
AR	6.35120	.012
WK	.68768	.407
PC	2.20549	.138
AS	3.25287	.071
MK	4.09735	.043
MC	2.03011	.154
EI	.00016	.990

Table 5d: Alternate Stepdown Procedure #2, Subtest Order of CS, NO, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
CS	16.99279	.000
NO	194.22345	.000
GS	2.48880	.115
AR	6.35120	.012
WK	.68768	.407
PC	2.20549	.138
AS	3.25287	.071
MK	4.09735	.043
MC	2.03011	.154
EI	.00016	.990

as contributing to significant differences between the education groups, then analyzing the remaining subtests according to operational administration procedure. By doing this, the other subtests were examined for any significant effects characteristic of them without being influenced by any significance associated with the NO and CS subtests. The adjustment in the subtest order for analysis purposes reconfirmed the significant differences associated with NO and CS and demonstrated AR and AS reflected significant differences in the two formats at the $\alpha \leq .05$ level. The results from these alternate stepdown F-test procedures are displayed in Tables 5c and 5d.

Univariate T-score contrasts were completed for each ASVAB subtest to examine the answer sheet format variable. Significant differences between the circular and rectangular response formats were confirmed for scores from the speeded subtests, NO and CS. Both T-scores were significant at the $\alpha < .01$ level; NO showed a high T-score of 13.633, and CS a lower one of 4.243. The nonsignificant T-scores ranged from 0.287 to 1.566. The T-scores are cited here to show the magnitude of the significance associated with NO and CS in comparison to the other subtests.

The overall hypothesis for the second MANOVA procedure addressed the possibility of significant departures among mean vectors representative of the answer sheet formats (rectangular and circular item responses) and the three recruit ability levels (low, medium and high), and the possibility of any significant interaction present between both independent variables. As with the first MANOVA procedure, if the omnibus tests revealed significant effects associated with the independent variables based upon the subtest scores, follow-up tests would be completed to examine which subtests were sources of the significant departures. These significance tests are presented in Tables 6a and 6b. The assumptions for this MANOVA procedure are similar to those of the first. Again, the homogeneity of variance assumption was not fulfilled for this procedure.

The three omnibus multivariate tests of significance used to consider the MANOVA hypothesis -- Pillai's Trace, Hotelling's Trace and Wilks's Lambda -- showed significant differences at the $\alpha < .01$ level among the two independent variables, as reflected in the subtest scores. Once again, no significant interaction was noted between the independent variables.

Univariate F-test procedures were used to identify which subtests were responsible for the significant differences associated with the mean vectors for the recruit ability group variable. All ten subtests demonstrated very high F-scores, significant at the $\alpha < .01$ level, which indicated each subtest showed significant differences among the mean vectors for the three groups.

Table 6a: Omnibus Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Ability Level Variables.

Omnibus Test	Recruit Ability Level and Answer Sheet Format	Recruit Ability Level	Answer Sheet Format
Pillai's Trace	.00194	.62835	.03605
F(Pillai's Trace)	.60438	285.80692	23.33155
df(Pillai's Trace)	20,12478	20,12478	10,6238
Significance of F(Pillai's Trace)	.913	.000	.000
Hotelling's Trace	.00194	1.62650	.03740
F(Hotelling's Trace)	.60443	507.22339	23.33155
df(Hotelling's Trace)	20,12474	20,12474	10,6238
Significance of F(Hotelling's Trace)	.913	.000	.000
Wilks' Lambda	.99806	.37823	.96395
F(Wilks' Lambda)	.60440	390.50215	23.33155
df(Wilks' Lambda)	20,12476	20,12476	10,6238
Significance of F(Wilks' Lambda)	.913	.000	.000

The stepdown F-test procedures also confirmed the above findings, demonstrating significant departures among the group mean vectors for all subtests at the $\alpha < .01$ level. Again, the F-test values were very large for each subtest in the stepdown procedure. Shared significant contributions to variable group differences seemed very possible here, as the stepdown F-values were much smaller for most subtests than the univariate F-values, which do not account for subtest intercorrelations, indicate.

Table 6b: Follow-Up Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Ability Level Variables

Ability Level

Subtest	Univariate F-Test	Stepdown F-Test
GS	2032.38231**	2032.38231**
AR	2173.31728**	926.64669**
WK	1467.17157**	135.18235**
PC	727.01747**	4.70480**
NO	212.33071**	44.17249**
CS	259.92671**	33.34904**
AS	938.62348**	146.38375**
MK	1612.04597**	217.82935**
MC	1537.07129**	54.51793**
EI	1466.90814**	37.67874**

Answer Sheet Format

Subtest	Univariate F-Test	Stepdown F-Test
GS	.08761	.08761
AR	.45816	.40145
WK	1.26308	1.07582
PC	.30381	.00014
NO	189.34123**	204.19126**
CS	16.99279**	17.39343**
AS	.00131	3.25287
MK	.36819	4.09735
MC	.10041	2.03011
EI	.17696	.00016

** = Significant at $\alpha \leq .01$.

Scheffe's S test procedures were also used to specify which group mean comparisons were the sources of the reported significant differences. For all ten subtests, the comparisons between the low and medium, low and high, and medium and high groups all displayed significant departures between their mean vectors.

Univariate F-test procedures indicated significant differences again between the two answer sheet formats at the $\alpha < .01$ level for the mean vectors associated with the NO and CS subtests. As in the first MANOVA, the speeded subtests recorded very high F-values of 189.34123 for NO and 16.99279 for CS, in comparison to the low values shown for the power subtests.

The stepdown F-test procedures also confirmed the significant departures ($\alpha < .01$) between the two format mean vectors associated with the NO and CS subtests and indicated F-values similar in size for these subtests as shown by the univariate F-tests. However, stepdown F-tests where NO and CS were alternately placed at the beginning of each subtest analysis sequence were completed to examine any other subtests for significant differences (unassociated with NO or CS) in the mean vectors for the answer sheet formats. When the two speeded subtests were alternately placed first in the sequence for the stepdown F-test analyses, both were significant at the $\alpha < .01$ level. The AR and MK subtests also indicated significant departures at the $\alpha \leq .05$ level. Shared variance among these two subtests is likely also, as their intercorrelation is a high one. These results are displayed in Tables 6c and 6d.

Table 6c: Alternate Stepdown Procedure #1, Subtest Order of NO, CS, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
NO	159.24446	.000
CS	16.23413	.000
GS	.40582	.524
AR	5.37522	.020
WK	1.56169	.211
PC	2.11243	.146
AS	4.17498	.041
MK	2.93410	.087
MC	.71848	.397
EI	.33615	.562

Table 6d: Alternate Stepdown Procedure #2, Subtest Order of CS, NO, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
CS	18.40429	.000
NO	157.00268	.000
GS	.40582	.524
AR	5.37522	.020
WK	1.56169	.211
PC	2.11243	.146
AS	4.17498	.041
MK	2.93410	.087
MC	.71848	.397
EI	.33615	.562

The third MANOVA procedure examined the hypothesis of no significant differences among the mean vectors for the answer sheet format variable (rectangular and circular item responses) and the six categories of recruit ethnicity (American Indian, Hispanic, Asian, Black, White and Other Ethnic) as reflected by the ASVAB subtest scores, and whether or not significant interaction emerged between the independent variables. If any of the three omnibus measures (Pillai's Trace, Hotelling's Trace or Wilks's Lambda) indicated significant differences were particular to the independent variables, post-hoc significance tests would be completed to examine which dependent variables were the source of the departures. The omnibus and follow-up significance tests are displayed in Tables 7a and 7b. For this MANOVA procedure, all assumptions were satisfied except for homogeneity of variance.

Pillai's Trace, Hotelling's Trace and Wilks's Lambda all demonstrated that significant differences were particular to the independent variables at the $\alpha < .01$ level, but no significant interaction was apparent between the independent variables.

Once again, univariate and stepdown F-test procedures were used to examine which of the subtests were sources of any significant differences between the mean vectors for the recruit ethnicity variable. Both procedures indicated each subtest was associated with significant differences ($\alpha \leq .01$) occurring among the mean vectors for some of the six ethnic groups. A wide range of F-values characteristic of all subtests is noted for both procedures. Significant differences indicative of unique effects, unassociated with shared variance between highly intercorrelated subtests which accounted for significant departures between mean vectors, were illustrated in the stepdown F-test results.

Table 7a: Omnibus Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Ethnic Group Variables.

Omnibus Test	Recruit Ethnic Group and Answer Sheet Format	Recruit Ethnic Group	Answer Sheet Format
Pillai's Trace	.00812	.31860	.00588
F(Pillai's Trace)	1.00398	42.01777	3.64686
df(Pillai's Trace)	50,30870	50,30870	10,6170
Significance of F(Pillai's Trace)	.466	.000	.000
Hotelling's Trace	.00814	.43418	.00591
F(Hotelling's Trace)	1.00417	53.56408	3.64686
df(Hotelling's Trace)	50,30842	50,30842	10,6170
Significance of F(Hotelling's Trace)	.465	.000	.000
Wilks' Lambda	.99190	.69077	.99412
F(Wilks' Lambda)	1.00408	47.55982	3.64686
df(Wilks' Lambda)	50,28142	50,28142	10,6170
Significance of F(Wilks' Lambda)	.465	.000	.000

Scheffe's S test procedure was completed to investigate where significant differences could be located among the six ethnic groups for each subtest.

Table 7b: Follow-Up Significance Tests for the 2 X 3 MANOVA Procedure for the Answer Sheet Format and Recruit Ethnic Group Variables.

Recruit Ethnic Group

Subtest	Univariate F-Test	Stepdown F-Test
GS	209.57499**	209.57499**
AR	120.03793**	24.30921**
WK	110.85498**	13.09776**
PC	48.67871**	2.99847**
NO	3.81835**	14.25501**
CS	7.20660**	6.04404**
AS	394.47879**	179.13676**
MK	20.40944**	12.51400**
MC	278.17645**	18.94868**
EI	207.49603**	4.53913**

Answer Sheet Format

Subtest	Univariate F-Test	Stepdown F-Test
GS	.72462	.72462
AR	2.77149	2.05858
WK	.10332	.32991
PC	3.60062	2.38320
NO	32.11202**	28.22817**
CS	6.75201**	1.31412
AS	.02389	.00383
MK	1.94323	.68384
MC	.02274	.67774
EI	.00085	.05298

** = Significant at $\alpha \leq .01$.

For the GS subtest, eight significant contrasts were noted. The White group was significantly different from the other five groups. The Black group also showed significant departures from the Hispanic, Asian, and American Indian groups.

Six significant contrasts were also identified for the AR subtest. The White group departed significantly from the American Indian, Hispanic, Black and Other Ethnic groups, while the Black group again showed significant differences from the Hispanic and Asian groups.

For the WK and PC subtests, the only significant contrasts identified were those for the White group, which was significantly different from the other ethnic groups for both subtests.

Among the speeded subtests (NO and CS), the only significant difference on NO was found between the Asian and Hispanic groups. No other significant contrasts were noted. For CS, the Asian group significantly diverged from the Hispanic and the Other Ethnic groups, while the Black and White groups also showed significant differences.

The AS subtest revealed significant departures between the White group and the Hispanic, Asian, Black and Other Ethnic groups. The Black group was significantly different from all other groups, and the Asian group likewise diverged significantly from the American Indian group.

Five significant contrasts were discovered for the MK subtest. The Asian group displayed significant departure from the Hispanic, Black and Other Ethnic groups. Also, the White group demonstrated significant differences in comparisons between the Hispanic and Black groups.

The MC subtest had eight significant contrasts associated with it. The Black group was significantly different from all other groups again, as the White group also departed significantly from the Hispanic, Asian and Other Ethnic groups.

Lastly, the EI subtest displayed eight significant contrasts among the six groups. The White group emerged as significantly different from all other groups once again. The Black group also indicated significant departures from the American Indian, Hispanic and Asian groups as well.

For the answer sheet format variable, the univariate F-tests identified both NO and CS as showing significant departures between the mean vectors for the two formats at $\alpha \leq .01$. Again, the NO subtest displayed a relatively high significant F-value and the CS subtest a smaller, but still suitably high significant F-score.

Surprisingly, the stepdown F-tests identified only NO as associated with significant differences between the two answer sheet formats' mean vectors. NO showed a consistently high significant F-value from the univariate F-test to the stepdown F-test procedures (32.11202 vs. 28.22817). Conversely, for the univariate F-test procedure, CS showed a relatively high significant F-value of 6.75201, then a much lower nonsignificant (at $\alpha \leq .05$) F-value of 1.31412 reported from the stepdown F-test procedure.

Significant differences in the groups' mean vectors attributed to CS based on the univariate F-test score were probably based on shared variance with the NO subtest, because they share a high intercorrelation as previously noted. To explore this finding, the subtest order in the stepdown F-tests was repeated as in the earlier MANOVA analyses; NO and CS each were respectively placed first and second in the subtest analysis sequence for the stepdown F-tests, with the other subtests maintaining their standard order. When NO was placed first in the procedure, it was the only subtest which demonstrated significant differences between the mean vectors for the answer sheet formats. However, when CS was placed first, it, along with NO, showed significant departures between the two answer sheet formats associated with it. Therefore, the significant effects associated with CS are probably associated with the intercorrelation it shares with NO versus any significant effects associated with it in its own right. These effects are reflected in Tables 7c and 7d.

Table 7c: Alternate Stepdown Procedure #1, Subtest Order of NO, CS, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
NO	32.11202	.000
CS	1.18474	.276
GS	.14561	.703

Table 7c: Alternate Stepdown Procedure #1, Subtest Order of NO, CS, GS, AR, WK, PC, AS, MK, MC, EI (Cont'd).

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
AR	.15523	.694
WK	.63247	.426
PC	.84607	.358
AS	.00383	.951
MK	.68384	.408
MC	.67774	.410
EI	.05298	.818

Table 7d: Alternate Stepdown Procedure #2, Subtest Order of CS, NO, GS, AR, WK, PC, AS, MK, MC, EI.

Subtest	Stepdown F-Score	Significance of Stepdown F-Score
CS	6.75201	.009
NO	26.51783	.000
GS	.14561	.703
AR	.15523	.694
WK	.63247	.426
PC	.84607	.358
AS	.00383	.951
MK	.68384	.408
MC	.67774	.410
EI	.05298	.818

DISCUSSION AND CONCLUSIONS

The first MANOVA procedure examined for significant departures among the levels of the answer sheet format (circular response vs. rectangular response) and the recruit education level (non-high school graduates, high school graduates, and high school graduates with some college education) variables as reflected by the ASVAB subtest scores, and any significant interaction which resulted from the combination of the above independent variables. Multivariate omnibus measures of significance (Pillai's Trace, Hotelling's Trace, and Wilks's Lambda) confirmed overall significant differences characteristic of one or both of the independent variables; however, no significant interaction was displayed by the independent variable at the $\alpha \leq .05$ level.

Univariate and stepdown F-tests were used to specifically determine which ASVAB subtests were associated with significant differences between the mean vectors for the groups of each independent variable. The Scheffe's S test procedure was used to examine for the significant differences among the groups of the independent variables.

In examining the univariate F-tests for the recruit education level variable, all subtests except MC indicate individual significant contributions to the multivariate effect. The stepdown F-tests showed all subtests except for EI demonstrated significant contributions to the multivariate composite. Scheffe's S test demonstrated significant differences between two of the three groups on GS, AR, WK (all for non-high school graduates vs. high school graduates with some college education) and EI (the non-high school graduates vs. high school graduates); significant departures among all three groups on the PC, NO, CS, AS and MK; and no significant departures on MC.

In conclusion, the above results indicated that all subtests except for MC or EI were sources of significant differences among the three recruit education categories. Based on both sets of F-scores and results of the contrasts, WK, NO, and MK appeared to have the most significant impacts upon the three groups. The remaining subtests, excluding MC and EI, were moderately associated with the above significant differences. MC and EI, because of their lower F-values (in comparison with the above subtests) probably provide the least noteworthy contributions to the differences among the three levels of the recruit education variable.

The above results indicated varying levels of education are viable sources of significant effects as reported upon the ASVAB subtests. These results only make sense, as it is more likely that persons graduating from high school and, in some cases, continuing their education after high school would perform better overall on an aptitude test which measures mathematics, verbal or technical abilities than a person who dropped out of high school or one who is knowledgeable or experienced in one specific subject area, but has neglected learning in other areas.

For the answer sheet format variable, the univariate and stepdown F-tests showed NO and CS were responsible for significant differences among the two response formats. Additionally, the initial stepdown F-test procedure identified AS as being associated with significant departures between the two formats to a lesser extent. When the subtest sequence was changed, alternately placing one of the speeded subtests first, then using the operational administration order for GS through EI for the stepdown F-test procedure, significance associated with AR was also identified.

Univariate contrasts were also used to examine for significant differences between the two answer sheet formats associated with the subtests. As with the two F-test procedures, the NO and CS subtests showed highly significant departures noted for the two format types on each subtest. The AS and AR subtests showed no significant departures between the answer sheet formats.

Reviewing the above data, significant differences among the answer sheet formats -- circular vs. rectangular response -- were reflected by both NO and CS. The other subtests, aside from AS and AR on a negligible basis, were not associated with any significant differences between the mean vectors representative of the two answer sheet formats. These results correspond highly with results cited for previous studies of answer sheet format effects, even when all ten subtests were considered through MANOVA procedures.

The second MANOVA procedure examined the independent variables of answer sheet format (circular response vs. rectangular response) and the recruit's ability level, as defined by their prior enlistment ASVAB test scores, for any significant differences among mean vectors characteristic to each, along with any instances of significant interaction between them. The multivariate omnibus significance measures (Pillai's Trace, Hotelling's Trace, and Wilks's Lambda) confirmed significant differences among the different levels of both variables, but displayed no significant interaction between them.

Again, univariate F-tests, stepdown F-tests, and Scheffe's S test procedures were examined to locate where significant departures occurred relative to each independent variable.

In examining the univariate F-tests for the recruit ability level variable, significant departures among the three levels were noted for all ten subtests, according to the F-values, which were very high. The stepdown F-tests likewise confirmed significant differences associated with each subtest. Once again, shared significant effects between intercorrelated subtests not accounted for in the univariate F-test values were better reflected in the stepwise F-test values, as the more important contributors such as the WK or AS subtests were shown. Lastly, the Scheffe's S test procedures displayed significant departures among all three ability levels for all ten subtests.

For the recruit ability level, all subtests, to some extent, significantly contributed to differences among the recruit ability groups. The GS, AR, WK, AS and MK subtests accounted for the most unique significant differences associated with the three levels. Noting this, it appears that verbal, mathematics and/or science attributes are major sources of departure among ability levels for this sample. These differences might be representative of recruits who differed in the quality of their respective educations, the amount of learning about these subjects acquired out of the classroom, the preparation each undertook to complete the ASVAB battery, or even their physical or mental condition on the day of the test. In other words, certain immeasurable variables could have influenced the recruit's ASVAB performance. However, it appears that differences in the recruit's verbal, math and science (to a lesser extent) abilities for this sample are reflected by the ASVAB subtest scores.

The answer sheet format variable for this MANOVA procedure reflected much of the same information as for the first MANOVA. Both speeded subtests showed significant differences between answer sheet formats confirmed by the univariate and stepdown F-test values. To examine for any other significant departures between the two formats apart from unique effects associated with NO and CS, the subtest order used to examine for significant effects associated with AR and AS in the first MANOVA procedure was repeated. Significant departures associated with AR and MK (at $\alpha \leq .05$) were noted for this analysis. It appeared that the significant unique impacts noted for AR and MK noted in the above stepdown F-test procedure might have been missed in the univariate F-tests, due to shared variance resulting from the high correlation between them unaccounted for in the univariate F-values.

For this second analysis, significant differences between the answer sheet formats were mostly associated with NO and CS. It was only after the significant effects of NO and CS were accounted for that significant impacts associated with the AR and MK subtests appeared. The significance associated with these subtests is probably not the primary source of differences between the circular and rectangular formats. The results for NO and CS in this analysis were, once again, consistent with prior studies confirming differences in speeded subtest scores arising from differences in answer sheet formats.

The last MANOVA procedure examined for significant differences between the mean vectors associated with the answer sheet format (circular response vs. rectangular response) and the recruit's ethnic background (American Indian, Hispanic, Asian, Black, White and Other Ethnic) independent variables, as reflected by the subtest scores, and any significant interaction apparent between them. Pillai's Trace, Hotelling's Trace, and Wilks's Lambda were used to investigate for overall significant differences peculiar to one or both of the independent variables; significant departures between mean vectors were confirmed among the groups for the independent variables, yet no significant interaction among them was noted.

Univariate F-tests, stepdown F-tests, and Scheffe's S test procedures were utilized to identify which group comparisons reflected significant differences and which subtests were closely tied to the aforementioned significant differences.

First of all, significant departures between the six ethnic groups as demonstrated in the subtest scores were examined. Univariate F-tests indicated all subtests contributed to the significant departures among levels of the ethnic group variable. The GS, AR, WK, AS, MC and EI subtests appeared to account for the most significant effects observed for the six ethnic groups. Effects associated with the speeded subtests showed the lowest significant F-scores. The stepwise F-tests also showed each subtest contributed significant unique effects to departures between the ethnic groups. The stepwise F-values were much smaller in most instances than the univariate F-values for the subtests; they accounted for intercorrelations among the subtests and examined each subtest's unique contribution to significance and excluded shared variance unaccounted for in the univariate F-values.The subtests which showed the most unique contributions to significant effects between the six groups were the GS, AR, WK, NO, AS and MC subtests. NO showed a higher significant F-score from the stepwise procedure than from the univariate procedure, indicating more unique effects were associated with it than revealed by the univariate F-test.

For the recruit ethnic group independent variable, the Scheffe's S test procedure identified numerous significant differences among the six groups. The White group was significantly different than all other groups on four subtests -- GS, WK, PC and EI; the Black group showed significant departures with all other groups on AS and MC. Significant differences between the American Indian and, respectively, the Asian group on AS; the Black group on GS and EI; and the White Group on AR were also noted. The Hispanic group showed significant differences with the Asian group on NO, CS and MK; the Black group on GS, AR and EI and the White group on AR, AS, MK and MC, aside from the earlier differences cited; and the Other Ethnic group on the CS subtest. Besides the earlier contrasts noted, the Asian group showed significant differences from the Black group on the GS, AR and MK subtests; the White group on the AS and MC subtests, and the Other Ethnic group on MK. The Black group was also significantly different from the White group on the AR, CS and MK subtests, with the White group showing significant differences with the Other Ethnic group on the AS and MC subtests.

In conclusion for the recruit ethnicity variable, it appeared all subtests, except NO and maybe CS, are sources of significant departures between the six groups of the variable. Consistency between the F-test results and the Scheffe's S test procedures indicated that the GS, AR, WK, AS, MC and EI subtests all were major sources of significant differences among the groups. The PC and MK subtests also showed smaller amounts of significant effects among the ethnic groups, and the speeded subtests, NO and CS, accounted for the least amounts of significant effects associated with each ethnic group.

The differences on subtest performance might have resulted from a number of sources particular to the recruit, such as the type or quality of education, the type of cultural background he or she is from, or the type of opportunities for learning through experience outside of the classroom available to he or she. Likewise, language barriers relative to the recruit's ethnicity might also have influenced the scores related to this sample, as well as environmental or socioeconomic differences in the regions they were reared in.

The answer sheet format variable was examined using univariate F-test and stepwise F-test procedures. The univariate F-tests showed both NO and CS contributed to significant departures between the answer sheet format mean vectors. However, the stepdown F-tests only demonstrated NO as a source of significant differences between the mean vectors. NO consistently showed a high F-value for both F-test procedures, yet CS demonstrated a moderately high F-value for the univariate procedure and a much lower one associated with the stepwise procedure. This indicated shared significant contribution to differences in the mean vectors, because of the high intercorrelation between NO and CS, was associated with CS in the univariate F-score. Later stepwise analysis, repeating the alternate subtest order for examination as in the first two MANOVA analysis, showed that when the significant effects responsible for notable differences in the two formats displayed by NO were accounted for, CS showed little unique contribution to the significant differences between the circular and rectangular answer sheet formats. These results once again concurred with the previous MANOVA procedures completed for this study, and the prior studies cited in the literature review.

REFERENCES

Bloxom, B., McCully, R., Branch, R., Waters, B.K., Barnes, J., & Gribben, M. (1991). Operational Calibration of New Optical-Mark-Reader Answer Sheets for the Armed Services Vocational Aptitude Battery. Unpublished manuscript, Defense Manpower Data Center, Monterey, CA.

Bray, J.H. & Maxwell, S.E. (1982). Analyzing and interpreting significant MANOVAs. Review of Educational Research, 52, 3, 340-367.

Earles, J.A., Guiliano, T., Ree, M.J. & Valentine, L.D., Jr. (1983). The 1980 youth population: An investigation of speeded subtests. Unpublished manuscript, Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks Air Force Base, TX.

Haase, R.F. & Ellis, M.V. (1987). Multivariate analysis of variance. Journal of Counseling Psychology, 34, 4, 404-413.

Hays, W.L. (1981). Statistics (3rd Ed.). New York, NY: CBS College Publishing.

McWilliams, H.A. (1980). Profile of American Youth: Field Report. Chicago, IL: National Opinion Research Center.

Olson, C.L. (1976). On choosing a test statistic in multivariate analysis of variance. Psychological Bulletin, 83, 579-586.

Ree, M.J. & Wegner, T.G. (1990). Correcting differences in answer sheets for the 1980 Armed Services Vocational Aptitude Battery reference population. Military Psychology, 2, 3, 157-169.

Roy, J. (1958). Step-down procedure in multivariate analysis. Annals of Mathematical Statistics, 29, 1177-1187.

Sims, W.H. & Maier, M.H. (1983). The appropriateness for military application of the ASVAB subtests and score scale in the new 1980 reference population (Memorandum 83-3102). Alexandria, VA: Center for Naval Analyses.

Valentine, L.D., Jr. & Cowan, D.K. (1974). Comparability study of Armed Services Vocational Aptitude Battery scores from answer sheet and answer card administration (AFHRL-TR-73-55, AD-775 048). Air Force Human Resources Laboratory, Personnel Research Division, Brooks Air Force Base, TX.

APPENDIX A: MULTIPLE REGRESSION PROCEDURES

Multiple regression procedures were also used with this study to further examine the relationships among the independent variables of recruit ability level, education level and ethnic background, and answer sheet type (circular vs. rectangular item response) with the dependent variables of the ASVAB subtests.

Each subtest was examined on a univariate level to determine which independent variables were most associated with significant differences displayed by the subtest scores. Both forced entry and stepwise regression procedures were used to consider the above inquiry, with each procedure yielding very similar results.

For GS, 40.1 percent of all available variance was accounted for. The recruit ability and recruit ethnicity variables both assumed significant portions of the available variance; whereas the answer sheet format and recruit education variable did not, displaying nonsignificant F-values representative of their beta weights.

40.8 percent of the available variance for AR was accounted for, primarily by the recruit education, recruit ability and recruit ethnicity variables. These variables were sources of significant effects associated with the subtest. The answer sheet format variable indicated no significant impact upon AR, as a very small, nonsignificant F-value was associated with its beta weight.

WK had 33.2 percent of its available variance accounted for. Once again, the recruit education, recruit ability and recruit ethnicity variables assumed significant portions of available variance; the answer sheet format variable did not, again demonstrating a nonsignificant F-value associated with its beta weight.

PC demonstrated 19.4 percent of its total variance was explained. The three recruit characteristic variables all showed significant impacts upon this subtest's scores. The answer sheet format variable once again was the odd one out, displaying a minuscule F-value which indicated no significant effect upon PC.

For NO, 10.3 percent of total variance was explained. All independent variables displayed significant effects upon this subtest, as indicated by significant F-values associated with each of their beta weights. Among the variables, the recruit ability group variable accounted for the highest portion of total variance, with the answer sheet variable representing the next largest amount.

CS showed about 9.0 percent of available variance accounted for. Again, all four independent variables assumed significant portions of available variance; though for this subtest, answer sheet format was responsible for the second smallest significant amount of available variance.

27.4 percent of total variance was reported for AS. Again, the three recruit characteristic variables (recruit education, ability and ethnicity) represented significant portions of available variance; the answer sheet format variable did not, revealing a very small, nonsignificant F-value associated with its beta weight.

For MK, 37.0 percent of overall variance was explained. As with AR, WK, PC and AS, the recruit education, recruit ability and recruit ethnicity variables reflected significant impacts upon this subtest's scores. The answer sheet format variable did not exhibit any significant effect upon MK scores.

On MC, 34.5 percent of total variance was represented by the three recruit characteristic variables, which all assumed significant portions of available variance. The answer sheet format variable, displaying a small, nonsignificant F-value associated with its beta weight, again had no viable effect demonstrated by MC subtest scores.

33.0 percent of available variance was reported for EI. The troika of recruit characteristic variables were identified as providing a significant impact upon this subtest. The answer sheet format variable demonstrated a minuscule, non-significant F-value associated with its beta weight.

For each subtest, less than half of available variance is accounted for by these independent variables. The range consisted of around 9 percent accounted for on CS to 40.8 percent accounted for on AR. Generally, about a third of available variance was accounted for in the academic and technical subtests, except for AS (27.4 percent) and PC (19.4 percent). NO (10.3 percent) and CS (9.0 percent), the speeded subtests, accounted for the lowest proportions of available variance.

The recruit ability group variable accounted for the largest significant amounts of available variance among all subtests. The recruit ethnicity variable assumed the next largest proportions of variance for GS, AR, WK, AS, MC and EI; the applicant education level variable did the same for PC, CS and MK; and the answer sheet format variable for NO. The amounts of available variance accounted by the recruit ability variable were substantially higher than those variables assuming the next larger proportions of variance for each subtest. For the power subtests, answer sheet format variable was the only variable which did not significantly account for any available variance; answer sheet format variable and recruit education variable did not for GS; and all four variables accounted for significant amounts of available variance on NO and CS.

It does not appear that shared variance occurs among the independent variables which individually account for portions of available variance. The recruit ability variable shows very small positive correlations with the answer sheet format, recruit education and recruit ethnicity variables respectively. The correlations between the answer sheet format variable and the three recruit characteristic variables do not confirm any viable relationships among them respectively. The recruit education variable, as previously mentioned, shows a very small positive intercorrelation with the recruit ability variable and no evident relationships with the answer sheet format or recruit ethnicity variables. Based on this data, it appears that variance associated with each of the independent variables is probably an accurate reflection of each variable's effect upon the subtest scores.

Conceptually speaking, one examining this data might conclude that recruit ability, reflected by the mean of the subtest Z-score totals derived from each recruit's enlistment subtest scores, demonstrates the highest degree of influence upon each of the subtest scores. Therefore, one would surmise that subtest score differences should be accurately reflected by differences in ability levels; this inference was confirmed earlier by the MANOVA procedures.

Similarly, recruit ethnicity also was associated with significant effects demonstrated upon GS, AR, WK, AS, MC and EI; the variable indicated apparent significant effects upon these subtest scores among the six ethnic groups. These results among ethnic groups could be due to cultural differences in educational methods, varying quality of regional educational opportunities, differences in perspective regarding education based upon cultural background and language-based or environmentally-based differences.

The education level variable demonstrated significant effects associated with the PC, CS and MK subtests. One might infer from this result that differences among the recruits' respective education levels would be significantly reflected by scores from the three subtests. This inference was also confirmed by the MANOVA procedures examined previously.

Lastly, the answer sheet format variable was only associated with significant differences reflected in the NO and CS subtest scores; this would signify that recruit differences in these subtest scores might be directly attributable to the circular vs. rectangular item response answer sheet formats. Though, the amount of variance accounted for among the speeded subtests is small, a link between the answer sheet format variable and performance upon the NO and CS subtests may be inferred.

The remaining variance to be accounted for might reflect significant differences between variables such as recruit service affiliation, gender or test number (i.e., second or third test, with a possible retest effect occurring) which were not addressed in this study.

APPROVAL SHEET

The thesis submitted by Mark Johnson has been read and approved by the following committee:

Dr. Jack Kavanagh, Director, Professor, Education, Loyola

Dr. Steven Miller, Professor, Education, Loyola

The final copies have been examined by the director of this thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

July 7, 1992
Date

Jack A. Kavanagh
Signature