FEBRUARY 1993

INCREMENTAL VALIDITY OF EXPERIMENTAL COMPUTERIZED TESTS FOR PREDICTING TRAINING CRITERIA IN MILITARY TECHNICAL SCHOOLS

Norm M. Abrahams John J. Pass
Jeffrey W. Kusulas
Darlene R. Cole
William F. Kieckhaefer

RGI Consultants
Lloyd G. Humphreys
Edward F. Alf

RGI, Incorporated
Personnel Systems Division
591 Camino de la Reina, Suite 917
San Diego, CA 92108

Prepared for
Navy Personnel Research and Development Center
53335 Ryne Road
San Diego, CA 92152-7250

Contract # N66001-90-D-9502
Delivery Order 7J13
Deliverable CDRL A007
Final Report

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
February 1993

3. REPORT TYPE AND DATE COVERED
FINAL

4. TITLE AND SUBTITLE
Incremental Validity of Experimental Computerized Tests for Predicting Training Criteria in Military Technical Schools

5. FUNDING NUMBERS
Program Element 0602233N,
Work Unit RM33M20.04
Contract Number N66001-90-D-9502

6. AUTHOR(S)
Norm M. Abrahams, John J. Pass, Jeffrey W. Kusulas, Darlene R. Cole, William F. Kieckhaefer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
RGI, Incorporated --- Personnel Systems
3111 Camino Del Rio North ---Suite 802
San Diego, CA 92108

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Navy Personnel Research and Development Center
53335 Ryne Road
San Diego, CA 92152-7250

10,. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Functional Area: Personnel Systems
Product Line: Computerized testing
Effort: Enhanced Computer Administered Testing (89-036)

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

12B. DISTRIBUTION CODE
A

13. ABSTRACT *(Maximum 200 words)*

This report presents analyses of two previously collected data sets to determine the increase in validity from adding new computerized aptitude tests to the Armed Services Vocational Aptitude Battery (ASVAB). The first data set contained test measurements on technical training students from only the Navy, and used an experimental battery called the Computerized Test Battery (CTB) along with the ASVAB. The second data set consisted of tests of students in the Army, Navy, and Air Force using the Enhanced Computer Administered Tests (ECAT) and the ASVAB. The study investigated the incremental validity of individual experimental tests, composites of experimental tests, and estimates of *g* which were based on the experimental tests.

Conclusions: (1) The Assembling Objects and Two-Hand Tracking tests can outperform some ASVAB tests in predicting school criteria as measured primarily by Final School Grade (FSG). (2) While the ECAT and CTB tests provide a relatively small amount of incremental validity over the ASVAB when averaged across schools, they provide substantial validity gains for some of the schools. (3) The ECAT Spatial dimension is the most promising dimension for increasing FSG validity over the ASVAB.

13. SUBJECT TERMS
Computerized Testing, ASVAB, ECAT, Aptitude Test, Test Validity, Incremental Validity

15. NUMBER OF PAGES
96

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT
UNCLASSIFIED

20. LIMITATION OF ABSTRACT
UNLIMITED

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

ACKNOWLEDGEMENTS

We wish to acknowledge Ms. Kathleen Moreno for serving as the Contracting Officer's Technical Representative for this project. We also want to recognize Dr. David Alderton, Mr. Jerry Larson, and Mr. John Wolfe for their development of the technical requirements contained in the contract statement of work.

CONTENTS

LIST OF TABLES

INTRODUCTION

The Department of Defense and the military services have an on-going objective to improve their personnel selection and classification systems. The Armed Services Vocational Aptitude Battery (ASVAB) is the current battery of aptitude tests used to select and classify enlisted personnel for military service. The military services use a composite of ASVAB tests to measure general cognitive ability. This composite, called the Armed Forces Qualification Test (AFQT), is used to determine eligibility for the military. Other ASVAB composites are used to classify AFQT-qualified individuals into specific occupational training schools. Research has demonstrated the substantial predictive power of the ASVAB tests (see Segall, Kieckhaefer, and Day, 1985, for a review of ASVAB validity studies).

The development of computer technology and advances in cognitive psychology have created opportunities for the development of new computerized ability tests. Such tests might be able to measure more effectively some of the same abilities measured by the ASVAB or measure additional abilities not measured by the ASVAB. Over the last several years, the military personnel labs have developed a number of new computerized ability tests. These tests measure a wide variety of abilities, including spatial, perceptual speed, and psychomotor. This study seeks to determine whether the new experimental tests can increase the predictive validity beyond that obtained by the ASVAB.

APPROACH

There are two ways by which validity increments to the ASVAB could occur. First, the new tests, by themselves or in concert with the ASVAB tests, could produce an estimate of general intelligence (g) that increases predictive validity beyond that currently available in the ASVAB. Second, in addition to enhancing g, the new tests could measure specific criterion-related abilities not included in the ASVAB.

Several studies have already examined some of the new tests developed by the services as well as some of the data analyzed in this study (Alderton and Larson, 1992; Larson and Alderton, 1992; Peterson, Oppler, and Rosse, 1992; Wolfe, Alderton, and Larson, 1992; Carey, 1992; Busciglio, 1990).

In this study, we investigated the incremental validity of individual experimental tests, composites of experimental tests, and estimates of g which were based on the experimental tests. For each level of analysis, we evaluated the incremental validity of the experimental predictors relative to some combination of ASVAB tests. In general, we based incremental validity analysis on multiple regression procedures. We obtained multiple correlations by first using an accretion procedure (to determine which new test or composites add any validity to the ASVAB) and then using a deletion procedure (to determine which new tests or composites are unique in regard to other tests or composites). Finally, we compared our findings to the relevant findings of some of the previously mentioned studies.

METHOD

Data Sets

We analyzed two independently-collected data sets provided by the Navy Personnel Research and Development Center (NPRDC). These data sets, referred to as the Navy Only data set and the Joint-Service data set, consist of ASVAB and experimental predictor and training school criteria scores for enlisted military personnel in a number of different training schools. Regarding criteria, Wolfe et al., (1992) and Busciglio (1990) have shown some limited evidence that certain specific abilities (e.g., psychomotor and spatial) can obtain higher incremental validities for performance as opposed to more knowledge-oriented criteria. Because of this evidence and findings reported in Wigdor and Green (1991) on the moderate relationship between knowledge and hands-on criteria, we conducted separate analyses on Final School Grade (FSG) and performance criteria when both measures were available. In many schools, FSG contains both knowledge and performance components which are uniquely weighted for each school.

Navy Only Data Set

Table 1 displays the Navy training schools represented in this data set. Table 2 lists the experimental predictors included in the Computerized Test Battery (CTB) and the ASVAB tests that were administered to Navy subjects. Wolfe et al., (1992) describe the test administration procedures, the predictors listed in Table 2, and the specific measures or test scores and criteria included in this data set. The CTB tests were administered to recruits along with the ASVAB prior to entry into a training school. The Space Perception (SP) and Figural Reasoning (FR) tests, as indicated by "P&P" in the Table, are computer-administered versions of paper-and-pencil tests.

The ASVAB scores obtained from this experimental post-enlistment administration were used only for data quality control purposes (e.g., to ensure consistency with results in the Wolfe et al. study, 1992). We used pre-enlistment ASVAB scores on the tests shown in Table 2 for all other analyses for two reasons. Most importantly, pre-enlistment ASVAB scores are the only ASVAB scores available for all subjects in the Joint Service data set (described below). Use of pre-enlistment ASVAB scores for both data sets permits more direct comparisons. Secondly, these scores reflect the motivational set present at the time of application, thus providing one necessary condition for generalization of results to future applicant populations. The ASVAB scores were operational standard scores for both data sets.

We show descriptive statistics for all predictors in Table 3 and all criteria in Table 4. The sample sizes displayed in Table 3 indicate the total number of individuals who had a score on each of the tests in the data set. As shown, not all individuals have complete sets of test scores. The sample sizes in Table 4 are the number of individuals remaining in the data set after a number of screening procedures were applied (see Appendix A for a description of these screening procedures).

Joint Service Data Set

Subjects from the Air Force, Army, and Navy are included in this data set. Table 5 displays the service training schools in this data set and identifies six Navy schools also represented in the Navy Only data set.

Table 6 presents the Enhanced Computer Administered Tests (ECAT) used in the Joint Service analyses. The ECAT tests and the specific test measures (scores) footnoted in the table are described in Alderton and Larson (1992). Again, the "P&P" in the table indicates that the test is a computerized version of a paper-and-pencil test. This table also shows that the first four predictors in the ECAT are included in the CTB. As previously stated, we used pre-enlistment ASVAB scores for analysis of the data. Table 7 shows descriptive statistics for the predictors.

Table 8 presents the criterion measures we selected and analyzed for the school samples included in this data set. These measures are a subset of the total set of criteria recommended by Kieckhaefer et al. (1992). In order to focus our analysis and to ensure results were reasonably interpretable, we limited our criteria to FSG and the single most comprehensive and reliable performance measure for those schools with performance criteria. We selected both an FSG and performance criterion for eight schools (i.e., AC, AE, AMS, AO, OS, 272(2), 732, 13F). Separate analyses were conducted with this subset of schools to determine the impact of criterion type on validity.

Analyses

While most analyses were conducted on both data sets, we performed several analyses only on the Joint Service data set. The Statistical Package for the Social Sciences was used for performing most of the regression analyses and the factor analyses. We used the MVCOR program (Sympson and Candell, 1989), as modified by McBride in 1989, to correct for multivariate restriction of range in the predictors. We developed software to perform other required analyses.

Data Verification and Data Screening

As one of the first steps, we obtained descriptive statistics and conducted a number of analyses on the data sets described above to evaluate the accuracy and completeness of the data. In addition, we conducted preliminary analyses to ensure that statistical results were consistent with previous findings (Wolfe et al., 1992). The data screening procedures that we used on both data sets are described in Appendix A.

Hypothesis Testing and Type I Error

In order to control Type I errors caused by multiple comparisons on the same data, we used the hierarchical approach employed by Wolfe et al. (1992) and based on procedures described in Cohen and Cohen (1983). For each school criterion, we calculated a multiple correlation for all ASVAB tests (R_{ASVAB}) and a multiple correlation for all ASVAB tests plus the experimental tests ($R_{ASVAB+CTB}$ or $R_{ASVAB+ECAT}$). The multiple Rs were computed on FSG criteria, with the exception of 11H and 19K for which performance criteria were used. The incremental validities (differences between multiple correlations) were determined along with their associated probabilities for each criterion using the F distribution as described in Cohen and Cohen (1983). We combined these individual probabilities and tested for overall significance using Fisher's chi-square test (Fisher, 1932). This procedure permits accumulating across separate samples, and determines whether the entire experimental battery significantly increments ASVAB validity across all schools.

If the overall hypothesis of no increment is rejected, the probability of a significant incremental validity for each new predictor and each new content dimension is calculated using the procedures mentioned above. If the null hypothesis is again rejected, significance tests may be applied to more specific comparisons.

Validity of Individual Tests

For each school criterion, we entered all ASVAB tests as a set in the first step of a multiple regression analysis. Next, we entered the full set of experimental predictors and evaluated the increment in validity.

Then, we added each experimental predictor alone (accretion procedure) to the set of ASVAB tests and the incremental validity was determined. Using regression deletion procedures, each new predictor was deleted from the full model of all predictors and the impact on validity was evaluated.

These analyses create the opportunity for the ASVAB tests to capitalize on chance and yield a composite that may remove more criterion variance than is truly predictable by ASVAB and reduces the opportunity for any experimental test or test composite to predict residual criterion variance. Under such circumstances, these analyses may be considered to be conservative since it is less likely for an experimental test to significantly increase validity relative to the ASVAB.

Validity of Content Dimensions

By combining various ASVAB and experimental test scores, we created two sets of content dimension scores for each subject, one for ASVAB and one for the experimental battery. The specific tests and weights associated with each dimension were provided by NPRDC and are referred to in the Results section.

For each school, all ASVAB content dimensions were entered as the first set of variables in multiple regression analysis. At stage two, all experimental battery based content dimensions (e.g. Spatial, Psychomotor, etc.) were entered and the incremental validity was evaluated. Accretion and deletion procedures as described above with the individual tests were used to determine which dimensions provide incremental validity and which provide unique incremental validity.

Operational Composite Analysis

For each school, we entered its operational selection composite (as provided by NPRDC) as a single variable into the first step of a regression analysis. The composites were based on

ASVAB standard scores. We did not convert any standard score composites to other scales such as percentiles. In the second stage of the regression analysis, we entered all the content dimensions derived from the experimental tests and assessed the resulting incremental validity. With the ASVAB operational composite as the first variable in multiple regression, accretion and deletion procedures were then applied to evaluate the incremental validity of each experimental content dimension.

Optimal School Composites

We developed an optimal predictor composite for each school criterion, allowing ASVAB and the experimental tests to enter freely into the equation. We calculated a full model (i.e., all ASVAB and all ECAT tests) regression equation for each school.

Derivation and Analysis of Methods to Estimate g

As indicated in the Approach, one way to increase validity is to develop an estimate of g which is more comprehensive and has more predictive power than ASVAB alone. We derived estimates by several different means.

One method used the first principal component of several sets of predictors. Concerning the choice of factor analysis methods, research (Ree and Earles, 1991) has shown almost no difference in the gs produced by different factor analysis methods applied to ASVAB test scores. For example, Ree and Earles (1991) report correlations of .996 and .994 between g scores from a hierarchical solution with g scores from the first principal component and the first principal factor, respectively. We factor-analyzed all ASVAB tests, all ASVAB tests plus the CTB predictors, and all ASVAB tests plus the ECAT predictors. Prior to component analysis, the intercorrelation matrices were corrected for range restriction using Lawley's (1943) multivariate correction procedure and unrestricted ASVAB score statistics for the 1991 joint service population of applicants. Test scores (including ASVAB scores) were standardized to z scores based on the 1991 DoD population for both data sets. We weighted these standardized test scores by the factor score coefficients derived for the respective first principal components.

In addition to using the component analytic method, we also computed composites called "psychological" g for the ASVAB plus CTB predictors and for the ASVAB plus ECAT predictors. To obtain these composites, we weighted standardized test scores by their unrestricted correlation with the Figural Reasoning test as specified by NPRDC. For both composites, NPRDC specified the weight for the Figural Reasoning test to be .75. Since the Figural Reasoning test is assumed to measure a basic psychological process, NPRDC refers to these composites as "psychological" g. The specific tests and weights associated with each equation are presented in the Results section.

AFQT scores, computed for both data sets using the current formula (2VE+1MK+1AR), and the g scores derived by the above methods were intercorrelated and each set of scores was correlated with the FSG (or other criteria if FSG was not available) for the schools in the respective data sets. The AFQT was included in this comparison since we perceive it to be the operational estimate of g used as a selection standard for entry into the military. Percentile scores for the AFQT were not used in these comparisons as we did not have percentile distributions for the other g score distributions.

Correcting and Averaging Validities

After completing all the incremental validity analyses and associated significance tests, we applied various corrections to the data. These consisted of Lawley's (1943) multivariate correction for range restriction in the predictors (based on the 1991 DoD population of applicants and used for both data sets), correction for shrinkage (Wherry procedure), and correction for range restriction and unreliability of the criteria (Gulliksen, 1987). The resulting fully-corrected coefficients were averaged across schools after weighting each coefficient by its degrees of freedom. For the Joint Service data set, results are tabulated separately for performance and FSG criteria for certain analyses.

The average validities are used to calculate the incremental validity for a variety of comparisons between the ASVAB and the experimental predictors. Incremental validity is thus assessed in terms of correlation points. We also calculated the percent gain in correlation points. However, care must be taken in regard to interpreting percent gain as it is dependent on the magnitude of the multiple correlation over which incremental validity is being evaluated (e.g., a .1 R gain over an R of .8 is 12.5%, but over an R of .2, it is 50%).

Final Battery

Failure to reduce the predictor space to an efficient and valid battery would be wasteful in terms of testing time and resources. From the total ASVAB and ECAT batteries, we identified the best possible "final" battery of approximately 10 tests in terms of predictive validity.

To identify the best subset of approximately 10 operational and experimental tests three different cumulative multiple regression analyses were performed. The first analysis was based on correlations corrected only for multivariate restriction in range. The second analysis was identical to the first, except the multiple correlations at each stage were shrunken using the Wherry correction. The third analysis was identical to the second, except that the Wherry-corrected multiple correlations were corrected for criterion unreliability.

The first step of each analysis selected the single test that had the highest weighted average validity across schools. In the next step, multiple correlations were computed for all possible combinations of the first test with each of the remaining tests. The average multiple correlation was computed for each test pair and that combination with the highest weighted average was selected. This process was repeated at each step until all the remaining tests were included. In this way, the tests were uniquely and optimally weighted for each school in each equation.

These analyses were initially applied to the 18 school samples using primarily FSG criteria. Then, to determine the impact of criterion type on test validity, these same procedures were applied to the eight schools with both FSG and performance criteria. These analyses were independently conducted on FSG and on performance criteria.

Averaging Validities by School Complexity

Research has demonstrated that the magnitude of validity is a function of the complexity of the job. More specifically, as complexity goes up, the validity of general mental ability increases and the validity for psychomotor ability decreases (Hunter, 1980; Schmidt, Hunter, & Dunn, 1987). Consequently, we would predict that the validities of the psychomotor test and any tests that added to the ASVAB g would vary as previously demonstrated in the literature. We also conducted analyses to determine if incremental validity of the new tests varies as a function of complexity.

We determined the complexity of schools by two approaches. First, we used a modification of the complexity coding system developed by Hunter (1986), which is based on the Data-People-Things rating scales of the Dictionary of Occupational Titles (DOT). (Also see Schmidt et al., 1987, for a description of the procedure.) We classified schools on the basis of ratings on the DOT Data scale for the job associated with the school. These ratings for military jobs have been developed by the DoD Crosswalk program. If the associated job for a school had a rating of 0 or 1, the highest ratings of complexity on the DOT scale, the school was categorized as "very high" in complexity. Schools with associated job ratings of 2 on the Data scale were categorized as "high" complexity, while schools with an associated job rating of 3 to 4 were classified as "medium" complexity. Schools were classified as "low" complexity if the associated job rating was 5 or higher. One school, BT/MM, had two associated jobs and ratings. In this case, we averaged the two ratings and rounded this average rating to the highest integer in order to assign a complexity level. After all the schools were categorized by complexity, we computed the weighted average validity for the schools in each category.

As our second approach, we determined complexity based on the length of training time allocated for the curricula of the different schools in the data set. We calculated a correlation coefficient to determine the relationship between length of training time and incremental validities across all schools.

RESULTS

The results of the analysis of the Navy Only data set (i.e., the CTB predictors) will be presented first followed by the results for the Joint Service data (i.e., the ECAT predictors).

Data Verification

We computed a number of preliminary regression analyses on the Navy Only data set to compare results with those previously obtained by Wolfe et al. (1992). We made these comparisons to ensure that we had properly identified the subjects, variables, and associated data in the Navy Only data set.

We calculated the multiple correlation between the ASVAB tests plus the CTB tests and the FSG criterion for eight schools in the data set, and compared these correlations to those calculated by Wolfe et al. (1992). As did these authors, we used post-enlistment ASVAB scores for these analyses. The results showed that six out of the eight multiple correlations differed by one correlation point or less and the other two multiple correlations differed by less than two correlation points. The discrepancies were due to small differences in the subjects included in school samples, which was, in turn, caused by different screening procedures.

Overall Tests of Significance for the Set of CTB Predictors

Table 9 displays the results of the regression analyses which added the set of CTB tests to the ASVAB for each school. The overall hypothesis of no significant increment across schools for these analyses was rejected ($p < .001$).

As shown in Table 9, we obtained findings of significant incremental validity for four schools (AV, GMG, OS, AT). Table 9 also displays multiple correlations that have been fully corrected; that is, corrected for range restriction, shrinkage, and unreliability in the criteria. The delta (▲) R in Table 9 is the difference between the multiple R obtained by ASVAB alone and the multiple R obtained by ASVAB plus the CTB tests. The percent gain in correlation points is also displayed.

The results of the regression analyses that added the set of CTB content dimensions to the set of ASVAB content dimensions are displayed by school in Table 10. We also footnote the content dimension weights in this table. The overall hypothesis of no incremental validity from adding the set of CTB content dimensions was rejected when tested across schools ($p < .001$). As expected, the same four schools which attained significance when the set of CTB tests was added to ASVAB also reached significance when adding the CTB content dimensions. In addition, the probability of a significant increment for the AO school equaled .053. In deciding how to treat the AO school in subsequent analyses, we considered this to be a significant value. It should be noted that this is the same set of five schools for which Wolfe et al. (1992) reported significant findings using post-enlistment ASVAB test scores. Of the remaining four schools that did not reach significance, two of them have sample sizes of about 100.

We also averaged, across schools, the fully corrected multiple Rs obtained for the full model of ASVAB plus all CTB predictors and for the full model of ASVAB content dimensions plus all CTB dimensions. We then calculated the overall incremental validity by subtracting the average multiple R based on the ASVAB component alone from the full model average multiple R. The resulting incremental validities of CTB content dimensions and tests equaled 1.9 correlation points (2.7% gain) and 1.7 correlation points (2.5% gain), respectively.

The findings of significant, incremental validity across schools for the addition of the set of CTB tests and the set of CTB content dimensions allowed us to look at more specific comparisons.

CTB Content Dimensions

Table 11 presents the results of adding each CTB content dimension to the set of ASVAB content dimensions, across schools. As shown, each of the CTB content dimensions added significant validity to the ASVAB. This table also presents the results of deleting each CTB content dimension from the full model of all ASVAB and all CTB content dimensions. The findings for the deletion procedure show that deleting each dimension from the full model resulted in a significant loss of variance. In other words, each dimension contributes significant, unique variance to the full model of ASVAB and ECAT tests.

We present the results of adding the content dimensions by school in Table 12. As displayed, the Working Memory and the Spatial content dimensions added significant validity in four and five schools, respectively. Table 12 also shows the results of deleting each CTB content dimension from the full model of all

ASVAB and CTB content dimensions. These deletion analyses indicate that each of the content dimensions added a significant amount of unique criterion variance for three of the five schools.

Individual CTB Tests

Tables 13 and 14 show the results of accretion and deletion of individual CTB tests across schools. Each of the six CTB tests demonstrate significant incremental validity when added to the ASVAB and analyzed across all schools. Based on the deletion analyses, three of the tests (Sequential Memory, Spatial Perception, and Integrating Detail) did not show any significant, unique variance across schools when compared to the ASVAB plus the remaining CTB predictors. The other three (Mental Counters, Perceptual Speed, and Figural Reasoning) demonstrated significant, unique variance across all schools. This finding indicates there is a substantial degree of correlation between some of the CTB predictors (See intercorrelation matrix presented in Appendix B).

The results of adding and deleting CTB tests for each school are presented in Table 14. This table displays only those schools for which the set of CTB tests yielded incremental validity.

Operational Composites

Table 15 displays the multiple regression results from adding the total set of CTB content dimensions to each school's operational selection composite. The increment attained significance for six of the nine schools. For these six schools, the increments range from 2.3 to 13.3 correlation points (or 2.7% - 28.2%). It should be noted that, of the three schools that did not demonstrate a significant increment, two had sample sizes of about 100. The results of the accretion and deletion analyses are presented in Table 16. Each of the content dimensions added significant, unique variance to the operational composite in three to five schools.

Comparison of Methods for Estimating g

Table 17 shows the test weights used to compute each of the four methods of estimating g. The average intercorrelations and validities corrected for range restriction for the four g scores are presented in Table 18. As shown, there is a very high relationship between the g scores derived from the "psychological g" and those derived from the ASVAB plus CTB g. Table 18 also shows that the AFQT obtains the lowest average validity coefficient. The gain in using ASVAB g over AFQT is 2.7 correlation points.

Overall Tests of Significance for the Set of ECAT Predictors

Table 19 presents the results of the regression analyses on the unrestricted samples which added the set of ECAT tests to the ASVAB for each school. The overall hypothesis of no significant increment across schools for these analyses was rejected ($p \leq .001$). As shown in this table, 10 of the 18 independent school samples obtained significant incremental validity. Table 19 also shows the fully corrected results (i.e., corrected for range restriction, shrinkage, and unreliability in the criteria).

Table 20 displays the results of the regression analyses that added the set of ECAT content dimensions to the set of ASVAB content dimensions by school. The overall hypothesis of no incremental validity across schools of adding the ECAT content dimensions to ASVAB content dimensions was rejected ($p \leq .001$). Nine of the ten schools which obtained significant results when the set of CTB tests were added to ASVAB tests also obtained significant results from adding the ECAT content dimensions to ASVAB content dimensions. The tenth school (BT/MM) obtained a significance level of $p = .133$.

We averaged, across schools, fully corrected multiple correlations for the full model of ASVAB plus all ECAT predictors, as well as the multiple correlations for the ASVAB alone. Average multiple correlations were also computed for the ASVAB content dimensions alone and for the full model of ASVAB and ECAT content dimensions. We then calculated the deltas or incremental validities by subtracting the average multiple correlations based on the ASVAB components alone from the full model average multiple correlations. The incremental validities equaled 1.2 and 1.1 correlation points or a 1.9% and 1.7% gain from adding all ECAT predictors and all ECAT content dimensions, respectively.

The findings of overall significance across schools from adding the set of ECAT predictors and the set of ECAT content dimensions allowed us to evaluate more specific comparisons.

ECAT Content Dimensions

Table 21 displays the results across schools of adding each ECAT content dimension to the set of ASVAB content dimensions. Each of the four ECAT content dimensions added significant incremental validity to the ASVAB content dimensions when tested across schools. In addition, each content dimension showed significant unique variance when deleted from the full model of all ASVAB and ECAT content dimensions.

Table 22 shows that different content dimensions significantly incremented validity in different schools. For example the RM school obtains a significant increment for Working Memory and Perceptual content dimensions but not for the Spatial or Psychomotor content dimensions.

Individual ECAT Tests

For each school, corrected zero-order validity coefficients for ECAT and ASVAB tests are presented in Appendix C. Average validities for each test are also provided. While incremental validities are the object of this investigation, these zero-order coefficients demonstrate substantial predictive power for ASVAB and ECAT tests. The intercorrelation matrix of ASVAB and ECAT predictors is presented in Appendix D. These correlation coefficients have been corrected for restriction in range.

Table 23 presents the results across schools of adding each ECAT test to the ASVAB and deleting each ECAT test from the full model of all ASVAB and ECAT tests. Each of the ECAT tests added a significant amount of validity to the ASVAB when the probability was calculated across all schools. In terms of adding unique variance, Figural Reasoning, Mental Counters and Target Identification reached significance. The remaining tests did not obtain significant results although the probability for the Two-

Hand Tracking test was equal to .055. It should be noted that, in the CTB analyses, Mental Counters, Figural Reasoning, and the CTB Perceptual Speed test also added significant validity to the full model of ASVAB plus new predictors.

Table 24 presents the results of the accretion and deletion analyses by school. The deletion results indicate that each of the ECAT tests provided a unique contribution to validity for one or more schools.

Operational Composites

The multiple regression results from adding the set of ECAT content dimensions to each school's operational composite are displayed in Table 25. We obtained significant increments for 12 school samples.

Table 26 presents the results of the accretion and deletion analyses by school. The working Memory and Spatial content dimensions added significant incremental validity for 11 or more of these school samples. The Perceptual Speed content dimension added significant validity for 7 of these 12 school samples and psychomotor yielded significant validity for 10 of the 12 school samples.

In terms of unique variance, the Spatial dimension obtained significance in the greatest number of schools, that is, 9 out of the 12 samples. It would appear that many of the operational composites could gain predictive power by including other measures. Since percent gain values depend upon the magnitude of the multiple correlation for an operational composite, caution should be used in their interpretation.

Optimal School Composites

Multiple regression analyses allowing all ASVAB and ECAT tests to enter the regression equation were calculated for each school. Table 27 shows, for each school, the ASVAB and ECAT tests with significant regression weights in these resulting equations. The pattern of weights shows the relatively greater strength of the ASVAB in relation to the ECAT and highlights those tests which capture significant, unique variance. Appendix E shows all of the standardized regression weights.

Performance Versus FSG Criteria

To evaluate the impact of criterion type on incremental validity, we analyzed the eight schools that had both performance and FSG criteria. We separately calculated incremental validities for each type of criteria. Table 28 displays the results of these analyses, conducted on both the set of ECAT tests and content dimensions. As shown, the gain in predicting performance criteria, when adding all the ECAT tests and all the ECAT content dimensions is 3.3 and 3.6 correlation points, respectively. The Working Memory and Spatial content dimensions add the most validity to the ASVAB, contributing gains of 2.5 and 2.3 correlation points, respectively. While the multiple Rs are always lower for the performance criteria, the increments are always larger, and usually substantially larger.

In terms of individual tests incrementing performance prediction, Sequential Memory adds the most validity (2.58%), followed by Assembling Objects (2.30%), and Mental Counters (2.20%). Table 29 presents the relative ranking of the ECAT tests based on the magnitude of their incremental validities using performance criteria versus FSG. The intercorrelations between FSG and Performance criteria for the eight schools are as follows: AC=.40, AE=.72, AO=.63, AMS=.75, OS=.89, 13F=.83, 732=.26, 272(2)=.04.

It should be noted that the gain (1.4 correlation points) shown in Table 28 from adding all the ECAT tests and using FSG was very close to the gain (1.2 correlation points) based on all the schools and using FSG.

Comparison of Methods for Estimating g

Table 30 shows the test weights used to compute each of the four methods of estimating g. For the four methods of estimating g scores, Table 31 presents the average intercorrelations and validities corrected for range restriction. As shown, there is an extremely high relationship (.999) between the "psychological g" and the g based on the component analysis of ASVAB plus ECAT predictors. The only note-worthy difference in average validities for the four methods is between AFQT and the other methods. These findings are highly similar to the comparable analyses for the Navy Only data.

Final Battery

Tables 32 and 33 show the results of applying the combined sample regression procedure to select the "best" tests. Table 32 provides the results of this procedure applied to the eight schools with FSG and performance criteria. Table 33 provides results of the analyses applied to all 18 school samples using primarily FSG. These tables display the order of entry for each ASVAB and ECAT test in the multiple regression procedure as well as the cumulative average R at each step. As described earlier, this procedure was applied to data subjected to various levels of correction.

As indicated in Table 32, several ECAT tests are among the first ten tests selected. It is interesting to note that the Two-Hand Tracking test is in the top ten tests for predicting performance criteria only. For the FSG criteria, the order of entry is identical regardless of level of correction. On the other hand, correction for unreliability of the performance criteria (which are more variable than FSG measures in reliability) produces a greater shift in the order in which tests enter the equations.

As Table 33 shows, the Assembling Objects and the Two-Hand Tracking tests are ECAT tests that appear in the top ten tests for all three applications of the procedure. The Figural Reasoning test also appears in the top ten for all but the analysis based on the fully corrected data, for which it enters as the eleventh test. It is interesting to note that the order of entry is nearly the same irrespective of statistical corrections.

Validity and School Complexity

Table 34 presents the classification of schools into DOT-based complexity categories, as described earlier. Table 35 displays the average, fully-corrected validities and incremental validities by complexity category. As expected, the average validity increases as complexity increases. However, the incremental validity from adding the new predictors decreases as complexity increases. As shown, the gains in correlation points are relatively substantial for the five schools in the medium and low complexity categories.

We also calculated the correlation between the amount of incremental validity and the amount of training time (in weeks) required for each school. The obtained correlation (-.48; p = .04) confirmed the negative relationship between ECAT incremental validity and complexity, where complexity is defined by the DOT rating-based approach.

SUMMARY OF RESULTS

1. The ASVAB is a relatively powerful test battery. For the Joint Service data set and the Navy Only data set, the fully corrected multiple correlation for predicting primarily Final School Grade (FSG), averaged over school samples, equaled .646 and .698, respectively.

2. The CTB and ECAT tests obtain substantial zero-order, corrected average validities (ranging from .20 to .41 for the ECAT tests). However, their intercorrelations with the ASVAB tests limit their capability to provide incremental validity.

3. Scores from the two methods to estimate g, which were based on the ASVAB plus experimental tests (ECAT and CTB), correlated .99 with each other. Estimates of g based on ASVAB or ASVAB plus experimental tests produced nearly identical validities. Estimates based on AFQT, which could be perceived to be an estimate of g, produced the lowest average validities.

4. For predicting school achievement (primarily FSG), the overall incremental validity of CTB and ECAT content dimensions was 1.9 correlation points (2.7% gain) and 1.1 correlation points (1.7% gain), respectively. For the Joint Service data, the Spatial dimension contributed more to validity than the other dimensions. For the Navy Only data, the Spatial and Working Memory dimensions contributed equally to validity.

5. For predicting school achievement (primarily FSG), the overall incremental validity from adding the entire CTB and ECAT batteries to the ASVAB is relatively small, when averaged over a number of schools. The CTB added 1.7 correlation points (2.5% gain) and ECAT added 1.2 correlation points (1.9% gain).

6. Each of the CTB and ECAT content dimensions added validity (and in some cases, relatively substantial validity) to the operational school selector composites.

7. The incremental validity of the ECAT tests was greater for performance criteria than for the more knowledge-based FSG criteria, but the absolute validities were smaller for performance-based criteria. The Working Memory content dimension and the Sequential Memory test added the most validity to the ASVAB when predicting performance criteria.

8. The Assembling Objects test and the Two-Hand Tracking test out-performed several ASVAB tests in predicting school achievement when measured primarily by FSG.

9. The amount of incremental validity from the ECAT tests correlated negatively with school complexity. The incremental gains were relatively large for lower complexity schools.

DISCUSSION

The results of this study and many others (Wolfe et al., 1992; Peterson et al. 1992; McHenry, Hough, Toquam, Hansen and Ashworth, 1990; and Busciglio, 1990) demonstrate that the ASVAB accounts for a very substantial amount of variance for school or job criteria. For the Joint Service data set in the present study, using only ASVAB tests, the average fully corrected validity coefficient for predicting school criteria (primarily FSG) across 18 schools equaled .646. While the CTB and ECAT tests obtain substantial zero-order, corrected average validities (e.g., for ECAT, they range from .20 to .41), their intercorrelations with the ASVAB tests limit their capability to provide incremental validity. The average incremental validity from adding the entire CTB or ECAT test battery to ASVAB was 1.7 and 1.2 correlation points, respectively. These results indicate that it is difficult to design predictors which will capture variance, <u>across schools</u>, not accounted for by the ASVAB.

As shown in Table 36, these findings are similar to the findings of several other studies that analyzed various ECAT tests or composites. This table shows incremental validities ranging from about 1 to 5 correlation points for various ECAT tests or composites relative to various combinations of ASVAB tests. These incremental validities were obtained on a number of different criteria (i.e., knowledge tests, hands-on performance tests and rating scales) for a variety of jobs and schools.

In the present study, the Spatial dimension added slightly more validity to the ASVAB than did other ECAT composites. In the CTB analyses, the Spatial dimension tied with the Working Memory dimension for the largest incremental validity. The Assembling Objects test (AO), one component of the Spatial dimension, was among the top ten tests in the combined sample regression analyses based on 18 school samples. In fact, AO preceded four ASVAB tests in entering the optimal equations. All five of the studies listed on Table 36 also found that the Spatial composite captured more incremental variance than did other cognitive composites. In addition, three of the five studies reported that the AO test increased validity more than did other ECAT tests. The Figural Reasoning test was also found to be a promising test in two of the studies (Busciglio, 1990; Peterson et al., 1992). This test was also empirically selected as one of the better tests for predicting criteria (primarily FSG) across schools in our study.

In addition to the spatial tests, Peterson et al. (1992) also indicated that the psychomotor tracking tests increased validity more than certain ASVAB tests, especially when evaluated in terms of differential validity. In the present study, the Two-Hand Tracking test was another one of the "best" 10 tests that were empirically selected in the combined sample regression analyses. In fact, this test preceded five ASVAB tests in order of entry.

It is important to note, however, that with the exception of the Peterson et al. (1992) study, none of the other investigators listed in Table 36 attempted to assess the benefits that the ECAT tests might provide when used for differential classification. A way to evaluate the differential validity of the ECAT tests would be to use simulated assignment algorithms. One analysis would be based on composites containing only ASVAB tests and a second analysis would use composites of ASVAB plus ECAT tests. Increases in mean predicted job or school performance would be attributed to the addition of ECAT tests. Since these results would be based on a limited set of schools/jobs, it would be important to estimate the utility of differential classifications for the entire population of military jobs. We could classify individual jobs in the sample and in the population of military jobs along various dimensions, such as complexity, by using procedures developed in our study. Then we could estimate the benefit for the entire population of military jobs from the sample jobs.

A finding in the present study concerns the dependence of results on the school criterion analyzed. This dependence was illustrated by the differences observed in individual test and composite incremental validities for predicting FSG versus performance criteria. We also observed differences in results by separately applying the combined sample regression procedure to FSG and performance criteria.

A weakness in most of the incremental validity research is that test evaluation and selection focuses solely on the magnitude of validity or incremental validity. Procedures developed in the present study could be refined to select tests based on joint functions of individual test validity, administration time, and other factors such as adverse impact.

Finally, while the Assembling Objects and the Two-Hand Tracking tests are promising new predictors, there are obviously several other technical and policy factors which must be considered before any new test is selected for implementation. For example, the specific weighting of tests for selection and classification composites, gender and racial effects, as well as practice effects must be considered. A discussion of some of these other factors is presented in Appendix F.

CONCLUSIONS

1. The Assembling Objects and Two-Hand Tracking tests can outperform some ASVAB tests in predicting school criteria (as measured primarily by FSG).

2. While the ECAT and CTB tests provide a relatively small amount of incremental validity over the ASVAB when averaged across schools, they provide substantial validity gains for some of the schools.

3. The ECAT Spatial dimension is the most promising dimension for increasing FSG validity over the ASVAB.

RECOMMENDATIONS

1. Evaluate the Classification efficiency of the ECAT tests.

2. Develop procedures to simultaneously evaluate individual test validity, administration time, and other factors such as adverse impact in relation to selecting a test battery.

REFERENCES

Alderton, D.L. (1991). Improved measurement of perceptual speed ability through computer administration. Paper presented at the Annual Meeting of the American Psychological Association, San Francisco, CA.

Alderton, D.L., & Larson, G.E. (1992). ECAT battery: Descriptions, constructs, and factor structure. Paper presented at the Centennial Convention of the American Psychological Association, Washington, D.C.

Arvey, R.D., & Faley, R.H. (1988). Fairness in selecting employees (2nd ed.). New York: Addison-Wesley Publishing Company.

Boldt, R.F., Centra, J.A., & Courtney, R.G. (1986). The validity of various methods of treating multiple SAT scores (College Board Rep. No. 86-4). New York: College Entrance Examination Board.

Brogden, H.E. (1951). Increased efficiency of selection resulting from replacement of a single predictor with several differential predictors. Educational and Psychological measurement, 11, 173-196.

Busciglio, H.H. (1990). The incremental validity of spatial and perceptual-psychomotor tests relative to the armed services vocational aptitude battery (Tech. Rep. No. 883). Alexandria, VA: U.S. Army Research Institute.

Carey, N.B. (1992, July). New predictors of mechanics' job performance: Marine corps findings. Paper presented at the American Psychological Association Symposium, Washington, D.C.

Cascio, W.F. (1989). Using utility analysis to assess training outcomes. In I. Goldstein (Ed.), Training and development in work organizations: Frontiers of industrial and organizational psychology. San Francisco: Jossey-Bass.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/ correlation for the behavioral sciences (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Dunnette, M.D. (1966). Personnel selection and placement. Belmont, CA: Brooks/Cole Publishing Company.

Fisher, R. (1932). Statistical methods for research workers (4th ed.). London: Oliver & Boyd.

Gulliksen, H. (1987). Theory of mental tests. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hartigan, J.A., & Wigdor, A.K. (Eds.) (1989). Fairness in employment testing. Washington, D.C.: National Academy Press.

Horst, P.A. (1954). A technique for the development of a differential prediction battery. Psychological Monograph, 68, No. 9 (Whole No. 380).

Horst, P.A. (1955). A technique for the development of a multiple absolute prediction battery. Psychological Monograph, 69, No. 5 (Whole No. 390).

Hunter, J.E. (1980). Test validation for 12,000 jobs: An application of synthetic validity and validity generalization to the general aptitude test battery (GATB). Washington, D.C.: U.S. Employment Service.

Hunter, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. Journal of Vocational Behavior, 29, 340-362.

Hunter, J.E., & Hunter, R.F. (1983). The validity and utility of alternative predictors of job performance (OPRD Report 83-4). Washington, D.C.: U.S. Office of Personnel Management.

Kieckhaefer, W.F., Ward, D.G., Kusulas, J.W., Cole, D.R., Rupp, L.M., & May, M.H. (1992). Criterion development for 18 technical training schools: The Navy, Army, and Air Force (Final Rep.). San Diego, CA: RGI, Inc.

Larson, G.E., & Alderton, D.L. (1992). Test/retest results for the enhanced computer administered tests (ECAT) battery. Paper presented at the Centennial Convention of the American Psychological Association, Washington, D.C.

Lawley, D. (1943). A note on Karl Pearson's selection formulae. Royal Society of Edinburgh, 62, (Sect. A), B28-30.

McHenry, J.J., Hough, L.M., Toquam, J.L., Hanson, M.A., & Ashworth, S. (1990). Project a validity results: The relationship between predictor and criterion domains. Personnel Psychology, 43, 335-354.

Oppler, S.H., Peterson, N.G., Whetzel, D.L., Steele, D., Childs, R.A., Park, R.A., Rosse, R.L., Rehling, J.F., Brantner, T.M., & Kieckhaefer, W.F. (1992). Selection and classification tests for critical military occupational specialties (Final Tech. Rep.). Washington, D.C.: American Institutes for Research.

Peterson, N., Oppler, S., & Rosse, R. (1992). Optimal battery analyses of project a/career force data. Report prepared for the ASVAB Technical Committee.

Ree, M.J., & Earles, J.A. (1991). The stability of g across different methods of estimation. Intelligence, 15, 271-278.

Schmidt, F.L., Hunter, J.E., & Dunn, W.L. (1987). Potential utility increases from adding new tests to the armed services vocational aptitude battery (Tech. Cont. No. 86-698). San Diego, CA: Navy Personnel Research and Development Center.

Segall, D.O., Kieckhaefer, W.F., & Day, L.E. (1985). Predictive utility evaluation of computerized adaptive testing in the Air Force. San Diego, CA: RGI Inc.

Sympson, J., & Candell, G. (1989). MVCOR: Correction for multivariate restriction of range. Computer Program. San Diego, CA: Navy Personnel Research and Development Center.

Thorndike, R.L. (1949). Personnel selection: Test and selection techniques. New York: John Wiley & Sons.

Wigdor, A.K., & Green, B.F., Jr. (Eds.). (1991). Performance assessment for the workplace (Vol. 1). Washington, DC: National Academy Press.

Willerman, L. (1979). The psychology of individual and group differences. San Francisco: W.H. Freeman and Company.

Wolfe, J.H., Alderton, D.L., & Larson, G.E. (1992). Incremental validity of New Computerized aptitude tests for predicting training performance in nine navy technical schools. Manuscript submitted for publication.

Wolfe, J.H., Alderton, D.L., Larson, G.E., & Held, J.D. (1993). Incremental validity of enhanced computer administered testing (ECAT). (Draft Tech. Rep.). San Diego, CA: Navy Personnel Research and Development Center.

Zeidner, J., & Johnson, C.D. (1989a). The utility of selection for military and civilian jobs. IDA Paper P-2239. Alexandria, VA: Institute for Defense Analyses.

Zeidner, J., & Johnson, C.D. (1989b). The economic benefits of predicting job performance. IDA Paper P-2241. Alexandria, VA: Institute for Defense Analyses.

APPENDIX A

DATA SCREENING PROCEDURES

DATA SCREENING PROCEDURES

A number of screening procedures were applied to the data. First, prior to our analysis, NPRDC screened a variety of score data generated by the new predictors. This screening eliminated scores from the two data sets based on an understanding of the test measures and their logical interrelationships. To eliminate the inappropriate scores, NPRDC inserted missing data values into the proper variable fields. See Wolfe, Alderton, Larson, and Held (1993) for a more specific description of these procedures. Although the Wolfe et al. (1993) elimination rules address the ECAT tests, NPRDC applied the same procedures to scores for corresponding tests in the Navy Only data set.

NPRDC also developed a procedure which compared pre-enlistment and post-enlistment ASVAB scores for Navy subjects and computed a probability for the size of the discrepancy between the two sets of scores. An NPRDC researcher described this procedure as follows:

> (A modification of)...a program that computes a X^2 estimate for the likelihood that ASVAB scores came from the same individual. Essentially, the post-enlistment scores are treated as observed while the pre-enlistment scores are treated as expected values...This procedure was used...to detect individuals that had large <u>drops</u> in post-enlistment scores. Since their focus was only on drops in post-enlistment scores, the resulting statistic was not a true X^2...For the Navy-only and Joint-Service studies, we modified the program to detect unexpected increases as well as decreases in performance. My logic was that unexpectedly large changes in either direction was a flag that either one set of scores was wrong or somehow did not reflect true

> performance and therefore should not be used. By changing the procedure, the resulting statistic became a true X^2 with 10 degrees of freedom. This X^2 value was calculated for every individual with pre and post-enlistment ASVAB data, whether it was paper-and-pencil or CAT-ASVAB. I then calculated the probability of obtaining a X^2 with 10 DoF that large or larger. This probability was then multiplied by the sample size (number of people who took both pre and post enlistment tests). If the expected frequency of this value was < 1 that person was deleted...This is a fairly conservative "motivational" screen. However, for the Navy-only study I looked carefully at some of the people that were screened out and their test scores and other aspects of their performance really did look odd, and most of them were people whose scores dropped dramatically for the post-enlistment test sessions. (Alderton, D.L., personal memo to Pass, J.J., Oct. 29, 1992).

We employed this procedure, excluding cases with a computed probability value less than 1.

Next, we screened cases for zero scores and out-of-range scores on the predictors and criteria. These cases, which also included school failures, were excluded from the analyses. We also eliminated two cases in the 13F school which had inconsistent test data.

In addition, we employed a list-wise deletion of cases for all analyses. Thus, missing data for any variable analyzed resulted in the exclusion of that case from the analysis. Finally, we ensured that we used the same sample across criteria. Subsequently, we deleted cases with missing criteria scores.

APPENDIX B

INTERCORRELATION MATRIX FOR ASVAB AND CTB TESTS

(CORRECTED FOR RANGE RESTRICTION)

INTERCORRELATION MATRIX FOR ASVAB AND CTB TESTS

(CORRECTED FOR RANGE RESTRICTION)

```
ANALYSIS NUMBER  1  LISTWISE DELETION OF CASES WITH MISSING VALUES      N = 4,387¹

CORRELATION MATRIX:

           GS1       AR1       WK1       PC1       NO1       CS1       AS1       MK1       MC1       EI1       XMC       XFR

GS1    1.00000
AR1     .61100   1.00000
WK1     .72010    .59620   1.00000
PC1     .60780    .57420    .73160   1.00000
NO1     .27500    .47030    .32430    .39590   1.00000
CS1     .24870    .39530    .32780    .38580    .64010   1.00000
AS1     .52010    .40040    .43660    .33900    .04690    .05820   1.00000
MK1     .55420    .70690    .49670    .49970    .49610    .40770    .19650   1.00000
MC1     .63760    .61340    .54720    .48520    .22780    .22110    .61800    .49380   1.00000
EI1     .62450    .48670    .53430    .44440    .14520    .14700    .66920    .36960    .63040   1.00000
XMC     .36550    .57750    .35240    .36370    .38130    .36500    .21470    .52570    .43340    .28190   1.00000
XFR     .50800    .60980    .48410    .43610    .32250    .31630    .31400    .57140    .53750    .38950    .57640   1.00000
XSM     .31780    .53040    .34660    .34470    .34650    .35730    .14680    .49400    .37450    .22330    .60010    .54180
XSP     .46220    .50130    .38530    .32170    .18310    .19530    .41800    .43380    .55790    .45940    .43860    .52910
XID     .52370    .59760    .47380    .42040    .26280    .26400    .38690    .53170    .58860    .44180    .53650    .62300
XPS     .18640    .25620    .23700    .28490    .42880    .47010   -.03050    .31740    .11380    .07500    .31540    .25010

           XSM       XSP       XID       XPS

XSM    1.00000
XSP     .38570   1.00000
XID     .49760    .58620   1.00000
XPS     .31210    .16260    .19700   1.00000

DETERMINANT OF CORRELATION MATRIX =      .0000744
```

---

[1] This N reflects the number of Ss in the restricted sample.

APPENDIX C

CORRECTED ZERO-ORDER VALIDITY COEFFICIENTS FOR ECAT AND ASVAB TESTS

CORRECTED ZERO-ORDER VALIDITY COEFFICIENTS

FOR ECAT AND ASVAB TESTS

	GS1:	AR1:	WK1:	PC1:	NO1:	CS1:	AS1:	MK1:	MC1:	EI1:	N:
AC:	0.612	0.702	0.510	0.581	0.455	0.482	0.322	0.745	0.494	0.488	72
AE:	0.577	0.566	0.514	0.448	0.296	0.345	0.436	0.530	0.517	0.539	173
AMS:	0.682	0.652	0.672	0.581	0.393	0.423	0.516	0.594	0.629	0.580	244
AO:	0.513	0.548	0.476	0.466	0.381	0.378	0.355	0.580	0.446	0.442	233
AV:	0.591	0.731	0.584	0.560	0.385	0.353	0.410	0.671	0.513	0.570	197
EM:	0.524	0.600	0.470	0.439	0.349	0.342	0.342	0.572	0.504	0.462	805
EN:	0.598	0.602	0.551	0.520	0.310	0.287	0.551	0.526	0.603	0.584	781
FC:	0.600	0.669	0.581	0.580	0.398	0.382	0.459	0.635	0.616	0.574	727
GM:	0.552	0.655	0.587	0.527	0.356	0.325	0.436	0.592	0.550	0.503	393
HMV:	0.230	0.231	0.244	0.263	0.092	0.107	0.266	0.249	0.303	0.265	554
ITV:	0.316	0.270	0.300	0.297	0.297	0.218	0.220	0.257	0.337	0.283	320
MM:	0.436	0.424	0.336	0.350	0.228	0.232	0.392	0.386	0.439	0.400	837
OS:	0.570	0.680	0.560	0.568	0.462	0.488	0.349	0.669	0.558	0.466	622
RM:	0.543	0.583	0.523	0.475	0.370	0.381	0.254	0.532	0.454	0.409	250
13F:	0.528	0.615	0.526	0.549	0.397	0.432	0.408	0.553	0.550	0.449	819
19K:	0.195	0.189	0.139	0.162	0.180	0.170	0.120	0.206	0.168	0.126	1106
272-1:	0.493	0.554	0.549	0.501	0.297	0.269	0.383	0.466	0.501	0.456	484
732:	0.606	0.669	0.604	0.596	0.439	0.390	0.354	0.624	0.503	0.435	421
AV*:	0.477	0.517	0.451	0.442	0.317	0.312	0.357	0.485	0.467	0.421	

	sm:	sr:	id:	ao:	or:	mc:	ti:	t1:	t2:	N:
AC:	0.489	0.446	0.538	0.521	0.383	0.579	-0.440	-0.314	-0.501	72
AE:	0.483	0.520	0.545	0.422	0.482	0.532	-0.298	-0.436	-0.361	173
AMS:	0.236	0.405	0.458	0.445	0.344	0.174	-0.077	-0.105	-0.073	244
AO:	0.175	0.357	0.387	0.385	0.386	0.268	-0.227	-0.110	-0.039	233
AV:	0.499	0.585	0.590	0.519	0.551	0.515	-0.315	-0.355	-0.430	197
EM:	0.389	0.467	0.405	0.395	0.441	0.452	-0.128	-0.261	-0.331	805
EN:	0.346	0.476	0.474	0.416	0.495	0.368	-0.249	-0.236	-0.304	781
FC:	0.342	0.466	0.495	0.434	0.528	0.431	-0.158	-0.270	-0.373	727
GM:	0.418	0.425	0.466	0.488	0.525	0.486	-0.229	-0.368	-0.420	393
HMV:	0.192	0.200	0.224	0.278	0.283	0.228	-0.175	-0.207	-0.254	554
ITV:	0.305	0.246	0.315	0.262	0.343	0.247	-0.263	-0.347	-0.386	320
MM:	0.248	0.368	0.373	0.412	0.305	0.296	-0.195	-0.230	-0.210	837
OS:	0.439	0.569	0.545	0.497	0.540	0.570	-0.297	-0.382	-0.427	622
RM:	0.406	0.430	0.458	0.378	0.384	0.382	-0.138	-0.217	-0.217	250
13F:	0.491	0.541	0.558	0.541	0.531	0.486	-0.318	-0.404	-0.402	819
19K:	0.162	0.114	0.171	0.172	0.134	0.168	-0.119	-0.145	-0.133	1106
272-1:	0.305	0.352	0.395	0.443	0.397	0.339	-0.173	-0.294	-0.317	484
732:	0.422	0.461	0.375	0.312	0.416	0.401	-0.115	-0.126	-0.176	421
AV*:	0.332	0.394	0.407	0.390	0.402	0.367	-0.200	-0.260	-0.289	

*Weighted average. R's weighted by (n-1).

APPENDIX D

INTERCORRELATION MATRIX FOR ASVAB AND ECAT TESTS

(CORRECTED FOR RANGE RESTRICTION)

INTERCORRELATION MATRIX FOR ASVAB AND ECAT TESTS

(CORRECTED FOR RANGE RESTRICTION)

```
ANALYSIS NUMBER  1  LISTWISE DELETION OF CASES WITH MISSING VALUES        N = 10,154[1]

CORRELATION MATRIX:

                GS1       AR1       WK1       PC1       NO1       CS1       AS1       MK1       MC1       EI1   SMPDCOR    SRPCOR
GS1         1.00000
AR1          .61100   1.00000
WK1          .72010    .59620   1.00000
PC1          .60780    .57420    .73160   1.00000
NO1          .27500    .47030    .32430    .39590   1.00000
CS1          .24870    .39530    .32780    .38580    .64010   1.00000
AS1          .52010    .40040    .43660    .33900    .04690    .05820   1.00000
MK1          .55420    .70690    .49670    .49970    .49610    .40770    .19650   1.00000
MC1          .63760    .61340    .54720    .48520    .22780    .22110    .61800    .49380   1.00000
EI1          .62450    .48670    .53430    .44440    .14520    .14700    .66920    .36960    .63040   1.00000
SMPDCOR      .36450    .53470    .36890    .37270    .34930    .34690    .16960    .49480    .38750    .24130   1.00000
SRPCOR       .50010    .59140    .47320    .44310    .30630    .28900    .30710    .54330    .52920    .38780    .54430   1.00000
IDPCOR       .50220    .56870    .43040    .38940    .25820    .25820    .38060    .51570    .57280    .43090    .49610    .59100
AOPCOR       .47350    .51090    .40020    .36240    .23910    .26840    .38890    .46540    .55430    .42700    .49010    .57540
ORPCOR       .49030    .53720    .44120    .39530    .22620    .24160    .39580    .48410    .56220    .43120    .45980    .54350
MCPCOR       .37170    .55780    .34440    .35450    .37180    .35100    .21050    .51550    .42740    .27140    .63080    .55920
TIDDT       -.31340   -.26280   -.25260   -.22420   -.17850   -.18900   -.22690   -.22830   -.32290   -.23200   -.28110   -.29050
T1MN        -.28800   -.29310   -.24260   -.22590   -.20240   -.19400   -.26260   -.25850   -.36780   -.26770   -.32170   -.34480
T2MN        -.33980   -.33370   -.29480   -.26090   -.19230   -.20970   -.32670   -.27700   -.43790   -.32670   -.33760   -.36970

             IDPCOR    AOPCOR    ORPCOR    MCPCOR     TIDDT      T1MN      T2MN
IDPCOR      1.00000
AOPCOR       .64540   1.00000
ORPCOR       .57260    .57790   1.00000
MCPCOR       .55220    .57190    .50750   1.00000
TIDDT       -.32660   -.36650   -.27730   -.29680   1.00000
T1MN        -.37730   -.37960   -.36750   -.37980    .36210   1.00000
T2MN        -.40340   -.42640   -.40820   -.39010    .38500    .75230   1.00000

DETERMINANT OF CORRELATION MATRIX =      .0000113
```

[1] This N reflects the number of Ss in the restricted sample.

APPENDIX E

STANDARD REGRESSION WEIGHTS FOR THE FULL MODEL OF ASVAB AND ECAT TESTS BY SCHOOL

STANDARD REGRESSION WEIGHTS FOR THE FULL MODEL OF ASVAB AND ECAT TESTS BY SCHOOL

	GS1:	AR1:	WK1:	PC1:	NO1:	CS1:	AS1:	MK1:	MC1:	EI1:	N:
AC:	0.192	0.196	-0.165	0.129	-0.069	0.216	0.067	0.445	-0.153	0.166	72
AE:	0.081	-0.000	0.115	-0.032	-0.040	0.109	0.116	0.180	-0.040	0.151	173
AMS:	0.148	0.077	0.206	0.003	0.020	0.158	0.145	0.158	0.097	0.079	244
AO:	0.072	0.100	0.003	0.094	0.021	0.129	0.081	0.307	-0.110	0.110	233
AV:	-0.005	0.274	0.087	0.082	-0.013	0.029	0.054	0.273	-0.172	0.275	197
EM:	0.096	0.192	0.024	-0.010	0.012	0.091	0.016	0.200	0.065	0.129	805
EN:	0.055	0.123	0.033	0.091	0.031	0.033	0.221	0.155	0.067	0.139	781
FC:	-0.014	0.122	0.032	0.127	0.043	0.072	0.070	0.245	0.084	0.178	727
GM:	-0.051	0.214	0.186	0.032	0.015	0.020	0.137	0.214	0.011	0.094	393
HMV:	-0.113	-0.096	0.010	0.151	-0.052	-0.031	0.139	0.155	0.046	0.052	554
ITV:	0.049	-0.093	0.015	0.051	0.214	-0.033	0.007	-0.010	0.086	0.067	320
MM:	0.117	0.035	-0.124	0.104	0.013	0.062	0.175	0.124	0.056	0.073	837
OS:	0.032	0.111	0.026	0.114	0.015	0.189	0.035	0.234	0.069	0.064	622
RM:	0.238	0.183	0.122	0.012	0.024	0.136	-0.089	0.011	0.082	0.076	250
13F:	-0.008	0.096	0.011	0.149	0.005	0.146	0.131	0.104	0.042	0.014	819
19K:	0.112	-0.006	-0.082	0.038	0.072	0.057	0.059	0.084	0.029	-0.031	1106
272-1:	-0.062	0.160	0.234	0.129	-0.001	0.001	0.060	0.059	0.040	0.078	484
732:	0.124	0.202	0.078	0.140	0.060	0.053	0.086	0.160	-0.003	0.005	421
AV*:	0.048	0.091	0.025	0.084	0.027	0.072	0.094	0.152	0.041	0.077	

	sm:	sr:	id:	ao:	or:	mc:	ti:	t1:	t2:	R:	N:
AC:	0.013	0.088	-0.120	0.121	-0.027	0.067	0.022	0.070	0.000	0.848	72
AE:	0.058	-0.016	0.111	0.047	0.040	0.115	-0.093	-0.118	0.118	0.736	173
AMS:	0.010	-0.003	0.052	0.021	-0.022	-0.033	0.000	0.003	0.011	0.814	244
AO:	-0.108	0.029	0.014	0.104	0.023	-0.013	-0.095	-0.154	0.100	0.691	233
AV:	0.032	-0.000	0.046	0.026	-0.049	0.079	-0.008	0.116	-0.149	0.818	197
EM:	-0.052	0.033	0.039	-0.028	0.014	0.098	0.067	-0.009	-0.003	0.682	805
EN:	-0.065	0.020	0.033	0.000	0.037	0.026	-0.029	-0.002	-0.012	0.747	781
FC:	-0.023	0.045	0.050	0.045	0.045	-0.014	0.053	0.069	-0.080	0.794	727
GM:	-0.068	0.057	0.006	-0.003	0.022	0.133	0.070	-0.025	-0.016	0.750	393
HMV:	0.012	-0.059	-0.029	0.086	0.096	0.040	-0.058	-0.006	-0.094	0.411	554
ITV:	0.140	-0.149	0.063	-0.017	0.101	-0.038	-0.063	-0.075	-0.180	0.522	320
MM:	-0.036	0.109	0.007	0.135	-0.077	-0.003	0.003	-0.113	0.111	0.562	837
OS:	0.018	0.037	0.026	-0.004	0.034	0.117	0.023	0.058	-0.052	0.793	622
RM:	0.099	-0.062	0.094	0.011	0.031	0.102	0.182	0.041	0.031	0.699	250
13F:	0.050	0.067	0.077	0.055	0.074	0.038	-0.019	-0.043	0.011	0.747	819
19K:	0.037	-0.101	0.011	0.059	-0.017	0.014	-0.012	-0.039	0.008	0.276	1106
272-1:	0.012	-0.024	-0.028	0.021	0.102	0.096	-0.076	-0.072	0.042	0.667	484
732:	0.100	0.083	0.033	-0.020	0.069	-0.005	0.028	0.050	0.083	0.780	421
AV*:	0.005	0.010	0.028	0.036	0.027	0.040	0.002	-0.018	-0.003	0.642	

*Weighted average. Entries weighted by (n-19).

APPENDIX F

OTHER ISSUES CONCERNING THE EVALUATION OF TESTS

OTHER ISSUES CONCERNING THE EVALUATION OF TESTS

Consideration of Other Factors

Several other factors may affect incremental validity and usefulness of the experimental tests but remain outside the scope of this project. In this section, we point out several of these factors and briefly discuss their relevance to this project.

Correction for Unreliability of Predictors. Normally when correcting for attenuation, selection researchers correct only for the unreliability of criteria. By doing so, we estimate the relationship of existing predictors with perfectly reliable job effectiveness or productivity. By not correcting predictors for unreliability, we accept that these predictors are the ones we intend to implement and use in making decisions.

In some instances, it may be a relatively simple matter to improve predictor reliability. In such cases, we propose correcting predictors to some reasonable level of reliability, such as .85 or so. Along with such analyses, we propose designing analyses to consider the cost in increased testing time versus the return in incremental validities.

Classification Efficiency. Writers on the use of a test battery for classification purposes point out the gains over a pure selection strategy both in human rewards and in increased job effectiveness. Thorndike (1949) formulates the classification problem where: (a) the goal is to contribute the maximum to the overall effectiveness of the organization as a result of assignments, (b) the number of individuals for assignment equals the number of jobs available, (c) a set of predictors exists, (d) validation data exist for each predictor in each job together with intercorrelations among the predictors and among the job criteria, and (e) weights exist for each job specialty indicating the importance of assigning individuals with the highest aptitude. He then describes procedures for assigning all individuals to maximize the weighted sum of the aptitudes of all individuals in all the jobs. Thorndike emphasizes that we can maximize differential prediction by using a set of predictors each of which: measures a relatively pure trait, remains completely unrelated to other traits, and has as much psychological meaning as possible.

Horst (1954, 1955) differentiates between multiple absolute prediction and differential prediction and presents quantitative procedures for addressing each. Horst (1955) defines multiple absolute prediction as the problem of "selecting from a battery of potential predictors that subset of specified size which will have the higher test prediction efficiency for all the criterion variables, irrespective of how well it differentiates among them". More recently, Zeidner and Johnson (1989a, 1989b) describe how the exclusive focus on predictive validity may reduce the efficiency of the ASVAB as a classification tool. Their work demonstrates the close relationship between classification efficiency and utility.

Brogden (1951) and Dunnette (1966) provide examples demonstrating the effect of differential placement over pure selection on mean job effectiveness scores. For a given predictor battery, higher mean effectiveness scores (i.e., predicted job performance) occur for differential placement (versus pure selection) the more there are: higher selection ratios, larger numbers of jobs, and lower correlations between predictors.

One of the examples presented in Dunnette (1966) describes a situation where: the correlation between predictors and the job performance index is .50., the number of jobs is 4, and the selection ratio is 25%. In that example, differential placement provided a mean job effectiveness score one-half a standard deviation greater than pure selection. To the extent these factors operate for military selection and classification, analyses of classification efficiency will provide different results than the pure selection analyses described above.

Subgroup Differences. Hunter and Hunter (1983) cite meta-analyses and validity generalization work to support their claim that ability tests which are valid for one subgroup are also valid for the others. Arvey and Faley (1988) support this view. Unfortunately, the same body of research they cite suggests considerable adverse impact on subgroups due to real subgroup differences in mean scores on both predictors and performance criteria.

Indeed, Larson and Alderton (1992) report significant mean decrements by gender for five of the ten ECAT tests (three psychomotor and two spatial tests). These findings appear consistent with the summary of literature on sex differences reported by Willerman (1979).

If these findings are typical of what we should expect, the question becomes what approaches the military services should consider in dealing with subgroup differences. One approach would point to the empirical results supporting that the subgroup differences are real and suggest that the services should continue making selection and classification decisions as in the past. A second approach would use only tests showing the least adverse impact to select into the services and reserve the other tests for classification.

A third approach would consider whether other combinations of other predictors (i.e., cognitive or non-cognitive) might provide less adverse impact with only minimal or no decrement in composite validities. This approach would allow the services to achieve more diversity within each occupational specialty.

A deficiency of this approach is its sole reliance on test validity and job performance criteria. A more comprehensive approach would also consider the other benefits and costs for using predictors that reduce adverse impact. One of these other benefits arises from increasing the opportunities for subgroup members to receive a broad range of technical training and experience. This may contribute to improved motivational conditions resulting in improved (i.e., better than prediction based only on aptitude) performance for subgroup members.

One of the other costs could be a negative perception by the majority group of the selection policy. This could result in frustration by affected majority group members.

A fourth approach would use utility analyses to assess the costs and benefits associated with using tests having relatively less adverse impact (and probably less validity) to increase diversity within occupational specialties. Such an approach would consider effect sizes associated with non-cognitive predictor variables such as personality and temperament, as well as the motivational effects of achieving work force diversity in the more complex jobs. This approach would provide a basis for a policy that would maximize both diversity and perceived fairness in the military occupational specialties.

Practice Effects. Here, too, recent research shows converging evidence. Working with ECAT tests, Larson and Alderton (1992) report that "score gains were greatest for speeded and/or psychomotor tests" (p. 8). Working with Army Project A tests, Oppler et al. (1992) report that "practice substantially improves performance on psychomotor and perceptual speed tests" (p. iii). Regarding the effects of test order, Oppler et al. (1992) report little difference in mean scores and reliability of measurement.

These studies leave unanswered questions regarding the meaning of these practice effects. There does exist research to support guarding against coaching on cognitive tests. For example, Boldt, Centra, and Courtney (1986) demonstrate that initial scores on the Scholastic Aptitude Test (SAT) predict college performance better than the highest SAT scores. Navy researchers found similar results using the SAT to predict performance in the Naval Academy. Researchers investigating practice effects on cognitive test batteries in other situations (like the Multiplex Controller Aptitude Test used to select air traffic controllers for the

Federal Aviation Administration) have found similar results. However, we know of little comparable research documenting the effects of practice or coaching on speeded or psychomotor tests. In a personal communication, Ree (1992) states that there are several unpublished Air Force studies and one published Air Force study that documents such effects. Nevertheless, a competing hypothesis is that coaching or practice on speeded and psychomotor tests may act like training and develop true increases in ability.

Number of Predictor Tests Versus Available Test Time. The statement of work directs our analyses of the complete battery to identify a final battery of approximately 10 tests. An alternative approach likely to demonstrate greater utility would focus on identifying a final battery of tests requiring less than three hours of test administration time.

Utility Analyses. Before implementing any new selection system, the Department of Defense should compare the costs of the new system with the marginal increase in performance attributable to that system. In this context, utility analysis (or decision theory) determines the degree to which use of the new selection system improves the quality of individuals selected (and, therefore, their resulting level of job performance) beyond what would have occurred had the services not used that device. This increase in performance due to the new selection procedure yields a marginal utility defined as the payoff (or savings) attributable to introducing the procedure.

In the subsections above, we suggest utility analyses as part of an approach for evaluating classification efficiency and for deciding how best to deal with subgroup differences. Such analyses would also contribute to an overall assessment of the economic benefits of any new selection or classification procedure.

However, there currently exist several lines of criticism oriented toward utility analyses in general. Hartigan and Wigdor (1989) express a number of these concerns in their summary of findings from the National Research Council Committee on the General Aptitude Test Battery. Managers tend to complain that "utility dollars" are ephemeral constructs that rarely materialize long enough to put to work.

While these arguments continue, managers must still make decisions about how best to invest limited resources. Cascio (1989) demonstrates how the capital budgeting approach used by many managers to make investment decisions can be applied to making investments in the human resource field. When combined with assessments of classification efficiency, Cascio's capital budgeting approach offers the most promise for expressing the marginal utility of the new selection and classification procedures using methods and terms familiar to managers.

Table 1

Navy Only: Schools for Navy Only Data Set

Abbreviation	Title
AD	Aviation Machinist's Mate
AMS	Aviation Structural Mechanic-Structures
AO	Aviation Ordinanceman
AV	Avionics Total, consisting of:
	Aviation Electronics Technician (AT)
	Aviation Fire Control Technician (AQ)
	Aviation Antisubmarine Warfare Technician (AX)
BT/MM	Boiler Tech/Machinist, consisting of:
	Boiler Technician (BT)
	Machinist's Mate (MM)
GMG	Gunner's Mate
HM	Hospitalman
HT	Hull Maintenance Technician
OS	Operations Specialist

Table 2

Navy Only: CTB and ASVAB Predictors.

Predictors
CTB Predictors[1]:
Mental Counters (MC)
Sequential Memory (SM)
Integrating Details (ID)
Space Perception (SP) (ASVAB Form 6) (P&P)
Perceptual Speed (PS)
Figural Reasoning (FR) (P&P)
ASVAB Predictors:
General Science (GS)
Arithmetic Reasoning (AR)
Word Knowledge (WK)
Paragraph Comprehension (PC)
Numerical Operations (NO)
Coding Speed (CS)
Auto and Shop Information (AS)
Mathematical Knowledge (MK)
Mechanical Comprehension (MC)
Electronics Information (EI)

---

[1] Scores employed for CTB tests are proportion correct, except Perceptual Speed, which uses the overall rate on three subtests (See Alderton, 1991).

Table 3

Navy Only: Descriptive Statistics for Predictor Variables: ASVAB and CTB Tests[1]

Variable	*n*	Mean	St.Dev.
ASVAB			
GS	4816	52.87	6.94
AR	4821	51.94	6.85
WK	4821	52.58	5.19
PC	4821	53.02	5.66
NO	4821	53.68	6.70
CS	4818	52.29	6.82
AS	4817	53.68	7.84
MK	4821	52.97	6.79
MC	4813	53.93	7.69
EI	4811	51.88	7.46
VE	4821	52.83	4.88
EXPERIMENTAL			
MC	4760	.724	.17
FR	4821	.661	.16
SM	4759	.688	.15
SP	4874	.528	.20
ID	4765	.725	.13
PS	4864	.673	.09

---

[1] Statistics pertain to the entire sample, before screening data.

Table 4

Navy Only: Descriptive Statistics for School Criteria.[1]

School	Criterion	n	Mean	Uncorrected St.Dev.	Uncorrected r_{xx}	Corrected St.Dev.	Corrected R_{xx}
AD	FSG	108	86.91	5.05	.950	6.45	.961
AMS	FSG	94	81.26	3.76	.900	4.61	.920
AO	FSG	112	82.25	6.76	.880	8.49	.904
AV	LAB1+LAB2	240	94.17	2.78	.617	3.02	.647
BT/MM	FSG	868	85.91	5.35	.810	6.11	.834
GMG	FSG	344	85.92	4.81	.920	6.02	.936
HM	FSG	517	82.18	4.02	.930	4.86	.942
HT	FSG	337	90.54	3.12	.910	3.55	.921
	LAB1+LAB2	337	95.32	1.08	.753	1.13	.764
OS	FSG	941	88.13	4.54	.900	5.79	.922

---

[1] Statistics pertain to the final sample, after applying screening procedures described in Appendix A.

Table 5

Joint-Service: Schools for Joint Service Data Set

	Abbreviation	Title
Navy:		
	AC	Air Controlman
	AE	Aviation Electrician's Mate
	AMS*	Aviation Structural Mechanic-Structures
	AO*	Aviation Ordinanceman
	AV*	Avionics Total, consisting of: Aviation Electronics Technician (AT) Aviation Fire Control Technician (AQ) Aviation Antisubmarine Warfare Technician (AX)
	EN	Engineman
	EM	Electrician's Mate
	ET	Electronics Technician[1]
	FC	Fire Control Technician
	GMG*	Gunner's Mate
	BT/MM*	Boiler Technician/Machinist's Mate
	OS*	Operations Specialist
	RM	Radioman
Air Force:		
	272	Air Traffic Controller
	732	Personnel Specialist
Army:		
	13F	Ft. Sill - Artillery Specialist
	11H	Ft. Benning - Tow Missile Specialist
	19K	Ft. Knox - M1 Tank Crewman

---

[1] Although ET is one of the schools in the Joint Service data set, it was not analyzed in this study.

* These schools are also in the Navy Only data set.

Table 6

Joint-Service: ECAT and ASVAB Predictors

Predictors
ECAT Predictors[1]:
Mental Counters (MC)*
Sequential Memory (SM)*
Figural Reasoning (FR) (P&P)*
Integrating Details (ID)*
Assembling Objects (AO) (P&P)
Spatial Orientation (SO) (P&P)
One-Hand Tracking (T1)
Two-Hand Tracking (T2)
Target Identification (TI)
ASVAB Predictors:
General Science (GS)
Arithmetic Reasoning (AR)
Word Knowledge (WK)
Paragraph Comprehension (PC)
Numerical Operations (NO)
Coding Speed (CS)
Auto and Shop Information (AS)
Mathematical Knowledge (MK)
Mechanical Comprehension (MC)
Electronics Information (EI)

---

[1] Scores employed for ECAT tests are proportion correct, except Target Identification which uses the Mean Clipped Easy/Difficult Decision Times, One-Hand Tracking which uses the TR1 Mean 1000*Log(1+RMS(Attempted)), and Two-Hand Tracking which uses the TR2 Mean 1000*Log(1+RMS(Attempted)).

* Also included in CTB.

Table 7

Joint-Service: Descriptive Statistics for Predictor Variables: ASVAB and ECAT Tests[1]

Variable	n	Mean	St.Dev.
ASVAB			
GS	11778	52.91	7.46
AR	11779	53.28	6.98
WK	11779	52.86	5.38
PC	11779	53.08	5.76
NO	11779	54.06	6.65
CS	11779	53.10	6.94
AS	11780	53.46	8.08
MK	11780	54.80	6.93
MC	11780	54.62	7.75
EI	11780	52.32	7.93
VE	11780	53.04	5.11
EXPERIMENTAL			
MC	11359	.716	.181
SM	11416	.681	.138
FR	11413	.658	.194
ID	10656[2]	.754	.128
AO	11451	.619	.196
SO	11260	.509	.248
T1	11458	2782.402	406.821
T2	11458	3657.661	479.458
TI	11441	1.861	.626

1 Statistics pertaining to the entire sample, before screening data.

2 NPRDC indicated that the relatively low n for ID was a result of software failure.

Table 8

Joint-Service: Descriptive Statistics for School Criteria

School Sample	Criterion[1]	n	Mean	Uncorrected St.Dev.	Uncorrected r_{xx}	Corrected St.Dev.	Corrected R_{xx}
AC	FSG	72	84.53	4.75	.900	6.74	.950
	Perf.(Perf.)	72	90.68	4.79	.590	5.17	.641
AE	FSG	173	83.44	5.95	.900	6.97	.930
	Perf.(Sum2)	173	86.63	11.59	.800	12.25	.821
AMS	FSG	244	83.52	4.23	.860	5.89	.930
	Perf.(Perf.)	244	85.11	3.44	.700	3.95	.848
AO	FSG	233	85.84	5.48	.850	6.43	.890
	Perf.(Practical)	233	93.28	5.24	.720	5.43	.739
AV[2]	FSG	197	89.91	4.17	.890	6.14	.950
EN	FSG	781	84.84	4.93	.874	5.93	.910
EM	FSG	805	87.90	4.73	.920	5.71	.950
FC	FSG	727	83.49	5.33	.920	7.68	.960
GMG	FSG	393	85.97	4.82	.920	6.41	.950
BT/MM	FSG	837	82.48	6.50	.890	7.03	.910
OS	FSG	622	88.58	4.50	.850	5.96	.910
	Perf.(Perf.)	622	89.75	4.90	.700	6.12	.808
RM	FSG	250	94.72	2.63	.800	3.03	.850
272(1)	FSG	484	83.26	5.52	.730	6.61	.812
272(2)[3]	FSG	283	83.76	5.21	.760	6.60	.851
	Perf.(Block5A)	283	0.12	1.71	.620	1.82	.665
732	FSG	421	81.67	6.28	.770	8.14	.860
	Perf.(WPM)	421	29.57	5.74	.950	8.41	.952
13F	FSG	819	90.41	4.13	.720	4.97	.810
	Perf.(Firing)	819	87.85	5.74	.580	6.50	.673
11H(H)[4]	Perf.(Evtsum)	554	1728.25	335.65	.960	344.31	.960
11H(I)[4]	Perf.(Evtsum)	320	1734.85	333.83	.960	348.31	.960
19K	Perf.(Average)	1106	1.88	0.11	—[5]	—	—

1 The descriptors in parentheses for the performance criteria correspond to criterion labels used in Kieckhaefer, et al., 1992.

2 We analyzed only those subjects from Group 1 of the AV school. See Kieckhaefer, et al., 1992 for definition of Group 1.

3 Subset of 272(1) sample, included for subsequent analyses.

4 11H school was separated into independent samples, reflecting separate school tracks.

5 Criterion reliability unavailable.

Note: Statistics pertain to the final sample, after applying screening procedures described in Appendix A.

Table 9

Navy Only: CTB Incremental Validities Over ASVAB.

			Uncorrected			Corrected		
School	Criterion	n	R_{ASVAB}[1]	$R_{ASVAB + CTB}$[2]	$p(\Delta R)$	R_{ASVAB}	ΔR	Percent Gain
AD	FSG	108	.512	.551	.487	.722	.000	0.0%
AMS	FSG	94	.498	.531	.719	.895	.000	0.0%
AO	FSG	112	.519	.588	.100	.737	.018	2.4%
AV	LAB1+LAB2	240	.286	.418	.001**	.537	.084	15.6%
BT/MM	FSG	868	.526	.534	.145	.726	.002	0.3%
GMG	FSG	344	.467	.519	.001**	.721	.019	2.6%
HM	FSG	517	.509	.517	.428	.715	.001	0.1%
HT	FSG	337	.427	.444	.449	.617	.001	0.2%
	LAB1+LAB2	337	.369	.441	.001**	.494	.051	10.3%
OS	FSG	941	.453	.499	.000**	.742	.017	2.3%

1 Multiple Correlation for the ten ASVAB tests.
2 Multiple Correlation for the ten ASVAB tests and all six CTB predictors.
** $p \le .01$.

Table 10

Navy Only: CTB Incremental Validities Over ASVAB for Content Dimensions[1].

			Uncorrected			Corrected		
School	Criterion	n	R_{ASVAB}[2]	$R_{ASVAB+CTB}$[3]	$p(\Delta R)$	R_{ASVAB}	ΔR	Percent Gain
AD	FSG	108	.359	.379	.642	.680	.000	0.0%
AMS	FSG	94	.414	.450	.315	.882	.003	0.3%
AO	FSG	112	.485	.538	.053	.744	.014	1.9%
AV	LAB1+LAB2	240	.235	.382	.000**	.518	.100	19.3%
BT/MM	FSG	868	.513	.518	.157	.721	.001	0.1%
GMG	FSG	344	.449	.494	.000**	.719	.017	2.4%
HM	FSG	517	.493	.495	.697	.710	.000	0.0%
HT	FSG	337	.417	.426	.366	.621	.000	0.1%
	LAB1+LAB2	337	.348	.408	.001**	.492	.045	9.1%
OS	FSG	941	.435	.482	.000**	.737	.018	2.4%

1 Test scores comprising Content Dimensions are standardized scores based on the total sample. Equations for the Content Dimensions are as follows: ASVAB Content Dimensions: Verbal = (WK + PC + .5(GS)), Math = (AR + MK), Clerical = (CS + NO), Technical = (AS + EI + MC + .5(GS)); CTB Content Dimensions: Working Memory = (SM + MC + .5(FR)), Spatial Ability = (ID + SP + .5(FR)), Perceptual Speed = (PS).

2 Multiple Correlation for all four ASVAB Content Dimensions.

3 Multiple Correlation for all four ASVAB and all three CTB Content Dimensions.

** $p \le .01$.

Table 11

Navy Only: Fully Corrected Incremental Validities Over ASVAB for Content Dimensions.

Predictor[2]	$R_{ASVAB+1}$	Entering ΔR	P to Enter	$R_{ASVAB+CTB-1}$	Deletion ΔR	P to Delete
Working Memory	.702	.011	.000**	.706	.004	.000**
Spatial Ability	.702	.011	.000**	.707	.003	.039*
Perceptual Speed	.696	.005	.001**	.706	.003	.032*

1 Probabilities pertain to differences between uncorrected multiple correlations.

* $p \leq .05$.
** $p \leq .01$.

Table 12

Navy Only: Fully Corrected Incremental Validities of ASVAB + CTB Over ASVAB for Content Dimensions (by School).

School	Criterion	$R_{ASVAB+1}$	Entering ΔR	P to Enter[1]	$R_{ASVAB+CTB-1}$	Deletion ΔR	P to Delete[1]
Working Memory Composite:							
AO	FSG	.743	.000	.327	.761	.000	.755
AV	LABS1&2	.583	.066	.000**	.604	.014	.044*
GMG	FSG	.731	.012	.001**	.732	.004	.040*
HT	LABS1&2	.511	.019	.009**	.532	.004	.122
OS	FSG	.754	.016	.000**	.748	.008	.000**
Spatial Ability Composite:							
AO	FSG	.764	.020	.006**	.740	.018	.010**
AV	LABS1&2	.571	.053	.001**	.612	.005	.151
GMG	FSG	.727	.008	.005**	.735	.001	.150
HT	LABS1&2	.521	.030	.001**	.521	.015	.014*
OS	FSG	.746	.009	.000**	.754	.001	.033*
Perceptual Speed Composite:							
AO	FSG	.741	.000	.987	.761	.000	.786
AV	LABS1&2	.564	.046	.001**	.590	.027	.006**
GMG	FSG	.725	.006	.013*	.732	.004	.024*
HT	LABS1&2	.496	.004	.124	.523	.013	.020*
OS	FSG	.740	.003	.007**	.755	.000	.190

---

1 Probabilities pertain to differences between uncorrected multiple correlations.

* $p \leq .05$.

** $p \leq .01$.

Table 13

Navy Only: Fully Corrected Incremental Validities Over ASVAB.

Predictor	$R_{ASVAB+1}$	Entering ΔR	P to Enter	$R_{ASVAB+CTB-1}$	Deletion ΔR	P to Delete
Individual Tests:						
Mental Counters	.706	.008	.000**	.713	.003	.013*
Sequential Memory	.702	.004	.000**	.716	.000	.567
ASVAB-6 Space	.700	.002	.021*	.716	.000	.417
Integrating Details	.706	.007	.000**	.715	.001	.115
Perceptual Speed	.703	.005	.003**	.710	.006	.050*
Figural Reasoning	.707	.008	.000**	.714	.002	.011*

1 Probabilities pertain to differences between <u>uncorrected</u> multiple correlations.

* $p \leq .05$.

** $p \leq .01$.

Table 14

Navy Only: Fully Corrected Incremental Validities for Accretion and Deletion of CTB Tests.

School	Criterion	Mental Counters	Sequential Memory	ASVAB-6 Space	Integrating Details	Perceptual Speed	Figural Reasoning
Accretion:							
AO	FSG	.000	.000	.016*	.023**	.000	.000
AV	FSG	.015*	.025*	.003	.036**	.044**	.051**
GMG	FSG	.011**	.002	.001	.009**	.005*	.006**
HT	Labs1&2	.029**	.000	.011*	.023**	.006	.004
OS	FSG	.010**	.005**	.002*	.004**	.002**	.012**
Deletion:							
AO	FSG	.000	.000	.007	.013*	.000	.000
AV	FSG	.000	.000	.000	.005	.029**	.018*
GMG	FSG	.004*	.000	.000	.002	.006*	.001
HT	Labs1&2	.021**	.000	.003	.006	.014*	.000
OS	FSG	.003**	.000	.000	.000	.000	.004**

* $p \leq .05$ for uncorrected coefficients.
** $p \leq .01$ for uncorrected coefficients.

Table 15

Navy Only: CTB Incremental Validities Over ASVAB Operational Composites.

				Uncorrected			Corrected	
School	n	Criterion	R_{ASVAB}[1]	$R_{ASVAB+CTB}$[2]	p(ΔR)	R_{ASVAB}	ΔR	Percent Gain
AD	108	FSG	.387	.411	.492	.709	.000	0.0%
AMS	94	FSG	.296	.328	.587	.861	.000	0.0%
AO	112	FSG	.453	.521	.026*	.734	.020	2.7%
AV	240	LABS 1&2	.129	.361	.000**	.472	.133	28.2%
BT/MM	868	FSG	.433	.460	.000**	.658	.023	3.5%
GMG	344	FSG	.341	.456	.000**	.683	.041	6.0%
HM	517	FSG	.464	.470	.326	.702	.000	0.0%
HT	337	FSG	.250	.282	.110	.512	.021	4.1%
	337	LABS 1&2	.181	.283	.001**	.312	.071	22.8%
OS	941	FSG	.382	.476	.000**	.715	.037	5.2%

1 Correlation coefficient for the school-specific Operational Composite (OP) and the noted criterion. Tests comprising the OP for each school are the following: For AV, AD, GMG, and AO schools, OP = (AR+MK+EI+GS); for BT/MM, OP = (MK+AS); for AMS, OP = (AR+MC+AS); for HT, OP = (VE+MC+AS); for HM, OP = (VE+MK+GS); for OS, OP = (VE+MK+CS).

2 Multiple correlation for the school-specific Operational Composite and all three CTB Content Dimensions with the noted criterion.

* $p \leq .05$.
** $p \leq .01$.

Table 16

Navy Only: Fully Corrected Incremental Validities of ASVAB + CTB Over ASVAB Operational Composites (by School).

School	Operational Composite[1]	Criterion	$R_{ASVAB+1}$	Entering ▲R	P to Enter	$R_{ASVAB+CTB-1}$	Deletion ▲R	P to Delete
Working Memory Composite:								
AD	OP1	FSG	.703	.000	.218	.705	.000	.443
AMS	OP3	FSG	.857	.000	.694	.856	.001	.203
AO	OP1	FSG	.745	.011	.327	.756	.000	.567
AV	OP1	LABS 1&2	.563	.091	.000**	.592	.013	.038*
BT/MM	OP2	FSG	.678	.020	.027*	.676	.005	.010**
GMG	OP1	FSG	.714	.031	.001**	.717	.007	.004**
HM	OP5	FSG	.701	.000	.501	.702	.000	.514
HT	OP4	FSG	.534	.022	.432	.526	.007	.192
HT	OP4	LABS 1&2	.302	.000	.009**	.388	.000	.354
OS	OP6	OP6	.747	.033	.000**	.740	.012	.000**
Spatial Ability Composite:								
AD	OP1	FSG	.704	.000	.576	.706	.000	.665
AMS	OP3	FSG	.857	.000	.180	.857	.000	.268
AO	OP1	FSG	.753	.019	.006**	.743	.011	.025*
AV	OP1	LABS 1&2	.539	.066	.001**	.591	.015	.074
BT/MM	OP2	FSG	.674	.016	.398	.679	.003	.037*
GMG	OP1	FSG	.703	.020	.005**	.720	.004	.041*
HM	OP5	FSG	.701	.000	.697	.702	.000	.588
HT	OP4	FSG	.527	.016	.811	.532	.001	.271
HT	OP4	LABS 1&2	.332	.020	.001**	.352	.031	.009**
OS	OP6	FSG	.739	.024	.000**	.747	.005	.001**
Perceptual Speed Composite:								
AD	OP1	FSG	.706	.000	.732	.701	.002	.276
AMS	OP3	FSG	.858	.000	.573	.857	.000	.684
AO	OP1	FSG	.734	.001	.987	.755	.000	.537
AV	OP1	LABS 1&2	.533	.061	.001**	.573	.032	.003**
BT/MM	OP2	FSG	.662	.004	.441	.681	.001	.545
GMG	OP1	FSG	.698	.015	.013*	.717	.007	.010**
HM	OP5	FSG	.702	.000	.376	.701	.001	.221
HT	OP4	FSG	.512	.000	.089	.535	.000	.552
HT	OP4	LABS 1&2	.339	.027	.124	.332	.051	.008**
OS	OP6	FSG	.715	.000	.007**	.752	.000	.449

---

[1] OP1: (AR+MK+EI+GS) OP2: (MK+AS) OP3: (AR+MC+AS) OP4: (VE+MC+AS) OP5: (VE+MK+GS) OP6: (VE+MK+CS)

[2] Probabilities pertain to uncorrected multiple correlations.

Table 17

Navy Only: Test Weights for Methods of Estimating g.[1]

Test	ASVABg	AFQTg	PSYCHg	ASVABCTBg
ASVAB Tests:				
GS	.161	–	.508	.105
AR	.159	1.00	.610	.114
WK	.158	–	.484	.102
PC	.149	–	.436	.097
NO	.100	–	.323	.071
CS	.094	–	.316	.068
AS	.117	–	.314	.073
MK	.141	1.00	.571	.103
MC	.151	–	.538	.104
EI	.139	–	.390	.089
VE	–	2.00	–	–
CTB Tests:				
MC	–	–	.576	.092
SM	–	–	.542	.085
ID	–	–	.623	.102
SP	–	–	.529	.089
PS	–	–	.250	.051
FR	–	–	.750	.102

---

[1] For all methods except AFQT, the test scores were standardized to z scores based on the 1991 DoD applicant population.

Table 18

Navy Only: Average Corrected Intercorrelations and Validities of g Score Composites for All Schools.

	r[1]	ASVABg	AFQTg	PSYCHg
ASVABg	.685	--		
AFQTg	.658	.942	--	
PSYCHg	.690	.945	.914	--
ASVABCTBg	.694	.962	.924	.998

---

[1] Average validity coefficient using primarily FSG across all schools.

Table 19

Joint-Service: ECAT Incremental Validities Over ASVAB

School Sample	Criterion	n	Uncorrected			Corrected		
			$R_{ASVAB}{}^1$	$R_{ASVAB+ECAT}{}^2$	$p(\Delta R)$	R_{ASVAB}	ΔR	Percent Gain
AC	FSG	72	.627	.659	.921	.827	.000	0.0%
	Perf.	72	.351	.508	.416	.409	.012	2.9%
AE	FSG	173	.536	.607	.025*	.695	.026	3.8%
	Perf.	173	.291	.404	.121	.398	.041	10.3%
AMS	FSG	244	.582	.586	.997	.833	.000	0.0%
	Perf.	244	.364	.411	.368	.606	.002	0.4%
AO	FSG	233	.491	.530	.233	.690	.006	0.8%
	Perf.	233	.294	.355	.377	.385	.005	1.3%
AV	FSG	197	.504	.533	.581	.818	.000	0.0%
EN	FSG	781	.594	.600	.510	.776	.000	0.0%
EM	FSG	805	.455	.471	.087	.687	.003	0.5%
FC	FSG	727	.456	.482	.007**	.800	.005	0.6%
GMG	FSG	393	.442	.478	.064	.748	.006	0.8%
BT/MM	FSG	837	.410	.446	.000**	.556	.018	3.3%
OS	FSG	622	.571	.592	.007**	.818	.006	0.7%
	Perf.	622	.545	.585	.000**	.818	.017	2.1%
RM	FSG	250	.510	.564	.023*	.704	.020	2.9%
272(1)	FSG	484	.403	.451	.005**	.705	.016	2.3%
272(2)	FSG	283	.446	.482	.239	.753	.004	0.5%
	Perf.	283	.261	.380	.007**	.469	.065	13.8%
732	FSG	421	.536	.584	.000**	.811	.017	2.0%
	Perf.	421	.300	.420	.000**	.332	.094	28.4%
13F	FSG	819	.544	.599	.000**	.793	.029	3.7%
	Perf.	819	.444	.472	.003**	.740	.013	1.7%
11H(H)	Perf.	554	.297	.351	.013*	.353	.028	7.8%
11H(I)	Perf.	320	.334	.455	.000**	.405	.080	19.9%
19K	Perf.	1106	.208	.227	.358	.243	.002	0.6%

1 Multiple Correlation for the ten ASVAB subtests.

2 Multiple Correlation for the ten ASVAB subtests and all nine ECAT predictors.

* $p \leq .05$.

** $p \leq .01$.

Table 20

Joint-Service: ECAT Incremental Validities Over ASVAB for Content Dimensions[1]

School Sample	Criterion	n	Uncorrected			Corrected		
			R_{ASVAB}[2]	$R_{ASVAB + ECAT}$[3]	p(▲R)	R_{ASVAB}	▲R	Percent Gain
AC	FSG	72	.577	.593	.777	.814	.000	0.0%
	Perf.	72	.270	.426	.093	.421	.059	13.9%
AE	FSG	173	.517	.573	.006**	.695	.029	4.1%
	Perf.	173	.236	.367	.006**	.389	.084	21.6%
AMS	FSG	244	.565	.567	.909	.831	.000	0.0%
	Perf.	244	.339	.359	.443	.584	.000	0.0%
AO	FSG	233	.468	.493	.121	.689	.007	1.0%
	Perf.	233	.261	.307	.172	.396	.016	4.0%
AV	FSG	197	.454	.462	.797	.807	.000	0.0%
EN	FSG	781	.589	.592	.412	.818	.004	0.5%
EM	FSG	805	.447	.456	.064	.686	.003	0.4%
FC	FSG	727	.441	.460	.005**	.797	.004	0.5%
GMG	FSG	393	.423	.445	.061	.747	.004	0.6%
BT/MM	FSG	837	.395	.407	.133	.550	.003	0.5%
OS	FSG	622	.558	.576	.001**	.814	.006	0.7%
	Perf.	622	.524	.564	.000**	.809	.019	2.3%
RM	FSG	250	.482	.532	.003**	.700	.024	3.4%
272(1)	FSG	484	.388	.422	.003**	.703	.013	1.8%
272(2)	FSG	283	.414	.432	.262	.742	.002	0.3%
	Perf.	283	.217	.343	.000**	.468	.075	15.9%
732	FSG	421	.532	.571	.000**	.814	.015	1.9%
	Perf.	421	.265	.380	.000**	.329	.089	27.0%
13F	FSG	819	.527	.586	.000**	.785	.034	4.3%
	Perf.	819	.425	.448	.001**	.730	.013	1.7%
11H(H)	Perf.	554	.259	.305	.004**	.333	.027	8.2%
11H(I)	Perf.	320	.308	.418	.000**	.402	.076	18.8%
19K	Perf.	1106	.190	.194	.739	.238	.000	0.0%

[1] Test scores comprising Content Dimensions are standardized scores based on the 1991 DoD applicant population.
Equations for the Content Dimensions are as follows: ASVAB Content Dimensions: Verbal = (WK + PC + .5(GS)), Math = (AR + MK), Clerical = (CS + NO), Technical = (AS + EI + MC + .5(GS)); ECAT Content Dimensions: Working Memory = (SM + MC + .5(FR)), Spatial Ability = (ID + AO + SO + .5(FR)), Perceptual Speed = (TI), Psychomotor = (T1 + T2).

[2] Multiple Correlation for the four ASVAB Content Dimensions.

[3] Multiple Correlation for the four ASVAB Content Dimensions and all four ECAT Content Dimensions.

** $p \leq .01$.

Table 21

Joint-Service: Fully Corrected Incremental Validities over ASVAB for Content Dimensions

Predictor	$R_{ASVAB+1}$	Entering ▲R	P to Enter[1]	$R_{ASVAB+ECAT-1}$	Deletion ▲R	P to Delete[1]
Working Memory	.647	.005	.000**	.653	.000	.000**
Spatial Ability	.649	.007	.000**	.652	.002	.001**
Perceptual Speed	.645	.002	.000**	.652	.001	.000**
Psychomotor Skill	.648	.005	.000**	.651	.002	.001*

[1] Probabilities pertain to differences between <u>uncorrected</u> multiple correlations.

* $p \leq .05$.

** $p \leq .01$.

Table 22

Joint-Service: Fully Corrected Incremental Validities of ASVAB + ECAT over ASVAB for Content Dimensions (by School).

School	Criterion	$R_{ASVAB+1}$	Entering ΔR	P to Enter[1]	$R_{ASVAB+ECAT-1}$	Deletion ΔR	P to Delete[1]
Working Memory Composite:							
AE	FSG	.721	.025	.001**	.721	.003	.146
	Perf.	.477	.088	.000**	.454	.020	.050*
FC	FSG	.797	.000	.422	.801	.000	.364
OS	FSG	.820	.006	.000**	.818	.003	.005**
	Perf.	.826	.017	.000**	.824	.005	.002**
RM	FSG	.710	.010	.014*	.718	.006	.047*
272(1)	FSG	.710	.007	.006**	.714	.001	.185
732	FSG	.821	.007	.001**	.826	.004	.009**
	Perf.	.380	.051	.000**	.400	.018	.007**
13F	FSG	.809	.024	.000**	.816	.002	.012*
	Perf.	.741	.011	.000**	.740	.002	.061
11H(H)	Perf.	.338	.006	.069	.362	.000	.855
11H(I)	Perf.	.418	.016	.019*	.480	.000	.664
Spatial Ability Composite:							
AE	FSG	.720	.024	.002**	.720	.004	.128
	Perf.	.457	.068	.002**	.468	.005	.212
FC	FSG	.800	.003	.003**	.797	.003	.001**
OS	FSG	.818	.004	.001**	.820	.001	.178
	Perf.	.824	.015	.000**	.825	.003	.012*
RM	FSG	.705	.006	.052	.723	.001	.227
272(1)	FSG	.711	.009	.003**	.714	.001	.166
732	FSG	.818	.004	.014*	.828	.001	.117
	Perf.	.371	.042	.000**	.405	.013	.011*
13F	FSG	.816	.031	.000**	.810	.009	.000**
	Perf.	.738	.009	.001**	.743	.000	.364
11H(H)	Perf.	.348	.016	.007**	.357	.003	.149
11H(I)	Perf.	.421	.018	.011*	.480	.000	.586

1 Probabilities pertain to differences between uncorrected multiple correlations.
* $p \leq .05$.
** $p \leq .01$.

Table 22 (Continued)

School	Criterion	$R_{ASVAB+1}$	Entering ΔR	P to Enter[1]	$R_{ASVAB+ECAT-1}$	Deletion ΔR	P to Delete[1]
<u>Perceptual Speed Composite:</u>							
AE	FSG	.718	.023	.035*	.723	.001	.209
	Perf.	.404	.015	.101	.476	.000	.475
FC	FSG	.797	.001	.122	.799	.001	.029*
OS	FSG	.814	.000	.729	.820	.000	.420
	Perf.	.810	.001	.102	.829	.000	.953
RM	FSG	.707	.007	.029*	.710	.014	.004**
272(1)	FSG	.709	.006	.010*	.713	.002	.096
732	FSG	.814	.000	.378	.829	.000	.303
	Perf.	.326	.000	.964	.420	.000	.706
13F	FSG	.787	.002	.014*	.819	.000	.616
	Perf.	.731	.001	.098	.743	.000	.581
11H(H)	Perf.	.340	.008	.045*	.359	.001	.248
11H(I)	Perf.	.423	.021	.008**	.476	.002	.236
<u>Psychomotor Skill Composite:</u>							
AE	FSG	.704	.008	.048*	.726	.000	.879
	Perf.	.412	.023	.060	.480	.000	.850
FC	FSG	.797	.000	.491	.801	.000	.710
OS	FSG	.814	.000	.290	.821	.000	.948
	Perf.	.812	.002	.030*	.829	.000	.719
RM	FSG	.698	.000	.829	.724	.000	.441
272(1)	FSG	.706	.004	.039*	.715	.000	.319
732	FSG	.818	.004	.011*	.822	.008	.001**
	Perf.	.342	.013	.027*	.385	.033	.000**
13F	FSG	.793	.008	.000**	.817	.000	.367
	Perf.	.737	.007	.002**	.741	.001	.119
11H(H)	Perf.	.356	.023	.001**	.350	.010	.024*
11H(I)	Perf.	.480	.077	.000**	.436	.042	.000**

[1] Probabilities pertain to differences between <u>uncorrected</u> multiple correlations.

* $p \le .05$.

** $p \le .01$.

Table 23

Joint-Service: Fully Corrected Incremental Validities over ASVAB (Individual Tests)

Predictor	$R_{ASVAB+1}$	Entering ▲R	P to Enter	$R_{ASVAB+ECAT-1}$	Deletion ▲R	P to Delete
Individual Tests:						
Mental Counters	.646	.004	.000**	.654	.000	.020*
Sequential Memory	.645	.003	.000**	.654	.000	.063
Figural Reasoning	.645	.003	.000**	.653	.002	.019*
Integrating Details	.645	.003	.000**	.655	.000	.534
Assembling Objects	.647	.004	.000**	.654	.000	.405
Spatial Orientation	.646	.004	.000**	.654	.001	.081
One-Hand Tracking	.647	.004	.000**	.654	.000	.086
Two-Hand Tracking	.647	.004	.000**	.654	.001	.055
Target Identification	.645	.002	.000**	.654	.001	.006**

1 Probabilities pertain to differences between uncorrected multiple correlations.
* $p \leq .05$.
** $p \leq .01$.

Table 24

Joint-Service: Fully Corrected Incremental Validities for Accretion and Deletion of ECAT Tests

School	Criterion	MC	SM	FR	ID	AO	SO	T1	T2	TI
Accretion:										
AE	FSG	.024**	.015**	.007	.014*	.012*	.008*	.015**	.002	.011*
	Perf.	.054**	.051**	.054**	.023	.028*	.015	.022	.007	.011
FC	FSG	.000	.000	.001*	.002	.002*	.001	.000	.000*	.000
BT/MM	FSG	.001	.000	.009**	.002	.010**	.000	.002	.000	.000
OS	FSG	.007**	.002*	.002*	.002*	.001	.001	.000	.000	.000
	Perf.	.014**	.007**	.004**	.005**	.006**	.006**	.001	.002	.000
RM	FSG	.005	.006	.000	.002	.000	.001	.000	.000	.012**
272(1)	FSG	.010**	.003	.000	.001	.004*	.011**	.004*	.002	.007**
732	FSG	.001	.008**	.007**	.001	.000	.004*	.003*	.004*	.000
	Perf.	.022**	.041**	.018*	.024**	.021**	.012*	.003	.032**	.000
13F	FSG	.014**	.012**	.014**	.016**	.015**	.014**	.008**	.005**	.002*
	Perf.	.004*	.010**	.005**	.003*	.004*	.012*	.009**	.003*	.002
11H(H)	Perf.	.008*	.002	.000	.001	.015**	.014**	.012*	.021**	.009*
11H(I)	Perf.	.002	.019*	.000	.010*	.002	.016*	.058**	.071**	.017*
Deletion:										
AE	FSG	.001	.000	.000	.002	.000	.000	.002	.002	.003
	Perf.	.000	.000	.003	.000	.000	.000	.000	.000	.000
FC	FSG	.000	.000	.001	.001	.001	.000	.001	.002*	.001*
BT/MM	FSG	.000	.000	.005**	.000	.006**	.002	.005**	.004*	.000
OS	FSG	.004**	.000	.000	.000	.000	.004	.000	.000	.000
	Perf.	.005**	.000	.000	.000	.000	.002*	.000	.000	.000
RM	FSG	.002	.002	.000	.001	.000	.000	.000	.000	.018**
272(1)	FSG	.003	.000	.000	.000	.000	.005*	.001	.000	.003
732	FSG	.000	.003*	.002*	.000	.000	.001	.000	.001	.000
	Perf.	.000	.011*	.000	.006	.000	.000	.000	.036**	.000
13F	FSG	.000	.001	.001	.002*	.001	.002*	.000	.000	.000
	Perf.	.000	.003*	.000	.000	.000	.000	.003*	.000	.000
11H(H)	Perf.	.000	.000	.000	.000	.003	.004	.000	.003	.002
11H(I)	Perf.	.000	.008	.008	.000	.000	.002	.000	.011*	.001

* $p \leq .05$ for uncorrected coefficients.

** $p \leq .01$ for uncorrected coefficients.

Table 25

Joint-Service: ECAT Incremental Validities Over ASVAB Operational Composites

			Uncorrected			Corrected		
School Sample	Criterion	n	R_{ASVAB}[1]	$R_{ASVAB + ECAT}$[2]	p(▲R)	R_{ASVAB}	▲R	Percent Gain
AC	FSG	72	.542	.573	.504	.822	.000	0.0%
	Perf.	72	.058	.285	.245	.294	.000	0.0%
AE	FSG	173	.451	.546	.000**	.652	.054	8.2%
	Perf.	173	.230	.366	.000**	.421	.070	16.7%
AMS	FSG	244	.386	.401	.504	.744	.001	0.1%
	Perf.	244	.238	.283	.192	.535	.007	1.4%
AO	FSG	233	.434	.459	.178	.674	.002	0.3%
	Perf.	233	.181	.235	.250	.355	.000	0.0%
AV	FSG	197	.437	.462	.238	.803	.003	0.3%
EN	FSG	781	.523	.543	.000**	.738	.031	4.3%
EM	FSG	805	.401	.421	.003**	.665	.007	1.1%
FC	FSG	727	.381	.422	.000**	.773	.010	1.3%
GMG	FSG	393	.352	.403	.002**	.721	.014	1.9%
BT/MM	FSG	837	.353	.378	.001**	.527	.013	2.4%
OS	FSG	622	.486	.558	.000**	.780	.032	4.1%
	Perf.	622	.436	.546	.000**	.758	.059	7.8%
RM	FSG	250	.414	.484	.001**	.666	.030	4.5%
272(1)	FSG	484	.303	.392	.000**	.697	.030	4.5%
272(2)	FSG	283	.303	.379	.003**	.702	.015	2.2%
	Perf.	283	.173	.324	.000**	.434	.102	23.4%
732	FSG	421	.269	.436	.000**	.654	.083	12.7%
	Perf.	421	.212	.356	.000**	.293	.108	36.9%
13F	FSG	819	.503	.568	.000**	.739	.064	8.7%
	Perf.	819	.402	.428	.000**	.701	.019	2.7%
11H(H)	Perf.	554	.237	.289	.003**	.318	.029	9.0%
11H(I)	Perf.	320	.251	.394	.000**	.363	.097	26.6%
19K	Perf.	1106	.158	.172	.300	.218	.002	0.7%

[1] Correlation coefficient for the school-specific Operational Composite (OP) and the noted criterion. Tests comprising the composition of OP for each school are the following: For AV, AO, and GMG schools, OP = (AR+MK+EI+GS); for AE, AC, and EM, OP = (AR+(2*MK)+GS); BT/MM, OP = (MK+AS); for AMS, OP = (AR+MC+AS); for RM and OS, OP = (VE+MK+CS); for 11H(H), 11H(I)H, and 19K, OP = (CS+AR+MC+AS); for 13F, OP = (AR+CS+MC+MK); for 732, OP = (NO+CS+VE); for 272(1) and 272(2), OP = (VE+AR).

[2] Multiple correlation for the school-specific Operational Composite and all four ECAT Content Dimensions with the noted criterion.

* $p \le .05$.

** $p \le .01$.

Table 26

Joint-Service: Fully Corrected Incremental Validities of ASVAB + ECAT Over ASVAB Operational Composites (by School).

School Sample	Operational Composite[1]	Criterion	$R_{ASVAB+1}$	Entering ▲R	P to Enter	$R_{ASVAB+ECAT-1}$	Deletion ▲R	P to Delete
Working Memory Composite:								
AE	OP2	FSG	.686	.034	.000**	.704	.002	.187
	OP2	Perf.	.489	.068	.001**	.475	.016	.054
EN	OP3	FSG	.755	.017	.015*	.769	.000	.548
EM	OP2	FSG	.669	.005	.003**	.671	.001	.106
FC	OP1	FSG	.774	.002	.032*	.783	.000	.553
GMG	OP1	FSG	.733	.012	.000**	.731	.004	.042*
BT/MM	OP3	FSG	.535	.007	.004**	.541	.000	.857
OS	OP5	FSG	.804	.024	.000**	.807	.004	.001**
	OP5	Perf.	.803	.045	.000**	.810	.007	.000**
RM	OP5	FSG	.681	.015	.003**	.695	.002	.109
272(1)	OP9	FSG	.675	.009	.001**	.697	.000	.294
732	OP8	FSG	.712	.059	.000**	.730	.007	.002**
	OP8	Perf.	.338	.045	.000**	.373	.029	.002**
13F	OP7	FSG	.788	.050	.000**	.800	.003	.025*
	OP7	Perf.	.717	.015	.000**	.717	.003	.077
11H(H)	OP6	Perf.	.319	.001	.120	.349	.000	.720
11H(I)	OP6	Perf.	.378	.015	.025*	.462	.000	.724
Spatial Ability Composite:								
AE	OP2	FSG	.697	.045	.000**	.695	.011	.034*
	OP2	Perf.	.476	.056	.002**	.485	.006	.200
EN	OP3	FSG	.768	.030	.000**	.757	.012	.001**
EM	OP2	FSG	.670	.006	.003**	.670	.002	.064
FC	OP1	FSG	.781	.009	.000**	.776	.007	.000**
GMG	OP1	FSG	.730	.009	.002**	.734	.001	.180
BT/MM	OP3	FSG	.542	.014	.000**	.535	.005	.009**
OS	OP5	FSG	.808	.028	.000**	.804	.007	.000**
	OP5	Perf.	.810	.052	.000**	.806	.011	.000**
RM	OP5	FSG	.685	.019	.002**	.686	.011	.026*
272(1)	OP9	FSG	.688	.022	.000**	.692	.004	.018*
732	OP8	FSG	.726	.073	.000**	.711	.026	.000**
	OP8	Perf.	.305	.012	.007**	.398	.003	.074
13F	OP7	FSG	.799	.060	.000**	.791	.012	.000**
	OP7	Perf.	.714	.013	.000**	.719	.001	.245
11H(H)	OP6	Perf.	.336	.018	.003**	.340	.007	.060
11H(I)	OP6	Perf.	.386	.023	.006**	.462	.000	.452

---

1 OP1: (AR+MK+EI+GS) OP2: (AR+(2*MK)+GS) OP3: (MK+AS) OP4: (AR+MC+AS) OP5: (VE+MK+CS)
OP6: (CS+AR+MC+AS) OP7: (AR+CS+MC+MK) OP8: (NO+CS+VE) OP9: (VE+AR).

2 Probabilities pertain to uncorrected multiple correlations.

Table 26 (Continued)

School Sample	Operational Composite[1]	Criterion	$R_{ASVAB+1}$	Entering ▲R	P to Enter	$R_{ASVAB+ECAT-1}$	Deletion ▲R	P to Delete
<u>Perceptual Speed Composite:</u>								
AE	OP2	FSG	.669	.017	.009**	.703	.003	.146
	OP2	Perf.	.432	.011	.067	.493	.000	.422
EN	OP3	FSG	.738	.000	.026*	.769	.000	.341
EM	OP2	FSG	.664	.000	.278	.670	.002	.025*
FC	OP1	FSG	.772	.000	.277	.781	.002	.022*
GMG	OP1	FSG	.720	.000	.410	.733	.002	.077
BT/MM	OP3	FSG	.530	.003	.028*	.540	.000	.444
OS	OP5	FSG	.781	.001	.119	.812	.000	.403
	OP5	Perf.	.764	.006	.005**	.818	.000	.965
RM	OP5	FSG	.664	.000	.176	.687	.010	.011*
272(1)	OP9	FSG	.681	.015	.000**	.692	.004	.039*
732	OP8	FSG	.655	.001	.439	.737	.000	.415
	OP8	Perf.	.287	.000	.587	.403	.000	.527
13F	OP7	FSG	.766	.027	.002**	.803	.000	.257
	OP7	Perf.	.704	.003	.033*	.720	.000	.319
11H(H)	OP6	Perf.	.324	.006	.036*	.347	.000	.255
11H(I)	OP6	Perf.	.395	.031	.002**	.455	.005	.134
<u>Psychomotor Skill Composite:</u>								
AE	OP2	FSG	.677	.025	.004**	.708	.000	.729
	OP2	Perf.	.443	.023	.028*	.497	.000	.946
EN	OP3	FSG	.748	.010	.005**	.768	.001	.210
EM	OP2	FSG	.665	.000	.203	.672	.000	.589
FC	OP1	FSG	.774	.002	.041*	.782	.001	.261
GMG	OP1	FSG	.724	.003	.033*	.735	.000	.413
BT/MM	OP3	FSG	.530	.002	.037*	.541	.000	.675
OS	OP5	FSG	.785	.006	.003**	.812	.000	.706
	OP5	Perf.	.773	.015	.000**	.817	.000	.314
RM	OP5	FSG	.662	.000	.632	.698	.000	.553
272(1)	OP9	FSG	.681	.014	.000**	.694	.002	.092
732	OP8	FSG	.654	.001	.855	.733	.004	.009**
	OP8	Perf.	.323	.030	.003**	.341	.061	.000**
13F	OP7	FSG	.770	.031	.000**	.803	.000	.277
	OP7	Perf.	.709	.007	.001**	.719	.001	.109
11H(H)	OP6	Perf.	.340	.021	.002**	.337	.010	.025*
11H(I)	OP6	Perf.	.456	.093	.000**	.412	.048	.525

---

[1] OP1: (AR+MK+EI+GS) OP2: (AR+(2*MK)+GS) OP3: (MK+AS) OP4: (AR+MC+AS) OP5: (VE+MK+CS) OP6: (CS+AR+MC+AS) OP7: (AR+CS+MC+MK) OP8: (NO+CS+VE) OP9: (VE+AR).

[2] Probabilities pertain to <u>uncorrected</u> multiple correlations.

Table 27

Joint-Service: Significance Levels of Standardized Regression Weights in Each School's Full Model Equation

		ASVAB TESTS										ECAT TESTS								
School	n	GS	AR	WK	PC	NO	CS	AS	MK	MC	EI	SM	SR	ID	AO	OR	MC	TI	T1	T2
AC	72								*											
AE	173																			
AO	233								**										*	
AMS	244	*		*			**		*											
AV	197		**						**	*	**									
EM	805	*	**				*		**	*	**							*		
EN	781		**		*			**	**		**									
FC	727		*		**		*		**	*	**							*		*
GMG	393		**	**				**	**								*			
BT/MM	837	*		*	*			**	*				**		**				**	*
OS	622		*		*		**		**								**			
RM	250	**	*															**		
13F	819		*		**		**	**	**					*		*				
11H (H)	554				*			*	*											
11H (I)	320					*														*
19K	1106	*											*							
272 (1)	484		*	**												*				
732	421	*	**		*				**			*	*							

Table 28

Joint-Services: ECAT validities for Predicting Performance and FSG Criteria[1].

Predictor	Criteria	R_{ASVAB}	$R_{ASVAB + ECAT}$	ΔR	Percent Gain
All Tests/Content Dimensions Added to ASVAB:					
Individual Tests:	FSG	.7872	.8011	.0139	1.77%
	Performance	.6033	.6364	.0331	5.49%
Content Dimensions:	FSG	.7828	.7983	.0155	1.98%
	Performance	.5955	.6320	.0365	6.13%
Individual Content Dimensions added to ASVAB:					
Working Memory:	FSG	.7828	.7933	.0105	1.34%
	Performance	.5955	.6207	.0252	4.23%
Spatial:	FSG	.7828	.7949	.0121	1.55%
	Performance	.5955	.6182	.0227	3.81%
Perceptual:	FSG	.7828	.7853	.0025	.32%
	Performance	.5955	.5979	.0024	.41%
Psychomotor:	FSG	.7828	.7859	.0031	.40%
	Performance	.5955	.6033	.0078	1.31%

---

[1] Multiple correlations are fully corrected and averaged across eight schools (i.e., AC, AE, AMS, AO, OS, 272(2), 732, and 13F).

Table 29

Joint-Service: Relative Ranking and Percent Gain Over ASVAB for Each ECAT Test for FSG and Performance Criteria[1]

Test	FSG	Performance
SM	3 (.76)	1 (2.59)
FR	4.5 (.75)	4 (2.14)
ID	4.5 (.75)	5 (1.56)
AO	6 (.72)	2 (2.30)
OR	2 (.77)	6 (1.46)
MC	1 (1.17)	3 (2.20)
TI	8.5 (.24)	9 (.41)
T1	7 (.51)	7 (1.23)
T2	8.5 (.24)	8 (1.19)

---

[1] Percent gain over ASVAB in correlation points is displayed in parenthesis

Table 30

Joint-Services: Test Weights for Methods of Estimating *g*.[1]

Test	ASVAB*g*	AFQT*g*	PSYCH*g*	ASVABECAT*g*
ASVAB Tests:				
GS	.161	–	.500	.091
AR	.159	1.00	.591	.096
WK	.158	–	.473	.086
PC	.149	–	.443	.081
NO	.100	–	.306	.057
CS	.094	–	.289	.055
AS	.117	–	.307	.066
MK	.141	1.00	.543	.086
MC	.151	–	.529	.092
EI	.139	–	.388	.077
VE	–	2.00	–	–
ECAT Tests:				
MC	–	–	.559	.082
SM	–	–	.544	.077
FR	–	–	.750	.088
ID	–	–	.591	.089
AO	–	–	.575	.087
SO	–	–	.544	.085
T1	–	–	-.345	-.062
T2	–	–	-.370	-.068
TI	–	–	-.291	-.054

---

1 For all methods except AFQT, the test scores were standardized to z scores based on the 1991 DoD applicant population.

Table 31

Joint-Services: Average Corrected Intercorrelations and Validities of g Score Composites for All Schools.

	r[1]	ASVABg	AFQTg	PSYCHg
ASVABg	.633	–		
AFQTg	.603	.938	–	
PSYCHg	.626	.923	.878	–
ASVABECATg	.629	.936	.882	.999

1 Average validity coefficient using primarily FSG criteria across all schools.

Table 32

Joint-Services: Combined Sample Regression Analyses for FSG and Performance Criteria.[1]

	Wherry-Corrected				Fully Corrected			
	FSG		Performance		FSG		Performance	
Order of Entry	Test	Average R	Test	Average R	Test	Average R	Test	Average R
1	AR	.631	AR	.441	AR	.677	AR	.509
2	WK	.672	CS	.474	WK	.721	CS	.547
3	MK	.697	sm	.497	MK	.747	sm	.572
4	ao	.711	AS	.517	ao	.762	AS	.594
5	CS	.722	MK	.528	CS	.774	MK	.607
6	AS	.732	t2	.537	AS	.785	t2	.617
7	PC	.736	id	.542	PC	.790	fr	.623
8	sm	.740	WK	.547	sm	.793	WK	.628
9	so	.742	mc	.549	so	.796	ao	.631
10	GS	.743	PC	.551	GS	.797	t1	.633
11	t1	.744	GS	.552	t1	.798	PC	.634
12	fr	.745	t1	.553	fr	.799	GS	.636
13	EI	.746	fr	.554	EI	.800	mc	.637
14	mc	.746	MC	.555	mc	.800	MC	.638
15	MC	.746	ao	.555	MC	.801	ti	.638
16	ti	.746	ti	.556	ti	.801	so	.638
17	t2	.747	so	.556	t2	.801	NO	.638
18	id	.747	NO	.555	id	.801	id	.638
19	NO	.746	EI	.554	NO	.801	EI	.637

[1] ECAT predictors are shown in lowercase.

Table 33

Joint-Services: Order of Entry, Tests, and Average Multiple Correlations for Combined Sample Regression Analyses.[1]

Order of Entry	Analysis One[2]		Analysis Two[3]		Analysis Three[4]	
	Test	Average R	Test	Average R	Test	Average R
1	AR	.517	AR	.517	AR	.543
2	GS	.558	GS	.555	GS	.583
3	MK	.574	MK	.570	MK	.598
4	AS	.593	AS	.588	AS	.617
5	CS	.604	CS	.598	CS	.628
6	t2	.611	t2	.605	t2	.635
7	PC	.617	PC	.609	PC	.640
8	ao	.621	ao	.613	ao	.644
9	EI	.625	EI	.616	EI	.647
10	fr	.628	fr	.618	fr	.649
11	WK	.630	WK	.619	WK	.650
12	so	.632	so	.621	so	.652
13	sm	.634	sm	.622	sm	.653
14	ti	.636	ti	.6224	ti	.654
15	NO	.638	NO	.6231	NO	.6543
16	mc	.639	t1	.6234	mc	.6546
17	t1	.640	mc	.6237	t1	.6550
18	MC	.641	MC	.6239	MC	.6552
19	id	.642	id	.6235	id	.6548

---

[1] ECAT predictors are shown in lowercase.
[2] Coefficients are corrected for predictor restriction of range.
[3] Coefficients are corrected for predictor restriction of range and shrinkage.
[4] Coefficients are corrected for range restriction, shrinkage, and criterion unreliability.

Table 34

Joint-Service: Schools Classified by Complexity Level

Complexity Level			
Low	Medium	High	Very High
11H (I & H)	BT/MM	AE	AC
19K	73230	AMS	OS
	13F	AO	27230
		AV	
		EN	
		EM	
		FC	
		GMG	
		RM	

Table 35

Joint-Service: Average and Incremental Validities by Level of School Complexity.

Complexity Level[1]	R_{ASVAB}	$R_{ASVAB + ECAT}$	ΔR	Percent Gain
Very High (3)	.7961	.8054	.0093	1.17%
High (9)	.7509	.7553	.0044	.59%
Medium (3)	.7006	.7227	.0221	3.15%
Low (3)	.2993	.3201	.0208	6.95%

---

[1] Number of schools in the complexity level appears in parentheses.

Table 36

Summary of Findings from Studies that Analyzed ECAT Tests.

Study	Jobs/Schools	Sample Size	ASVAB Base	Experimental Tests Added	Criteria	Range of Average Corrected Increments (in Correlation Points)	Best Performing Tests
McHenry et al., 1990	9 Army Jobs	4039	ASVAB Composites	Spatial Composite Psychomotor Perceptual Other Cognitive and Non-Cognitive Composites	Knowledge, Hands on, Ratings	1 - 4	Spatial Composite
Busciglio, 1990	As Above	4039	Significant ASVAB Tests	Spatial Composite Psychomotor Composite AO FR SO TI T1, T2	Knowledge, Hands on, Ratings	1.4 - 5.1 (median)	AO FR TI
Peterson et al., 1992	As Above	3759	AR and WK	Other ASVAB tests All ECAT except ID, MC, and SM	Knowledge, Hands on, Ratings	1.1 - 4.5	AO FR T1, T2 (especially for discriminant validity)
Carey, 1992	2 Marine Corp Jobs	698	All ASVAB Tests	All ECAT tests excepts MC, GATB Short term memory	Hands On	1.2 - 1.9	AO
Wolfe et al., 1992	9 Navy Schools	4989	All ASVAB Tests	MC, SM, FR, ID, Memory Composite Spatial Composite Perceptual Tests ASVAB Space	Knowledge, Hands on	1.6	Spatial Composite

---

[1] This table simplifies the results of the studies. The studies should be read to fully understand the findings presented above.