RELATIONSHIPS OF GENERAL ABILITY, SPECIFIC ABILITY, AND JOB CATEGORY FOR PREDICTING TRAINING PERFORMANCE

Malcolm James Ree
James A. Earles

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

September 1990

Approved for public release; distribution is unlimited.
NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Division

HAROLD G. JENSEN, Colonel, USAF
Commander
### ABSTRACT (Maximum 200 words)

The roles of general ability ($g$), specific abilities ($s$), and job category are investigated as predictors of final school grade for 78,041 Air Force enlistees in 82 jobs. General cognitive ability and specific abilities were defined by scores on the first and subsequent principal components of the enlistment selection and classification test, the Armed Services Vocational Aptitude Battery. A linear models analysis revealed that $g$ was a good predictor of all criterion measures and that $s$ added a small but statistically significant amount. Adjusting the final school grade criterion scale using job category variables significantly improved prediction.
SUMMARY

Many multiple-aptitude test batteries, including the Armed Services Vocational Aptitude Battery (ASVAB), used for assigning or classifying individuals to jobs or for occupational counseling have subtests covering a broad range of content such as science, mathematics, reading, vocabulary, clerical, mechanical, or technical knowledge. This content reflects a belief that performance in different jobs is best predicted by subtests whose content appears to be closely related to the jobs. It has been demonstrated that the subtests of a multiple-aptitude test battery all measure, at least in part, an examinee's general learning ability in addition to the specific abilities implied by the differing contents of the subtests.

The present effort investigated the utility of general learning ability and specific abilities for predicting final school grades in 82 Air Force technical training schools. Subjects were 78,041 Air Force enlistees. It was found that general learning ability was by far the best predictor for the training grades; however, specific abilities improved the predictive accuracy by a small amount.
This effort was conducted under work unit 77191867, Research on Air Force Selection and Classification. The authors would like to thank SSgt Steven Hoffer, AFHRL/SC, for his computer analyses. Gratitude is extended to Ms Jacobina Skinner for assistance with the formulation of the linear models and the counting of linearly independent vectors. For critical reading of the urtexts the authors thank Linda Sawin, Thomas Watson, and Lonnie D. Valentine, Jr., all of AFHRL/MA and Bill Phalen of AFHRL/MD. William Alley, AFHRL/MD, is owed a debt of gratitude for insightful discussions on the topic.

This effort was unique as it contained linear models with over 900 vectors and as such, was computationally cumbersome. The results were instructive and the authors believe Shakespeare would have written the following had he observed the results: "Lord, what tools these models be."
TABLE OF CONTENTS

I. INTRODUCTION .......................................................... 1

II. METHOD ............................................................................ 2
    Subjects ........................................................................ 2
    Measures ......................................................................... 2
    Procedures ....................................................................... 4

III. RESULTS AND DISCUSSION .............................................. 6

IV. CONCLUSIONS ............................................................... 7

REFERENCES ........................................................................ 9

LIST OF TABLES

Table | Page
--- | ---
1 | Educational and Demographic Description of the Sample | 2
2 | Subtests of the ASVAB | 3
3 | Unrotated Principal Component Weights for the ASVAB Subtests | 3
4 | Linear Models Used in the Statistical Tests | 5
5 | Correlations, Squared Multiple Correlations, and F Tests for the Models | 6
RELATIONSHIPS OF GENERAL ABILITY, SPECIFIC ABILITY, AND JOB CATEGORY FOR PREDICTING TRAINING PERFORMANCE

I. INTRODUCTION

The concept of general cognitive ability or psychometric $g$ first proposed by Galton in 1883 appeared in analyses early in this century (Spearman, 1904). It has been and remains the center of much controversy.

Early intelligence test developers such as Binet and Simon were proponents of $g$ but eventually the influence of multiple-ability theorists (Thurstone, 1938) was pervasive. This led to the development of multiple-aptitude batteries. The Differential Aptitude Tests (DAT), the General Aptitude Test Battery (GATB), and the Armed Services Vocational Aptitude Battery (ASVAB) were designed to measure specific abilities and to make specific predictions about employment or education. Sets of test scores are differentially selected or differentially weighted for each situation, fulfilling a proposal by Hull (1928). The different composites of subtests used by the military for job placement or the interpretation of score profiles in counseling are current examples of the application of multiple-ability theory. The use of differential weighting and different composites led to multiple-aptitude theory being termed "a theory of differential validity."

Recently the primacy of $g$ as a predictor has again become the subject of many studies. The December 1986 issue of Journal of Vocational Behavior (Gottfredson, 1986) provided much impetus for the renewed interest, as did the evidence emerging from validity generalization studies (Hunter, 1983, 1984a, 1984b, 1984c; Hunter, Crosson, & Friedman, 1985).

The ASVAB is an excellent data source for recent studies investigating the value of $g$ as a predictor, with over a million administrations and over 200,000 selections to job training each year. Jones (1988) correlated the average validity of the ASVAB subtests for predicting training performance with the $g$ saturation of the subtests. The validities were the correlations between a subject-weighted ($N =$ 24,482) average over 37 Air Force jobs, and the $g$ saturation was measured by the loadings on the unrotated first principal component (see Jensen, 1987). She found a rank-order correlation of .72, demonstrating a strong positive relationship between $g$ and validity.

Ree and Earles (1990a) investigated the predictive utility of both the general and specific components of the ASVAB by regressing Air Force technical school grades on the unrotated principal component scores of the ASVAB ($g$ was represented by the first principal component). Across 89 jobs (sample sizes ranged from 274 to 3,939 individuals), the average correlation of $g$ and the training criterion was .764 corrected for range restriction. When the specific components were added to the regressions, the increase in $R^2$ averaged .015.

The above two studies examined the relationships of general ability and specific ability, but neither statistically addressed differences in job category. A linear models analysis allows a statistical investigation of the relationships among these variables. It also allows investigation of the effects of using only the specific components ($s_1...s_n$) or only the $g$ component, as well as investigation of differences in the mean and dispersion of grade criteria of the individual training schools.
II. METHOD

Subjects

The subjects were 78,041 nonprior-service Air Force enlistees who had tested with ASVAB parallel Forms 11, 12, or 13 from 1984 through 1988 and who had completed both basic military training and technical training. They were mostly white, male, from 17 to about 23 years old and high school graduates. Table 1 shows the demographic characteristics of the sample. To ensure sufficient statistical power (Kraemer, 1983) for within-job regressions, no job with fewer than 274 cases was used. This is the same sample used earlier by Ree and Earles (1990a).

Table 1. Educational and Demographic Description of the Sample

<table>
<thead>
<tr>
<th>Gender</th>
<th>Proportion</th>
<th>Age</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>82.8</td>
<td>17-18</td>
<td>29.2</td>
</tr>
<tr>
<td>Female</td>
<td>17.2</td>
<td>19-20</td>
<td>37.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>Less than High School</td>
</tr>
<tr>
<td>Hispanic</td>
<td>High School Graduate</td>
</tr>
<tr>
<td>White</td>
<td>College Experience</td>
</tr>
<tr>
<td>Other</td>
<td>College Graduate</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>

Measures

The Armed Services Vocational Aptitude Battery is a multiple-aptitude test battery (DOD, 1984) composed of 10 subtests, as shown in Table 2. Except for the Numerical Operations and Coding Speed subtests, all are power tests. The ASVAB is used by all of the Armed Services for enlistment qualification and initial job assignment. It is normed on a weighted, nationally representative sample of 18- to 23-year-old youths (Maier & Sims, 1986; Wegner & Ree, 1985). The battery has been used in this current configuration since 1980, and is highly reliable (Palmer, Hartke, Ree, Welsh, & Valentine, 1988) and valid (Wilbourn, Valentine, & Ree, 1984).

There are three generally accepted ways of estimating the g component of a set of variables (Jensen, 1980). Ree and Earles (1990b) have shown for the ASVAB that estimates of g from these three methods—principal components, principal factors, and hierarchical factor analysis—all correlated greater than .996. Because of its mathematical simplicity and the high correlation among the various g estimates, the principal components methodology was chosen to represent g and specific measures of the ASVAB. Table 3 gives the 10 sets of ASVAB principal component score weights (Hotelling, 1933a, 1933b) derived previously (Ree & Earles, 1990a).

The method of principal components is a procedure for forming orthogonal linear composites of observed variables. Kendall, Stuart, and Ord (1983) noted that the uncorrelated principal component scores avoid problems of colinearity and are useful for regression analyses.
Table 2. Subtests of the ASVAB

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Number of Items</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Science (GS)</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>Arithmetic Reasoning (AR)</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>Word Knowledge (WK)</td>
<td>35</td>
<td>11</td>
</tr>
<tr>
<td>Paragraph Comprehension (PC)</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Numerical Operations (NO)</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Coding Speed (CS)</td>
<td>84</td>
<td>7</td>
</tr>
<tr>
<td>Auto and Shop Information (AS)</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>Mathematics Knowledge (MK)</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Mechanical Comprehension (MC)</td>
<td>25</td>
<td>19</td>
</tr>
<tr>
<td>Electronics Information (El)</td>
<td>20</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 3. Unrotated Principal Component Weights for the ASVAB Subtests

<table>
<thead>
<tr>
<th>Principal components</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS</td>
<td>.13808</td>
<td>-.11244</td>
<td>-.21982</td>
<td>-.29416</td>
<td>.19523</td>
<td>-.88893</td>
<td>-1.05107</td>
<td>.56764</td>
<td>.46367</td>
<td>-1.25618</td>
</tr>
<tr>
<td>AR</td>
<td>.13715</td>
<td>.03854</td>
<td>-.39912</td>
<td>.54694</td>
<td>-.02066</td>
<td>.26159</td>
<td>.58641</td>
<td>.25640</td>
<td>-1.51740</td>
<td>1.06178</td>
</tr>
<tr>
<td>WK</td>
<td>.13736</td>
<td>.06649</td>
<td>-.21381</td>
<td>-.64261</td>
<td>-.08976</td>
<td>.20343</td>
<td>-.35471</td>
<td>-.19392</td>
<td>-.122910</td>
<td>1.53259</td>
</tr>
<tr>
<td>PC</td>
<td>.12778</td>
<td>.16656</td>
<td>-.31273</td>
<td>-.71570</td>
<td>-.02359</td>
<td>.110958</td>
<td>.48914</td>
<td>-.18581</td>
<td>.83254</td>
<td>-.55741</td>
</tr>
<tr>
<td>NO</td>
<td>.11291</td>
<td>.38342</td>
<td>.42663</td>
<td>.23843</td>
<td>-.13676</td>
<td>-.11449</td>
<td>-.39672</td>
<td>-.29306</td>
<td>.20266</td>
<td>-.11527</td>
</tr>
<tr>
<td>CS</td>
<td>.09956</td>
<td>.44464</td>
<td>.75816</td>
<td>.03679</td>
<td>1.11560</td>
<td>.12448</td>
<td>-.30623</td>
<td>.21087</td>
<td>.39938</td>
<td>.36281</td>
</tr>
<tr>
<td>AS</td>
<td>.10878</td>
<td>-.43374</td>
<td>.60474</td>
<td>-.00918</td>
<td>-.34001</td>
<td>-.14894</td>
<td>.21734</td>
<td>.13184</td>
<td>-.06193</td>
<td>-.04099</td>
</tr>
<tr>
<td>MK</td>
<td>.12965</td>
<td>.12066</td>
<td>-.61466</td>
<td>.64452</td>
<td>.20353</td>
<td>.12857</td>
<td>-.29635</td>
<td>.14351</td>
<td>-.13640</td>
<td>-.00001</td>
</tr>
<tr>
<td>MC</td>
<td>.12448</td>
<td>-.30623</td>
<td>.21087</td>
<td>-.39938</td>
<td>.36281</td>
<td>.89768</td>
<td>-.119071</td>
<td>-.72807</td>
<td>-.02996</td>
<td>.28081</td>
</tr>
<tr>
<td>El</td>
<td>.12857</td>
<td>-.29635</td>
<td>.14351</td>
<td>-.13640</td>
<td>-.00001</td>
<td>-.78167</td>
<td>9.0823</td>
<td>.143032</td>
<td>.09391</td>
<td>-.06884</td>
</tr>
</tbody>
</table>

These 10 unrotated principal component scores were the measures of ability under investigation. The first unrotated principal component served as a measure of g (Jensen, 1980), and the other nine as the measures of specific ability ($s_1$ to $s_9$).

The criteria were the final school grades (FSG) received by the subjects in each of the technical training courses (see Ree & Earles, 1990a: Wilbourn, Valentine, & Ree, 1984).
most technical training schools, the FSG is the average of four fairly short multiple-choice
technical knowledge and procedures tests. However, to be eligible to take these tests, students
must first pass work sample tests (frequently called ‘performance checks’) In most technical
training schools, these performance checks may be repeated numerous times until the subject
succeeds. Although some subjects are removed from technical training for failure to pass
these performance checks, no easily accessible records of repeated performance check scores
exist. FSG is a complex criterion, reflecting more than printed tests

FSG is reported as a numerical grade from a lowest passing of 70 to a highest of 99, although some schools credit 60 as passing. The selectivity and difficulty of the school are not reflected in the criterion measures. Each criterion is scaled independently, the criterion means and standard deviations are in no orderly relation to one another. This tends to wrongly equate a grade of, say, 85 in each school. Without knowledge of differences in the criteria, proper interpretation is difficult and potentially misleading. For example, differences in the means of the criteria across schools may lead to differences in the regressions. In practice, this adds heterogeneity to the criterion and the accuracy of prediction is reduced.

**Procedures**

Seven linear models (Ward & Jennings, 1973) were constructed and tested to assess the
correlation of $g$ and $s_1...s_9$, and to adjust for job category in predicting the criteria. The
general linear model underlies all linear statistics, such as regression, analysis of variance,
canonical analysis, factor analysis, and discriminant function analysis. Testing of linear models
is based on establishing a ‘full model’ which contains as much information as is available
and then evaluating the loss of predictive accuracy resulting from the elimination of portions
of that information in simpler models, often called “restricted models.” The restricted models
must contain variables which are a subset of the variables in the full model. There is an $F$
statistic associated with the analysis (see Ward & Jennings, 1973, p. 67)

Table 4 describes the linear models and variables used in the present investigation. The
full model contained 902 linearly independent variables, allowing each job category to have its
own intercept and 10 slopes, one for each of the 10 principal component scores. There were
820 variables associated with the principal components (10 per job times 82 jobs) and 82 job
variables (the unit vector and 81 binaries—one for each job, less one to eliminate redundant
vectors).

Restricted model 1 (RM1) contained the unit vector, 81 job binaries, and 82 interaction
variables (JP1) to allow principal component 1 ($g$) to have its own relationship (slope) to FSG
for each job. No other aptitude information was included. There were 164 linearly independent
variables in this model.

Restricted model 2 (RM2) contained intercepts for each job (the unit vector, all but one of
the binaries for job membership) and the 10 principal component scores (P1 to P10), for a
total of 92 independent predictors. In this model the relationship (slope) between each aptitude
predictor and FSG was constrained to be the same for all jobs. That is, $g$ had a single
slope across all jobs; $s_1$ had a single slope across all jobs, as did $s_2$ and so on.

The third restricted model (RM3), with 83 predictors, had the unit vector, 81 job binaries
(JOBS), and the scores on principal component 1 ($g$). This model contained all the job
information but only one aptitude predictor, $g$, which was constrained to have the same slope
across all jobs.
Table 4  Linear Models Used in the Statistical Tests

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSG</td>
<td>Final School Grade, a continuous variable</td>
</tr>
<tr>
<td>JOBS</td>
<td>Categorical variable, 1 if in job, 0 otherwise</td>
</tr>
<tr>
<td>P1 to P10</td>
<td>Scores on principal components 1 to 10, continuous variables</td>
</tr>
<tr>
<td>JP1 to JP10</td>
<td>Product variables of categorical variable JOBS and P1 to P10</td>
</tr>
<tr>
<td>Unit Vector</td>
<td>A vector with every element equal to 1</td>
</tr>
</tbody>
</table>

The fourth restricted model (RM4) of 82 independent predictors had the unit vector and 81 job binaries. Prediction in this model was based only on knowledge of job category; no information on aptitude was used.

The fifth restricted model (RM5), with 11 linearly independent variables, contained the unit vector and the 10 principal component scores (P1 to P10). This model removed all job information from the full model, leaving only aptitude as predictors.

Restricted model 6 (RM6) contained only two linearly independent predictors, principal component 1 \((g)\) and the unit vector.

Each statistical test compared a restricted model to the full model to answer questions about the relationships of ability and job category to FSG. Some models controlled for criterion scaling and allowed estimation of criterion difference effects. All statistical tests were evaluated at the .01 Type I error rate.

While the linear models F test compares the difference between error sums of squares (or \(R^2\)s), the relative predictive efficiency of the models was evaluated using \(Rs\)  Brogden (1946)
presents a proof which demonstrates that predictive efficiency is linearly and directly related

to $R$ (or $r$ in the bivariate case). A correlation of .40 is half as efficient in prediction as a
correlation of .80. The predictive efficiency of restricted models will be expressed as a percent
of the predictive efficiency of the full model in all but one instance.

III. RESULTS AND DISCUSSION

Each of the seven linear models was determined to be significantly different from zero. Table 5 shows the correlations and their respective $F$ statistics.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R^2$</th>
<th>$df_1$</th>
<th>$df_2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>.62831</td>
<td>.39477</td>
<td>901</td>
<td>77139</td>
<td>55.77900</td>
</tr>
<tr>
<td>RM1</td>
<td>.60812</td>
<td>.36981</td>
<td>163</td>
<td>77877</td>
<td>280.37153</td>
</tr>
<tr>
<td>RM2</td>
<td>.60863</td>
<td>.37042</td>
<td>91</td>
<td>77949</td>
<td>503.98888</td>
</tr>
<tr>
<td>RM3</td>
<td>.60334</td>
<td>.36402</td>
<td>82</td>
<td>77958</td>
<td>544.17176</td>
</tr>
<tr>
<td>RM4</td>
<td>.46419</td>
<td>.21547</td>
<td>81</td>
<td>77959</td>
<td>264.34314</td>
</tr>
<tr>
<td>RM5</td>
<td>.42814</td>
<td>.18330</td>
<td>10</td>
<td>78030</td>
<td>1751.36070</td>
</tr>
<tr>
<td>RM6</td>
<td>.41803</td>
<td>.17475</td>
<td>1</td>
<td>78039</td>
<td>16525.06542</td>
</tr>
</tbody>
</table>

Each restricted model was tested against the full model, and all of their $R^2$s were found to be significantly lower. The standard error of estimate ($S_e$) of the full model was 5.09 and those for the restricted models were 5.16 or greater.

The test of RM1 against the full model yielded a statistically significant but small difference ($F = 4.31$, $df_1 = 731$, $df_2 = 77139$). There was an increase of .08 in the $S_e$. The $R$ difference of .02019 (.62831 - .60812 = .02019) represents a loss in relative predictive efficiency of 3% (.02019/.62831 = .03373). This restricted model contains all job information (the job binaries) and variables to allow $g$ to have a different slope for each job. No information on specific ability is included and the loss in relative predictive efficiency was very small.

The second restricted model (RM2) tested against the full model showed a statistically significant difference ($F = 3.83$, $df_1 = 810$, $df_2 = 77139$), but a relatively small $R$ difference of .01968 (a loss of 3% relative predictive efficiency) and a commensurately small increase in $S_e$ from 5.09 to 5.16. RM2 creates parallel regression lines for each principal component across jobs, but allows each of the 10 principal components to have its own slope. For example, the slope for principal component 1 is the same for each job and the slope for principal component 2 is the same for each job, but the slopes of principal components 1 and 2 are not necessarily the same. This model specifies that the predictive value of aptitude is the same for each job, differing only in level.

The test of RM3 versus the full model determines if constraining $g$ to a single slope across all jobs, removing specific abilities, and retaining job information (job binaries) is less predictive. This yielded a significant difference ($F = 4.78$, $df_1 = 819$, $df_2 = 77139$) in $R^2$. The $S_e$ increased by .10. There was a small loss in relative predictive efficiency, 4%, for giving up all the specific ability information and the individual slopes on $g$. 
The test of RM4 (which contains only the job binaries) against the full model allows estimation of the importance of aptitude. The statistical test found the models to be significantly different ($F = 27.86$, $df_1 = 820$, $df_2 = 77139$). The $R$ difference was .16412, with a loss of relative predictive efficiency of 26% and an $Se$ increase of .67. Giving up information on aptitude was costly.

Furthermore, RM4 has a predictive efficiency of 46% relative to perfect prediction. The criterion variable differed, at least to scale mean, for the jobs investigated. These differences added heterogeneity to the criterion. Interpretation of the regressions without consideration of the differences in the criterion would be misleading.

The next linear models test, comparing RM5 to the full model, allows each principal component to have a separate slope; but any given principal component has the same slope across all jobs. RMF eliminates all information on jobs and retains only constrained information on the aptitude predictors. There was a statistically significant difference ($F = 30.25$, $df_1 = 891$, $df_2 = 77139$) between the models. The $R$ difference was .20017, a large decrease with a loss in predictive efficiency of 32%. Additionally, the increase in $Se$ was from 5.09 to 5.88 from the full to the restricted model. Clearly, giving up information about jobs was costly.

The test of RM6 against the full model asks if $g$ is as good a predictor as all the information in the full model. The $F$ statistic ($F = 31.15$, $df_1 = 900$, $df_2 = 77139$) showed a significant difference between the models. There was a .21028 decrease in the $R$ (a 33% loss in relative predictive efficiency) and a substantial increase of .82 ($5.91 - 5.09$) in the $Se$.

Finally, a test of the difference between RM1 and RM3 was conducted to determine the statistical difference and the relative predictive efficiency loss for using a single slope for $g$ as opposed to using a different slope for each job. It was statistically significant ($F = 8.83$, $df_1 = 81$, $df_2 = 77877$). There was an $R$ difference of .00478 and a .02 $Se$ increase. Although there was not the same relationship between $g$ and each job, the differences among these relationships were small. The loss in relative predictive efficiency was less than 1%.

**IV. CONCLUSIONS**

These analyses disclosed several interesting results. First, all the principal components were useful in predicting the criteria, much as was found earlier (Ree & Earles, 1990a). In the ASVAB, the specific abilities, as represented by the second through tenth principal components, added to the accuracy of prediction, but only a small amount.

The results also indicated that interpretation of regressions without reference to job categories could be very misleading. It should be noted that a little more than two-thirds of the predictive efficiency of the full model came from the knowledge of job category.

Controlling job category information, the role of specific aptitude information accounted for an increase of about 1 (RM2 versus RM3) to 3 (full model versus RM1) percent in relative predictive efficiency. When only job categories and $g$ were used, the consequence of using a single slope for $g$ instead of a separate slope for $g$ for each job (RM1 versus RM3) was very small.

In the statistical process of aggregating correlations to describe relationships across jobs, accounting for job category was important. General cognitive ability ($g$) has been shown to be a good predictor for all of the jobs. The addition of measures of specific ability increased predictive efficiency but only a small amount. Personnel selection systems with few applicants would not be likely to detect the statistical effects of such small contributions of specific abilities.
Finally, consideration of the need for job-specific slopes for $g$ indicated that the use of a common slope, though not optimum, is not very costly.

The subtests of the ASVAB all contribute about equally to the measurement of $g$, but their non-$g$ contributions are small and unequal (Ree & Earles, 1990b). The present study demonstrates the utility of $g$ and the relative lack of utility of $s_1$...$s_9$ as predictors of FSG. It follows that any attempt to aggregate the ASVAB subtests into equally reliable composites for the purpose of creating differential prediction of FSG can yield a small amount of improvement. Additional cognitive or non-cognitive measures outside of the ASVAB will be required to produce such differential prediction.

Because FSG is not the only criterion of importance, this study should be replicated using job performance measures, supervisory ratings, and other criteria.
REFERENCES


