AD

U. S. NAVAL PERSONNEL RESEARCH ACTIVITY

SAN DIEGO, CALIFORNIA 92152

RESEARCH REPORT SRR 69-6 **SEPTEMBER 1968**

SELECTION OF MILITARY PERSONNEL FOR FOREIGN LANGUAGE TRAINING

Idell Neumann
Norman M. Abrahams
William H. Githens

NPRA
SAN DIEGO CALIFORNIA

AN ACTIVITY OF THE BUREAU OF NAVAL PERSONNEL

NOTE: The contents of this publication do not necessarily represent the official position or policy of the Department of the Navy.

AD

SELECTION OF MILITARY PERSONNEL FOR FOREIGN LANGUAGE TRAINING

Idell Neumann
Norman M. Abrahams
William H. Githens

September 1968

PF0160501B01
Research Report SRR 69-6

Submitted by

B. Rimland, Ph.D., Director, Personnel Measurement Research Department

Approved by

E. E. Dudek, Ph.D., Technical Director
G. W. Watson, Commander, USN
Commanding Officer

U. S. Naval Personnel Research Activity
San Diego, California 92152

SUMMARY

A. Problem

Because of the need for military personnel competent in foreign language skills, a program of research has been initiated for the Defense Language Institute (DLI) to develop tests and other procedures for improving selection of language trainees capable of high levels of language achievement.

B. Background

Currently selection of students is primarily based on the Foreign Language Aptitude Test (FLAT). Prior research, however, has demonstrated the importance of including both motivation and aptitude tests in predicting foreign language achievement. Since present DLI selection procedures do not include systematic measurement of trainee motivation, it was important that non-cognitive measures be considered for selecting foreign language students.

C. Approach

In addition to obtaining FLAT scores, several measures such as the Personal Data Questionnaire, the Navy Adjective List, and Instructor Ratings were gathered experimentally at the Defense Language Institute West Coast for validation as predictors of final class standing. Keys were empirically developed for the experimental tests in part of the sample and validated on the remainder. Multiple regression techniques were used to determine the best combinations of predictors. Where data were available, important findings were replicated on a small sample from the Defense Language Institute East Coast.

D. Findings, Conclusions, and Recommendations

The major finding of this research is that prediction of language achievement can be markedly improved by an instructor's rating obtained at the end of only one week of instruction--or even one day if need be. (Pages 5, 7). If "trial training" were implemented for the purpose of obtaining Instructor's Ratings prior to the inception of formal training, considerable expense may be avoided by eliminating those students considered substandard by the instructors.

If brief trial training proves to be infeasible and an ample number of potential trainees are available, some improvement may also be achieved if selection were based on paper and pencil tests, i.e., a combination of the Personal Data Questionnaire and Foreign Language Aptitude Test scores. (Page 10)

If either or both of the two procedures (i.e., paper and pencil tests and instructor's rating) are adopted for operational use, follow-up research is recommended on a larger sample of Navy personnel to improve the accuracy of the weights and cutting scores used, since the DLI included members of all the military forces.

REPORT USE AND EVALUATION

Feedback from consumers is a vital element in improving products so that they better respond to specific needs. To assist the Chief of Naval Personnel in future planning, it is requested that the use and evaluation form on the reverse of this page be completed and returned. The page is preaddressed and franked; fold in thirds, seal with tape, and mail.

Department of the Navy

Postage and Fees Paid
Navy Department

Official Business

Chief of Naval Personnel (Pers-A3)
Department of the Navy
Washington, D. C. 20370

Report Title & No: SELECTION OF MILITARY PERSONNEL FOR FOREIGN LANGUAGE TRAINING (Research Report SRR 69-6)

1. Evaluation of Report. Please check appropriate column.

FACTORS	RATING LOW	AVE	HIGH	COMMENTS
Usefulness of Data				
Timeliness				
Completeness				
Technical Accuracy				
Validity of Recommendations				
Soundness of Approach				
Presentation and Style				
Other				

2. Use of Report. Please fill in answers as appropriate.

a. What are your main uses for the material contained in the report?

b. What changes would you recommend in report format to make it more useful?

c. What types of research would be most useful to you for the Chief of Naval Personnel to conduct?

d. Do you wish to remain on our distribution list?

e. Please make any general comments you feel would be helpful to us in planning our research program.

NAME: _____ CODE: _____

ORGANIZATION: _____

ADDRESS: _____

CONTENTS

TABLES

SELECTION OF MILITARY PERSONNEL FOR FOREIGN LANGUAGE TRAINING

A. BACKGROUND AND PURPOSE

The Foreign Language Aptitude Test (FLAT) is a selection test having moderate to high validity for predicting success in foreign language training. The test requires the examinee to learn the vocabulary of an artificial language and certain grammatical principles of the artificial language, all of which are applied in the translation of sentences. The FLAT was instituted in January 1963, for selecting recruits from the naval training centers for foreign language training at the Defense Language Institute (DLI).

Even though prior research had indicated the FLAT alone to be useful as a selection instrument, a review of relevant research studies demonstrated the importance of including motivation as well as aptitude measures for predicting foreign language student achievement. Consequently, an extensive experimental test battery was given to students at the DLI, both the West Coast (DLIWC) and the East Coast (DLIEC) Branches, to determine the effectiveness of various cognitive and noncognitive measures in predicting success in foreign language training.[1] Comparisons were made to determine if selection using FLAT alone could be improved upon through the addition of (1) an instructor's rating obtained at the end of the first day and/or first week of the course, (2) interest and motivational questionnaires empirically keyed to predict foreign language achievement, and (3) other experimental and biographical indices such as: Pay Grade, Age, Education Level, and Vocabulary Learning Test scores. A detailed description of the procedures used in the analysis of the DLIWC data and the results obtained were presented in an earlier technical report (Neumann, Abrahams & Githens, 1968). The present report provides the DLIWC findings in a less technical manner and presents the results of a replication of the relevant findings on the DLIEC sample.

B. POPULATION

The primary population studied consisted of 660 men enrolled in a wide variety of language classes at the DLIWC, located at Monterey, California. Due to the relatively small sizes of the individual classes, the small proportion of naval personnel in attendance, and the need for sizeable groups for statistical analysis, the sample studied also included Army, Air Force, and Marine Corps students. Army personnel made up the largest

---

[1]Dr. Bob D. Rhea served as Project Director during the early stages of the study.

portion of the sample--472 men, or 71.5 per cent of the total. Navy men constituted 13.9 per cent of the total; the Air Force contributed 11.8 per cent, and the remaining 2.8 per cent came from the Marine Corps. The majority of the students, 88.8 per cent, were from the enlisted ranks, and the remainder were officers. Age ranged from 17 to 51 years, with a median of 20. The amount of formal education ranged from less than high school graduation to the completion of Master's degree requirements, with a median of two years of college. In addition to this DLIWC group, a smaller sample (N=129) was obtained from the DLIEC, located at Washington, D. C. for replication of relevant DLIWC findings.

C. CRITERION

Final Class Standing (FCS), adjusted for class size, was used as the criterion of foreign language achievement. Adjusting for class size made it possible to combine students from different classes and languages into "language groups" on a common scale to reflect each student's relative classroom achievement.

D. PREDICTORS

In addition to FLAT, the operational selection test, experimental predictors were assembled from other tests administered after selection but prior to language training, ratings secured from instructors after the first week of class, and information available from DLI records. Experimental predictors were classified into three types. Test scores based on existing scales and background information were labeled Set I predictors. The second category, Set II, consisted of specially constructed empirical keys. Instructor ratings made up the third category, Set III.

1. Operational Predictor: Foreign Language Aptitude Test (FLAT)

This test was originally called the Army Language Aptitude Test (ALAT) when it was developed by Dorcus, Mount, and Jones in 1952. Later Army studies determined the ALAT to be of operational utility (Berkhouse, Mendelson & Kehr, 1959). The FLAT is currently used as an aptitude screening test for selection to DLI.

2. Experimental Predictors

a. Set I. The Set I predictors consist of scores based on existing tests and background information records. These include the following measures:

(1) Insolence Scale. This test is assumed to be a measure of passive-aggressive personality structure and has been found to be related to the job performance of Navy third class enlisted men (Kipnis, 1965). Two subscales have been developed and scores on both were obtained to determine their effectiveness in predicting foreign language achievement.

(2) Hand Skills Test (HST). The HST is designed to measure motivation by testing the persistence of subjects in doing a simple and monotonous tally-marking task several hundred times in a timed situation (Kipnis & Glickman, 1958). Scores are derived by subtracting the practice trial score from the final trial score.

(3) Education Level. The amount of formal education completed prior to admission to the DLI was employed as a continuous scale, measured in years. It ranged from less than high school graduation to the completion of graduate college degrees, with a mean education level of two years of college.

(4) Pay Grade. Pay Grade is a code that is uniformly used by each of the services to reflect salary level. All grades were recoded to form a continuous measure. Grades E1 through E9 were coded 1 through 9, respectively, grades W1 through W4 as 10 through 13, respectively, and officer grades 01 through 06 as 14 through 19, respectively.

(5) Age. The school input age ranged from 17 to 51 years, and was used as a continuous variable.

(6) Vocabulary Learning Test. This test consists of 20 unusual English words and their definitions. It was the first test presented in the test battery at the DLI. Ten minutes were given for the students to learn the list. At the end of the test battery, only the definitions were presented and the students were given five minutes to supply the appropriate word from memory. Three experimental scores were derived as a means of measuring the ability to recall and match newly presented words to their meanings:

(a) The number of accurately spelled words recalled.

(b) The number of words for which at least the first two letters were correct.

(c) The number of words attempted.

b. Set II. This predictor set consisted of two questionnaires which were empirically keyed to predict foreign language achievement. The primary criterion of foreign language achievement was FCS. Two other less relevant criteria of achievement were available--the Listening Comprehension (L/C) and Reading Comprehension (R/C) scores of the Army Language Proficiency Test (ALPT). The construction of the empirical keys to predict these criteria was discussed in the technical report by Neumann, et al. (1968).

(1) Personal Data Questionnaire (PDQ). The 190 items of the PDQ include biographical, need for change, acceptance of social change, and study habit subtests which have been related to foreign language achievement in other studies (Altus, 1961; Hebenstreit, 1959; Heilbrun, 1962; Lambert, undated; Levy, 1962; Maier, 1959; Pimsleur, 1962; Preston, 1961).

(2) Navy Adjective List (NAL). "Need for achievement" has in past research been found to be related to school success (Barnette, 1961; Bendig, 1957, 1958; Hebenstreit, 1959). The NAL consists of 103 adjectives, many of which appear relevant to the "need for achievement" concept.

c. Set III--Instructor's first week ratings. To determine whether instructors can accurately predict students' ultimate achievement in language school from performance in early stages of instruction, each language instructor at DLIWC rated his students at the end of the first week of class. In the second study, conducted at DLIEC, first day ratings were also obtained, in addition to the first week ratings. Ratings were secured on seven-point scales to estimate a student's probable degree of language "success," his quality of "oral production," and his "motivation" to complete language training. The "oral production" rating was obtained in response to the suggestions of language school personnel that willingness to vocalize in the target language and the student's correctness of pronunciation may be related to skill in acquiring language facility. If ratings obtained early in training were valuable in predicting achievement, attempts may be made to secure such ratings at designated centers to be used in the selection of language school students prior to transfer to the DLI.

E. PROCEDURE

1. Language Groups

Insufficient sample size for any single language necessitated combining classes into language groupings. The DLI suggested six language groups based on language structure and grammar, comparable difficulty in acquiring a vocabulary, length of time required to achieve desired proficiency with the language, and ability required to discriminate tonal changes. Sufficient data were available to analyze the following three of the six recommended categories:

a. Indo-European (Western): Albanian, French, German, Greek, Italian, Portuguese, Romanian, and Spanish.

b. Indo-European (Eastern): Bulgarian, Czechoslovakian, Hungarian, Persian, Polish, Russian, and Serbo-Croatian.

c. Indo-Chinese: Burmese, Chinese, Japanese, Korean, Malayan, Thai, and Vietnamese.

2. Criterion Scores

Prior to statistical analyses of the data, it was necessary to make the criterion scores comparable for the various language classes combined into each language group. By assuming that an individual's class standing is not influenced by course length, that each student's interest and aptitude

were not dependent upon which language he was studying, and that the distribution of student aptitude was equal between language classes within a group, it was possible to include all classes regardless of course length or content in each language group for the prediction of FCS.

3. Statistical Analyses

a. Key construction. Special keys were built for the PDQ and the NAL to predict FCS. Two-thirds of the students were used for key construction, and the remaining randomly selected one-third were held out for cross-validation. These samples were used to construct and validate separate keys for each language group, and also to construct and validate one general key for the three combined groups.

b. Multiple regression. Statistical procedures were used to evaluate increases in validity as the number of predictors was increased. Multiple-regression procedures were used to derive four equations for four combinations of predictor sets within each language group. These equations were developed to predict FCS from: (1) FLAT and Set I, (2) FLAT, Set I, and Ratings, (3) FLAT, Set I, and Set II, and (4) FLAT, Set I, Set II, and Ratings. In addition, general equations were computed for the three language groups combined.

F. RESULTS AND DISCUSSION

1. DLIWC Results

The DLIWC data were analyzed through multiple regression to assess possible increases in validity for four combinations of predictor sets. The individual predictor validities are presented in Appendix Table A for the key construction and cross-validation samples. The complete inter-correlation matrices for each of the samples can be found in the technical report on the DLIWC data analysis (Neumann, et al., 1968).

The equations resulting from multiple-regression analyses are presented in Appendix Tables B and C. They were evaluated for predictive efficiency and the best equations selected for each language, one including instructor ratings and one excluding ratings. The validities of the selected equations are presented in Table 1, along with the validities of FLAT alone. In all instances, prediction of foreign language achievement can be improved by applying the appropriate composite rather than just using FLAT alone.

For the Indo-European languages, both Western and Eastern, the use of an instructor's rating combined with FLAT (equations 1 and 3) shows the largest gain in validity, 43 and 18 correlation points, when compared with that of FLAT alone in the respective language groups. However, should it not be administratively feasible to obtain first week instructor's ratings, selection of more successful Indo-European Western and Eastern language students can still be effected by the use of equations 2 and 4. For the

TABLE 1

Recommended Multiple-Regression Equations for the Prediction of Final Class Standing

Language Group	Regression Weights and Variables in Composite Equation	Validities on Cross-validation Samples	
		Composite	FLAT
Indo-European (Western)			
1. With Ratings	-1.165(1st Week "Oral Rating") -0.714(FLAT)	.70	.27
2. Without Ratings	-4.069(Education Level) -0.290(PDQ) - 0.271(NAL)	.56[a]	.42
Indo-European (Eastern)			
3. With Ratings	-1.165(1st Week "Oral Rating") - 0.714(FLAT)	.55	.37
4. Without Ratings	-1.121(FLAT) - 1.284(Vocabulary Learning Test)	.42	.37
Indo-Chinese			
5. With Ratings	-1.165(1st Week "Oral Rating") - 0.714(FLAT)	.49	.45
6. Without Ratings	-2.119(Pay Grade) - 1.359(FLAT)	.54	.45

Note --

[a]Regression weights and variables were determined on the smaller sample and cross-validated on the larger sample, due to constraints in the use of the data.

Western languages, a composite score for selection purposes may be based on a weighted combination of Education Level and scores on the empirically derived PDQ and NAL scales. Similarly, combining a Vocabulary Learning Test score with FLAT shows an increase in validity over that possible with FLAT alone for the Eastern language group.

For the Indo-Chinese languages, Pay Grade in combination with FLAT provides the greatest improvement with the predictors investigated in this study, i.e., a gain of 9 correlation points over the validity of FLAT alone. However, to eliminate potential students on the basis of Pay Grade may conflict with more essential needs of the service and prove to be impractical. In this case, provided that ratings are introduced as an operational selector for language school, then equation 5, identical to equations 1 and 3, would raise the validity from .45 for FLAT alone to .49 for the composite.

In order to illustrate the practical effects of the various predictors in terms of student achievement, separate analyses were directed toward identifying students who graduated in the upper half of their classes. Using a variety of possible selection cut-offs on FLAT and the composite predictors, i.e., those scoring in the upper 20 per cent, upper 40 per cent, upper 60 per cent, and upper 80 per cent, the percentage of "top half" students was computed (see Table 2). For example, with the Western language group, if selection were limited to the upper 20 per cent with respect to the composite predictor, which includes instructor ratings, 29 per cent more "top half" graduates could be expected when the composite is used than when FLAT is used as a single predictor. For the Western and Eastern languages, using any of the four cut-offs, the composite predictor is as good or better than FLAT alone. For the Indo-Chinese languages, however, some comparisons do not favor the composite predictors over the use of FLAT alone, as indicated by negative increments.

2. Replication of Relevant DLIWC Findings on DLIEC Data

Due to the relatively small input to DLIEC, data collection was limited to a sample of 129 subjects. When categorized by languages, it became apparent that only the Indo-European (Western) language group (N=75) was sufficiently large to permit meaningful analysis. Validities for predicting foreign language achievement, as measured by FCS, were computed for each available predictor and are presented in Appendix Table A.

In addition to the first week ratings, each student was rated by an instructor on each of the three scales ("success," "oral production," and "motivation"), after having been observed for one full day. A comparison is made in Table 3 between the validity of first day and first week ratings since, for selection purposes, a one day rather than a one week rating would be preferred. Although the comparison between first day and first week ratings is available only on this relatively small sample, the results are favorable to replacing first week with first day ratings. Even though the validity of the "success" rating scale is lowered from .71 based on a first week estimate, to .51 for the first day rating, and from .67 to .35 for the "oral production" rating, respectively, these first day validities

TABLE 2

Percentage Expected in Top Half of Class for
Various Language Groups

Cumulative Predictor Categories	Western (1)[a] FLAT	Western (1)[a] Composite (Rating)	Western (1)[a] Increment	Western (2) FLAT	Western (2) Composite (No Rating)	Western (2) Increment	Eastern (3) FLAT	Eastern (3) Composite (Rating)	Eastern (3) Increment	Eastern (4) FLAT	Eastern (4) Composite (No Rating)	Eastern (4) Increment	Chinese (5) FLAT	Chinese (5) Composite (Rating)	Chinese (5) Increment	Chinese (6) FLAT	Chinese (6) Composite (No Rating)	Chinese (6) Increment
Upper 20%	58	87	29[b]	85	85	0	83	85	2	83	93	10	75	83	8	75	77	2
Upper 40%	60	82	22	66	94	28	68	83	15	68	74	6	74	63	-11	74	67	- 7
Upper 60%	57	73	16	62	65	3	59	72	13	59	65	6	60	59	- 1	60	67	7
Upper 80%	54	61	7	57	57	0	55	63	8	55	60	5	55	57	2	55	61	6
Base Rate (No Selectors)	50	50	--	50	50	--	50	50	--	50	50	--	50	50	--	50	50	--

Notes --

[a]Number in parentheses refers to the equation number as presented in Table 1.

[b]The underlined numbers are increments, indicating the difference between results if the FLAT were used alone and if various composites were used.

TABLE 3

Comparison of the Validities Obtained Through the Combined Use of FLAT With a Rating Scale Versus Using FLAT Only
(N=75)

Composite (FLAT & Rating)	Composite Validity	Separate Validity		Increase of Composite Over FLAT Alone
		FLAT	Rating	
FLAT & 1st Day "success"	.55	.29	.51	.26**
FLAT & 1st Day "oral"	.38	.29	.35	.09*
FLAT & 1st Week "success"	.71	.29	.71	.42**
FLAT & 1st Week "oral"	.68	.29	.67	.39**

Notes --

*Significant beyond the .05 level.
**Significant beyond the .01 level.

Single variable coefficients are Pearson r's and are presented as positive for purposes of comparing with multiple R's.

still exceed the correlation of .29 between FLAT and the FCS criterion. If the instructors had designed their initial lessons to facilitate student selection, presumably even higher validities could be obtained.

For purposes of indicating any increase in prediction possible through the combined use of ratings with FLAT, multiple correlations were computed for each of the potentially useful ratings obtained at DLIEC. These composite validities are presented with FLAT validities for comparison in Table 3. Varying increments in validity are indicated in Table 3, depending upon which rating scale is being considered, with three of the four increases being significant beyond the one per cent level. For both first day and first week ratings, the largest contribution to validity is from the "success" rating, which is based on an instructor's estimate of each student's probable degree of language success.

In addition to validating first day instructor's ratings, the relevant DLIWC findings were replicated on the DLIEC data. The recommended predictor composites resulting from the DLIWC data are presented in equations 1 and 2 of Table 4. For the first equation, a validity of .62 was obtained on the DLIEC sample, a considerable improvement over the validity obtained using FLAT alone. Due to the unavailability of the Education Level information and NAL scores at DLIEC, it was not possible to cross-validate the second equation. However, in an attempt to provide an alternate selection procedure if ratings cannot be used, weights were determined on the one-third DLIWC sample for the combined use of FLAT and the PDQ and cross-validated on both the two-thirds DLIWC sample and the DLIEC sample. Applying regression equation 3 resulted in a significant increase for both samples in the Indo-European (Western) languages over the use of FLAT alone.

TABLE 4

Cross-validation on Data From DLIEC and DLIWC of Weights Determined on a Portion of DLIWC Data for The Prediction of Final Class Standing

Regression Weights and Variables in Equation	Validities			
	DLIWC		DLIEC	
	Composite	FLAT	Composite	FLAT
With Rating				
1. -1.165(1st Week "Oral" Rating) -0.714(FLAT)	.70[a]	.27[a]	.66[c]	.29[c]
Without Rating				
2. -4.069(Education Level)-0.290 (PDQ)-0.271(NAL)	.56[b]	.42[b]	Not available	
3. -0.393(PDQ)-0.655(FLAT)	.67[b]	.42[b]	.53[c]	.29[c]

Notes --

[a]Based on one-third sample (N=66).

[b]Based on two-thirds sample (N=139).

[c]Based on total sample (N=75).

Table 5 is an expectancy chart analogous to Table 2, designed to permit comparison between the percentages of expected "top half" students if selected on the basis of two composite scores or on the FLAT score alone. As an example, if selection were limited to the upper 20 per cent with respect to the predictor, 14, 13, and 9 per cent additional above-average students could be expected, depending upon which composite is used as opposed to using FLAT alone.

3. Recommendations for Selection

Foreign language achievement in an intensive language training course can be predicted with greater accuracy than is presently possible using FLAT alone. The following recommended selection procedures are optimal for samples similar to the ones analyzed in the present study, i.e., heterogeneous samples composed of members from the various branches of service:

a. If instructor ratings are obtained for use in selecting students, the recommended equations differ, depending on whether first week or first day ratings are used.

(1) The FLAT score combined with an instructor's first week rating results in improved prediction of foreign language achievement for the Indo-European Western and Eastern groups over that provided by the FLAT alone. Only slight improvement was found for the Indo-Chinese languages with this multiple and, therefore, it is recommended for use only when a uniform selection procedure seems advantageous or if more essential needs of the services conflict with the specific recommendation outlined in b(3) below.

(2) Results obtained on a small sample of DLIEC students on one language group indicate that a composite of first day ratings and FLAT is not as effective as that obtained with first week rating and FLAT. However, a significant increase in validity is possible over that obtained with FLAT alone. Since it would be more feasible for selection use to obtain first day rather than first week ratings, further research is recommended to assess the validity of first day ratings for a much larger sample composed of all language groups.

b. If it is not economical to obtain and use instructor ratings for selection, alternative equations are presented:

(1) For the Indo-European (Western) languages, findings based on samples from both the DLIWC and DLIEC schools suggest the use of a weighted combination of FLAT and the PDQ.

(2) For the Indo-European (Eastern) group, a weighted combination of the FLAT and the Vocabulary Learning Test is recommended.

(3) For the Indo-Chinese group, a weighted combination of Pay Grade and FLAT improves the prediction of success in foreign language training. This combination is recommended for use provided that selecting only men from the higher pay grades does not conflict with more essential needs of the services.

TABLE 5

Percentage Expected in Top Half of Class for
Western Language Group

Cumulative Predictor Categories	DLIEC					DLIWC		
	FLAT	Composite (Rating)	Increment (Over FLAT)	Composite (No Rating)	Increment (Over FLAT)	FLAT	Composite (No Rating)	Increment (Over FLAT)
Upper 20%	69	83	14	82	13	86	95	9
Upper 40%	60	70	10	67	7	65	79	14
Upper 60%	62	66	4	68	6	60	69	9
Upper 80%	57	62	5	55	- 2	55	60	5
Base Rate (No Selectors)	50	50	--	50	--	50	50	--

Note --

The underlined numbers are increments, indicating the difference between results if the FLAT were used alone and if various composites were used.

c. Since the differential prediction resulted from the grouping of languages, it is expected that similar results would be found for individual languages. The differential prediction of language achievement for a specific language was originally planned, but operational restrictions at DLIWC did not permit time for sufficient students to be tested for this research. Thus, it is felt that prediction could be improved even further if sufficient data were available for analyses of the more widely studied individual languages, such as French, German, Russian, or Chinese.

4. Limitations

Since the personnel needs of the various branches of service differ, FLAT had not been applied in a uniform manner for selection of students in classes used in the present study. Consequently, the resulting regression equations probably do not predict equally well for all service branches. They do, however, demonstrate the magnitude of increased validity possible with the additional predictors. The optimal equations for each service could, of course, be constructed only from sizeable samples from each service. Since this was not possible with the existing data, the prediction equations represent a necessary compromise, and again, indicate potential gains in predictive efficiency.

It should also be noted that the instructor ratings are probably underestimated as to validity, since they were not gathered with the express purpose of facilitating student selection.

G. CONCLUSIONS AND RECOMMENDATIONS

One of the major findings of this research is that the foreign language achievement of military trainees may be predicted with substantial accuracy using the predictors examined in this study. If selection were based on only paper and pencil tests such as the Foreign Language Aptitude Test (FLAT), or the Personal Data Questionnaire (PDQ), improvement may be achieved with less than two or three hours of testing. Thus, if ample potential trainees are available, only the most promising of a group of potential trainees need be selected for the Defense Language Institute.

Another major finding of this research is that trainee language proficiency at the end of the course of instruction can be fairly readily predicted by an instructor at the end of only one week of instruction--or even one day if need be.

The improvement in selection obtained by paper and pencil tests can itself be improved upon by using both tests and ratings. This procedure is recommended. It seems advisable to develop means for permitting language instructors to screen potential trainees through an intensive period of language training of perhaps two days duration. This could be accomplished through instructor travel to training centers, or by sending the trainees to the language school for a short trial training session. It is believed that the cost of such travel could be offset by the savings in training

expenses if these selection procedures were instituted. It seems likely that screening out the potential trainees with the lowest probability of success would permit the remainder of the group to complete training at markedly lower cost in time and dollars per graduate with considerably greater language proficiency on graduation as a bonus. Investigation of the operational feasibility of this procedure is recommended.

The selection of more promising foreign language trainees is possible if either of the two previously outlined procedures (i.e., paper and pencil tests alone or combined with an instructor's rating) are established for operational use. Either procedure will require further research on a sufficiently large sample consisting of Navy personnel only to establish exact weights and cutting scores to be used.

REFERENCES

Altus, W. D. Correlation data for first-semester grade averages at the University of California, Santa Barbara. J. genet. Psychol., 1961, 98, 303-305.

Barnette, W. L., Jr. A structured and a semi-structured achievement measure applied to a college sample. Educ. psychol. Measmt., 1961, 21, 647-656.

Bendig, A. W. The validity of two temperament scales in predicting student achievement in introductory psychology. J. educ. Res., 1957, 50, 571-580.

Bendig, A. W. Comparison of the validity of two temperament scales in predicting college achievement. J. educ. Res., 1958, 51, 605-609.

Berkhouse, R. G., Mendelson, M. A., & Kehr, C. Determining the minimum qualifying score on the Army Language Aptitude Test ALAT-1. Washington: Personnel Research Branch, Dept. of the Army, July 1959. (Technical Research Study 59-6)

Dorcus, R. M., Mount, G. E., & Jones, M. A. Construction and validation of foreign language aptitude tests. Washington: Personnel Research Branch, Department of the Army, June 1952. (Technical Research Report PRB-992)

Hebenstreit, W. J. The prediction of college success from biographical data and self-ratings. Univ. of Wash., Div. of Couns. and Test. Serv., 1959.

Heilbrun, A. B., Jr. Prediction of first year college drop-outs, using ACL need scales. J. counsel. Psychol., 1962, 9, 58-63.

Kipnis, D. The relationship between persistence, insolence and performance as a function of general ability. Educ. psychol. Measmt., 1965, 25, 95-103.

Kipnis, D., & Glickman, A. S. The development of a non-cognitive battery to predict enlisted performance. Washington: U. S. Naval Personnel Research Field Activity, August 1958. (Technical Bulletin 58-9)

Lambert, W. E., Gardner, R. C., Olton, R., & Tunstall, K. A study of the roles of attitudes and motivation in second-language learning. McGill University, undated, unpublished mimeo.

Levy, L. W. Age and personal need correlates of expectancy for change. Percep. mot. skills, 1962, 15, 351-356.

Maier, M. Follow-up on the validity of a forced-choice study activity questionnaire in another setting. J. appl. Psychol., 1959, 43, 293-295.

REFERENCES

Neumann, Idell, Abrahams, N. M., & Githens, W. H. Predicting foreign language achievement of officer and enlisted personnel. San Diego: Naval Personnel Research Activity, January 1968. (Technical Bulletin STB 68-8)

Pimsleur, P., Mosberg, L., & Morrison, A. L. Student factors in foreign language learning. Mod. lang. J., 1962, 46, 160-170.

Preston, R. C. Improving the item validity of study habits inventories. Educ. psychol. Measmt., 1961, 21, 129-131.

APPENDIX

TABLE A

Validities for the Prediction of FCS in Key Construction and Cross-validation Samples

Predictors	Language Group: Western, DLIWC, Key Construction (N=139)[a]	Western, DLIWC, Cross-Validation (N=66)	Western, DLIEC, Cross-Validation (N=75)	Eastern, DLIWC, Key Construction (N=169)[a]	Eastern, DLIWC, Cross-Validation (N=83)	Chinese, DLIWC, Key Construction (N=142)[a]	Chinese, DLIWC, Cross-Validation (N=61)	Total, DLIWC, Key Construction (N=450)[a]	Total, DLIWC, Cross-Validation (N=210)
Operational									
FLAT	-42**	-27*	-29**	-45**	-37**	-43**	-45**	-42**	-36**
Set I									
Insolence Score 2	04	15	-03	25**	08	22**	-03	17**	07
Insolence Score 1	-05	24*	03	24**	-05	16*	-03	12**	06
Hand Skills	09	-02	-29**	-10	04	-15	-23	-05	-05
Education Level	-11	-37**	--	-22**	-10	-37**	-19	-23**	-22**
Pay Grade	05	-18	-14	-10	-09	-22**	-13	-08	-11
Age	12	06	12	-12	-03	09	34**	06	15*
Vocabulary Learning 1	-46**	-32**	-60**	-39**	-31**	-25**	-07	-36**	-26**
Vocabulary Learning 2	-39**	-28*	-61**	-36**	-29*	-23**	-12	-32**	-25**
Vocabulary Learning 3	-35**	-22	-56**	-19*	-26*	-13	-07	-20**	-19**
Set II									
PDQ	71	51**	-36**	80	17	71	42**	52	32**
NAL	55	39**	--	43	12	69	24	37	24**
Set III--Instructor's Rating									
1st Week "Success"	-62**	-66**	-71**	-54**	-50**	-62**	-44**	-58**	-55**
1st Week "Oral"	-61**	-65**	-67**	-55**	-50**	-62**	-41**	-59**	-54**
1st Week "Motivation"	-48**	-56**	-46**	-44**	-34**	-58**	-33**	-49**	-42**
1st Day "Success"	--	--	-51**	--	--	--	--	--	--
1st Day "Oral"	--	--	-35**	--	--	--	--	--	--
1st Day "Motivation"	--	--	-17	--	--	--	--	--	--

Notes --

1. Decimal points have been omitted from correlations.
2. Due to the direction of the FCS criterion (i.e., low FCS score indicates high achievement), most validity coefficients are negative.
3. *p < .05.
4. **p < .01.

[a]Since these samples were used to construct the PDQ and NAL keys, significance levels are not appropriate and are therefore excluded for PDQ and NAL.

APPENDIX

TABLE B

Validity Coefficients for Predicting FCS With Operational and Experimental Set I Predictors[a]

Language Group & Predictor Category	Regression Weights & Variables in Equation	Weighting Sample N	Weighting Sample R	Cross-validation N	Cross-validation r
Western (Specific)					
FLAT	-0.920(FLAT)	139	.42[b]	66	.27
FLAT, Set I	-0.849(FLAT)-1.922(Vocabulary Learning Test)	139	.52	66	.35
FLAT, Set I, Ratings	-1.009(1st Week "Success")-1.659(Vocabulary Learning Test)	139	.68	66	.68
Western (General)[c]					
FLAT	-0.752(FLAT)	(Not Computed, See Note "c")		66	.27
FLAT, Set I	-1.374(FLAT)-1.508(Pay Grade)			66	.34
FLAT, Set I, Ratings	-0.714(FLAT)-1.165(1st Week "Oral")			66	.70
Eastern (Specific)					
FLAT	-0.700(FLAT)	168	.45[b]	83	.37
FLAT, Set I	-1.121(FLAT)-1.284(Vocabulary Learning Test)	168	.50	83	.42
FLAT, Set I, Ratings	-1.027(FLAT)-1.136(1st Week "Oral")	168	.64	83	.56
Eastern (General)[c]					
FLAT	-0.752(FLAT)	(Not Computed, See Note "c")		83	.37
FLAT, Set I	-1.374(FLAT)-1.508(Pay Grade)			83	.38
FLAT, Set I, Ratings	-0.714(FLAT)-1.165(1st Week "Oral")			83	.55
Chinese (Specific)					
FLAT	-0.860(FLAT)	142	.43[b]	61	.45
FLAT, Set I	-1.359(FLAT)-2.119(Pay Grade)	142	.54	61	.54
FLAT, Set I, Ratings	-1.427(1st Week "Success")-2.404(Education Level)	142	.65	61	.43
Chinese (General)[c]					
FLAT	-0.752(FLAT)	(Not Computed, See Note "c")		61	.45
FLAT, Set I	-1.374(FLAT)-1.508(Pay Grade)			61	.40
FLAT, Set I, Ratings	-0.714(FLAT)-1.165(1st Week "Oral")			61	.49

Notes --

[a]All single variable validity coefficients are presented as positive for purposes of comparing with multiple R's.

[b]Single variable coefficients are Pearson r's.

[c]General equations are based on the Western, Eastern, and Chinese weighting samples combined and, therefore, applied only to the cross-validation sample of each language group.

APPENDIX

TABLE C

Multiple-Regression Coefficients for Predicting FCS Based on Cross-validation Samples

Language Group & Predictor Category	Regression Weights & Variables in Equation	R
Western (N=66)		
FLAT	-0.557(FLAT)	.27[a]
FLAT, Set I	-3.742(Education Level)-1.389(Vocabulary Learning 1)	.44
FLAT, Set I, Ratings	-1.198(1st Week "Success")-3.053(Education Level)	.70
FLAT, Set I, Set II	-4.069(Education Level)-0.290(PDQ)-0.271(NAL)	.66
FLAT, Set I, Set II, Ratings	-1.064(1st Week "Success")-0.419(NAL)	.74
Eastern (N=83)		
FLAT	-0.601(FLAT)	.37[a]
FLAT, Set I	-0.933(FLAT)-1.002(Vocabulary Learning 2)	.42
FLAT, Set I, Ratings	-0.709(FLAT)-1.240(1st Week "Oral")	.55
FLAT, Set I, Set II	-0.794(FLAT)-0.541(PDQ)	.44
FLAT, Set I, Set II, Ratings	-1.284(1st Week "Oral")-0.566(PDQ)	.57
Chinese (N=61)		
FLAT	-0.986(FLAT)	.45[a]
FLAT, Set I	-1.460(FLAT)-1.816(Pay Grade)	.55
FLAT, Set I, Ratings	-1.460(FLAT)-1.816(Pay Grade)	.55
FLAT, Set I, Set II	-1.460(FLAT)-1.816(Pay Grade)	.55
FLAT, Set I, Set II, Ratings	-1.460(FLAT)-1.816(Pay Grade)	.55
Total (Western, Eastern and Chinese Combined)		
FLAT	-0.685(FLAT)	.36[a]
FLAT, Set I	-1.163(FLAT)-1.572(Pay Grade)	.42
FLAT, Set I, Ratings	-0.662(FLAT)-1.053(1st Week "Success")-1.200(Pay Grade)	.60
FLAT, Set I, Set II	-0.920(FLAT)-1.696(Pay Grade)-0.834(PDQ)	.46
FLAT, Set I, Set II, Ratings	-0.662(FLAT)-1.053(1st Week "Success")-1.200(Pay Grade)	.60

Note --

[a]Single variable coefficients are Pearson r's and are presented as positive for purposes of comparing with multiple R's.

DISTRIBUTION LIST

BUPERS (Pers-A3)	(25)
NPRL WASH	(3)
Army Personnel Research Office Washington	(3)
Personnel Research Laboratory Lackland AFB	(3)
Defense Documentation Center	(20)

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
U. S. Naval Personnel Research Activity San Diego, California 92152	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

SELECTION OF MILITARY PERSONNEL FOR FOREIGN LANGUAGE TRAINING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final

5. AUTHOR(S) (First name, middle initial, last name)

Idell Neumann, Norman M. Abrahams, William H. Githens

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1968	28	16

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. PF0160501B01	SRR 69-6
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Chief of Naval Personnel (Pers-A3) Department of the Navy Washington, D. C. 20370

13. ABSTRACT

A program of research has been initiated for the Defense Language Institute (DLI) to develop tests and other procedures for improving selection of language trainees.

Currently selection of students is primarily based on the Foreign Language Aptitude Test (FLAT). Since present DLI selection procedures do not include systematic measurement of trainee motivation, non-cognitive measures were considered for selecting foreign language students.

In addition to FLAT scores, several measures such as the Personal Data Questionnaire, the Navy Adjective List, and Instructor Ratings were gathered experimentally at the Defense Language Institute West Coast. Using final class standing as a criterion, the predictors were validated alone and in combination.

Prediction of language achievement can be markedly improved by an instructor's rating obtained at the end of only one week of instruction--or even one day if need be. (Pages 5, 7). If "trial training" were implemented for the purpose of obtaining Instructor's Ratings prior to the inception of formal training, considerable expense may be avoided by eliminating those students considered substandard by the instructors.

If brief trial training proves to be infeasible and an ample number of potential trainees are available, some improvement may also be achieved if selection were based on paper and pencil tests, i.e., a combination of the Personal Data Questionnaire and Foreign Language Aptitude Test scores. (Page 10)

If either or both of the two procedures (i.e., paper and pencil tests and Instructor's Rating) are adopted for operational use, follow-up research is recommended on a larger sample of Navy personnel to improve the accuracy of the weights and cutting scores used, since the DLI included members of all the military forces.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Motivation						
Foreign language						
Language achievement						
Non-cognitive predictors						
Selection						
Instructor's ratings						
Foreign Language Aptitude Test						