DOCUMENT RESUME

ED 281 862 TM 870 243

AUTHOR	Wilson, Kenneth M.
TITLE	The Relationship of GRE General Test Scores to First-Year Grades for Foreign Graduate Students: Report of a Cooperative Study.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Graduate Record Examinations Board, Princeton, N.J.
REPORT NO	ETS-RR-86-44; GREB-82-11P
PUB DATE	Dec 86
NOTE	82p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*College Entrance Examinations; Departments; *English (Second Language); *Foreign Students; *Grade Point Average; Graduate Students; Graduate Study; Higher Education; Language Proficiency; Native Speakers; *Predictive Validity; Scores; Student Characteristics
IDENTIFIERS	*Graduate Record Examinations; Test of English as a Foreign Language

ABSTRACT

The validity of the Graduate Record Examinations (GRE) was examined for foreign students enrolled in U.S. graduate schools. Subjects included 1,353 foreign students for whom English was a second language (ESL) and 42 foreign students whose native language was English. The relationships between college departments' scores on the GRE General Test and first year average grades were examined for three populations: (1) foreign ESL students who were heterogeneous with respect to linguistic, cultural, and educational background; (2) subgroups with homogeneous country of origin and background variables; and (3) subgroups classified according to English proficiency, as indicated by Test of English as a Foreign Language (TOEFL) scores; GRE verbal, analytical, and quantitative scores; and self-reported English language proficiency. The students were highly selected and represented mainly quantitative departments of study. The majority of students were Asian; about half were from India, Taiwan, or Korea. Results for these samples were comparable to validity data for American students. In quantitative fields of study, quantitative and analytical scores were the strongest predictors; in verbal fields, verbal scores were strongest. GRE verbal and TOEFL scores had parallel patterns of validity. (GDC)

GRE

GRADUATE RECORD EXAMINATIONS

THE RELATIONSHIP OF GRE GENERAL TEST SCORES TO FIRST-YEAR GRADES FOR FOREIGN GRADUATE STUDENTS: REPORT OF A COOPERATIVE STUDY

Kenneth M. Wilson

GRE Board Professional Report GREB No. 82-11P
ETS Research Report 86-44

December 1986

This report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

EDUCATIONAL TESTING SERVICE, PRINCETON, NJ

The Relationship of GRE General Test Scores to First-Year Grades for Foreign Graduate Students: Report of a Cooperative Study

Kenneth M. Wilson

GRE Board Professional Report GREB No. 82-11

December 1986

Acknowledgments

This study was made possible by the support and encouragement of the Graduate Record Examinations Board and the cooperation of participating graduate schools and departments.

Richard Harrison helped plan for data collection and file development. Ka Ling Chan helped in data analysis.

Henry Braun, Nancy Burton, and Charles Seckolsky made helpful reviews of the manuscript.

These contributions are gratefully acknowledged. The writer, however, is solely responsible for the contents of this report.

Abstract

The present study was designed as a GRE population validity study for foreign students enrolled in U.S. graduate schools. Primary interest was in validity for students for whom English is a second language (foreign ESL students). The objectives of the study were to obtain information regarding the typical within-department relationship of scores on the GRE General Test to first-year average grades (FYA) (a) in samples of foreign ESL students that are heterogeneous with respect to linguistic-cultural-educational background, (b) in subgroups that are homogeneous with respect to country of origin and associated background variables, and (c) in subgroups classified according to relative level of English proficiency, as measured by scores on the Test of English as a Foreign Language (TOEFL), GRE verbal and analytical performance relative to quantitative performance, and self-reported English language communication status.

The study was based on data for a total of 1,353 foreign ESL students and 42 foreign ENL (English-native-language) students from 97 graduate departments in 23 graduate schools. Eighty-six departments were primarily quantitative—either engineering, math/science, or economics; six were bioscience departments; and five were primarily verbal—four education and one political science. More than 90 different countries were represented in the sample. The majority of students were from Asia. The three largest national contingents (students from India, Taiwan, and Korea) accounted for about one half of all students. Students were highly selected both in terms of quantitative ability and in terms of English proficiency as measured by TOEFL.

The department-level samples were small (median N = 12). In order to obtain reliable estimates of within-department GRE/FYA relationships, data for similar departments were pooled. GRE and FYA variables were z-scaled by department before pooling—that is, scores were expressed as deviations from deparment-level means in department standard deviation units. Primary emphasis was given to analyses based on pooled data for the primarily quantitative departments, because representation of departments from the other fields was limited.

Average levels and patterns of GRE/FYA correlations based on pooled data for foreign ESL students from the 86 quantitative departments and the five "verbal" departments were found to be comparable to levels and patterns of coefficients that have been reported by the GRE Validity Study Service (VSS) for a sample composed primarily of U.S. citizens; results for the small sample of bioscience departments were anomalous, due probably to sampling effects. In quantitative fields, quantitative and analytical scores were the strongest predictors; in the verbal fields, verbal scores were strongest. GRE verbal and TOEFL total scores had parallel patterns of validity.

Study findings suggested that inferences based on GRE scores regarding the subsequent academic performance of applicants, especially those applying for admission to primarily quantitative departments, are likely to be as valid for foreign ESL applicants as for U.S. applicants. Questions regarding the comparative academic performance of U.S. and foreign students with comparable GRE scores, not addressed directly in this study, call for further research.

Table of Contents

Acknowledgments		i
Abstract		iii
Summary		S-1
Section I:	Background	1
	The Present Study	3
	Distribution of Students by School and Department	4
	National Origin and English-Language Background of Students in the Sample	7
Section II:	Description of Study Variables and Sample Performance on the Variables	11
	Variables Related to English Proficiency	12
	Performance on the Study Variables	14
	Intercorrelations of Selected Variables in the Total ESL Sample	18
Section III:	Study Methods and Procedures	21
	General Methodological Rationale	21
	Application in the Present Study	23
Section IV:	GRE/FYA Relationships for Foreign ESL Students, by Academic Area	27
	Interpretive Perspective	27
	GRE/FYA Validity Coefficients for Foreign ESL Samples	28
	Multiple Regression Results for Foreign ESL Students by Academic Area	30
	Validity of a Standard GRE Predictive Composite for Quantitative Departments	32
	Simple Correlation of Selected Non-GRE Variables with FYA	32
Section V:	Zgre/Zfya Relationships for Subgroups Differing in Performance on Selected English-Proficiency-Related Measures—Pooled Z-scaled Data for Quantitative Departments	37

Procedures Involved in Defining Subgroups 37

Results 38

Comparative Performance of Subgroups 40

Section VI: Correlation of a Standard Z(gre) Predictive Composite with Z(fya) for Foreign ESL Students Classified by National Origin and Graduate Major Area—Data for Quantitative Departments 41

Section VII: An Exploratory Analysis of the Comparative Performance of Regional Subgroups on FYA and GRE Variables—Students from All Quantitative Departments 45

Findings 47

Observed versus Predicted Criterion Standing of Regional Groups 49

References 53

Appendix A Summary of Department-Level Data 55

List of Tables and Figures

Table 1	Distribution of Foreign ESL (English Second Language) Students by School and Department	5
Table 2	Distribution of 97 Study Departments by Size of ESL Sample	6
Table 3	Distribution of Students by Country of Citizenship and Academic Area for 39 Countries Represented by Five or More Students	8
Table 4	Distribution of Foreign ESL Students by Country of Citizenship within "Regional" Classifications Based on Geographic and English-Background-Related Variables . . .	10
Table 5	Summary Statistics for the ESL Sample on Selected Study Variables, by Broad Academic Areas, with Comparative Data for all ENL Students	15
Table 6	Intercorrelations of Selected Study Variables in the Total FSL Sample	19
Table 7	Simple Correlation of GRE General Test Scores with FYA for Foreign ESL Students: Weighted (Size-Adjusted) Means of Department-Level Coefficients for Departments Grouped by Graduate Field and Area	29
Table 8	Regression Results for ESL Students Using Departmentally Standardized Data Pooled by Graduate Major Area	31
Table 9	Validity Coefficients for a Standard Composite of GRE Quantitative and Analytical Scores versus Coefficients for Individual GRE Scores: Size-Adjusted Averages of Department-Level Coefficients for Quantitative Departments Grouped by Fields and Major Areas	33
Table 10	Simple Correlations of Selected Background Variables with FYA for Foreign ESL Students: Size-Adjusted Means of Department-Level Coefficients for Departments Grouped by Graduate Field and Area	34
Table 11	GRE/FYA Relationships for Subgroups Defined by Relative Standing on Selected ESL-Proficiency-Related Test or Self-Report Variables in Analyses Using Departmentally Standardized Predictor/Criterion Data Pooled Across Quantitative Departments	39

Table 12	Correlation of a Composite of Z-Scaled GRE Quantitative (Zq) and Analytical (Za) Scores with Z-Scaled FYA (Zfya) for Subgroups Defined by Country and Region: All Quantitative Fields	42
Table 13	Correlation of a Composite of Z-Scaled GRE Quantitative (Zq) and Analytical (Za) Scores with Z-Scaled FYA (Zfya) for Subgroups Defined by Region and Major Graduate Area	42
Table 14	Means of Regional Groups on Regularly Scaled FYA and GRE Scores: Pooled Data for 1,213 Students in 86 Engineering, Math/Science, and Economics Departments . .	48
Table 15	Means of Foreign Students on FYA and GRE Variables, by Region, Expressed (a) as Deviations from the Means of Their Respective Departments of Enrollment and (b) as Deviations from Grand Means for all Students: Data for All Quantitative Departments	48
Table 16	Grand Mean FYA and Mean FYA' as Compared to Within-Department Mean Zfya and Mean Z'fya for Regional Groups: Pooled Data for All Quantitative Departments	50
Figure 1	Findings based on analysis of data across departments versus findings based on analysis of data within departments	51

Summary

The examinee population taking the GRE General Test and the samples used in standardization and calibration (scaling) are made up predominantly of U.S. citizens toward whom the test is oriented educationally, culturally, and linguistically. However, the test is also taken by foreign nationals, the majority of whom speak English as a second language (ESL) and whose cultural and educational backgrounds differ from those of U.S. examinees.

The average GRE quantitative performance of foreign ESL examinees is fully comparable to that of U.S. examinees in similar fields of study, but their average performance on the GRE verbal and analytical ability measures is markedly lower, due primarily to factors associated with their less-than-native level of English proficiency. For foreign examinees for whom English is the native language (ENL), verbal and analytical averages are consistent with average performance on the quantitative test.

Study Objectives

The present study was designed to obtain information regarding the typical within-department relationship of GRE scores and certain English-proficiency-related variables to first-year average grade (FYA) in selected graduate fields (a) in samples of foreign ESL students that are heterogeneous with respect to linguistic-cultural-educational background variables, (b) in subgroups that are homogeneous with respect to country of origin and associated background variables, and (c) in subgroups classified according to relative level of "English proficiency," as defined by the following variables:

- Total score of individual students on the TOEFL, the Test of English as a Foreign Language (available for about two-thirds of the students)

- Relative Verbal Performance Index (RVPI)—the discrepancy between observed GRE verbal score and that expected for U.S. GRE examinees with given quantitative scores, thought of as reflecting a type of general English proficiency deficit

- Relative Analytical Performance Index (RANPI)—the discrepancy between observed analytical score and that expected for U.S. GRE examinees with given quantitative scores, thought of as reflecting a somewhat different type of English proficiency than that indexed by the RVPI

- Self-reported better communication in English (BCE) status versus better communication in some other language

Study Sample and Data

The study analyzed data for cohorts of foreign students, without regard to visa status, who entered their respective departments in 1982-83 and 1983-84 as full-time students, who earned a first year average, and for whom GRE scores and information regarding country of citizenship were available.

Students with GRE scores from a test administration prior to October 1981 were not included in the study because the analytical test changed in October 1981.

Data were obtained for 97 departments from 23 graduate schools. The cooperating schools and departments supplied first-year average (FYA) grades and TOEFL total scores (as available). They also supplied information on country of citizenship, undergraduate origin (U.S. versus other), sex, and date of birth, if this information was not supplied by the student when registering to take the GRE.

Most of the samples (86 of 97) were from primarily quantitative departments: engineering (electrical, civil, chemical, industrial, and mechanical), mathematics, statistics, chemistry, physics, computer science, or economics. Six were from biological science departments, and 5 were from either education or political science departments (see text, Table 1, for detail).

The department-level samples typically were quite small: median N = 12 (see Table 2). However, across all departments, data were available for a total of 1,353 ESL students: 702 from combined engineering departments, 353 from math-science departments, and 138 from economics departments, plus 55 from bioscience departments and 76 from social science departments (primarily education). Data were also available for 42 ENL students.

The sample was extremely heterogeneous with respect to national origin (see Table 3 and Table 4). Over 90 countries were represented by at least one student, but about 90 percent of all the ESL students were accounted for by 39 countries with five or more nationals in the study sample.

- More than two-thirds of the students (67.8 percent) were from Asia, 11 percent were from Europe, 9 percent from the Americas (excluding Canada), 7 percent from the Mideast, and about 5 percent from Africa.

- The three largest national contingents were from Taiwan, India, and Korea. These contingents accounted for 666 of the 1,353 foreign ESL students.

The students were highly selected in terms of quantitative ability (see Table 5 for detailed summary of sample characteristics). For both ESL and ENL students, the GRE quantitative mean was 684—above the eightieth percentile in the score distribution for all GRE examinees. For ESL students in quantitative departments the mean was 698. Verbal and analytical means for ESL students were 382 and 486; for the ENL students, corresponding values were 546 and 592.

Quantitative means were lower for bioscience and social science students than for those in the primarily quantitative departments. However, for ESL students in all major areas the pattern of relative performance on the respective ability measures was the same: highest on quantitative, lowest on verbal, with analytical in between.

The ESL students were highly selected in terms of English proficiency as

measured by TOEFL: the total score mean, 567, for the sample was at approximately the 84th percentile for all graduate-level TOEFL examinees.

Mean FYAs for the ESL students, except for those in the small sample from bioscience departments, were comparable to those reported by the GRE Validity Study Service (VSS) for samples composed predominantly of U.S. citizens.

- The mean first-year average (FYA) grade for all ESL students was 3.45 (or approximately B+). For 1,213 students from quantitative departments, the mean was 3.49. The small samples of students from bioscience and social science departments had FYA means of 3.18 and 3.44, respectively.

In order to obtain general estimates of GRE/FYA relationships for foreign ESL students (data were not available for U.S. students), the principal analyses were based on data pooled across departments within particular fields. Since the study was primarily concerned with general trends in GRE/FYA relationships, and not with the development of operational prediction equations, it was convenient to standardize the scores of students on each continuous variable, including the GRE predictors. Predictor and criterion scores were z-scaled—that is, they were expressed as deviations from the respective department means in department standard deviation units prior to pooling. Data for the small sample of ENL students were z-scaled using means and standard deviations for ESL students in their respective departments. Pooled-sample coefficients based on the z-scaled variables are equivalent to averages of the corresponding department-level coefficients weighted according to size of sample. The coefficients reported in this study may be thought of as estimates of population values for ESL students in the groups of departments for which data were pooled.

Principal Findings

GRE/FYA Relationships For General Samples of Foreign ESL Students

To determine the typical levels of GRE/FYA correlations in general samples of foreign ESL students in the respective fields, size-adjusted means of department-level GRE validity coefficients were computed for various groups of departments:

a) chemical, civil, electrical, industrial, and mechanical engineering departments, respectively; all engineering departments

b) statistics, chemistry, physics, mathematics, and computer science departments, respectively; all math/science departments

c) economics departments

d) all 86 quantitative departments—(a) + (b) + (c)

e) bioscience departments (six)

f) social science departments (five)

The size-adjusted means of the resulting GRE/FYA correlations (see text, Table 7, for detail), except those for the six bioscience departments, were comparable to size-adjusted mean coefficients that have been reported by the GRE Validity Study Service (VSS) for samples composed predominantly of U.S. citizens.

- For all subgroups of quantitative departments, GRE quantitative scores and GRE analytical scores were more highly correlated with FYA than were GRE verbal scores.* For the pooled sample of 1,213 students from the 86 quantitative departments, size-adjusted mean coefficients were **.311, .275,** and **.097** for **quantitative, analytical,** and **verbal** scores, respectively. Trends were generally similar for engineering, math/science, and economics samples.

- However, for the small sample of 85 students from five social science departments, **verbal scores were most valid** and quantitative scores were least valid. Mean coefficients for **verbal, analytical,** and **quantitative** scores were **.253, .184,** and **.116,** respectively.

- For the six bioscience samples, average coefficients for quantitative and analytical scores were of about the same magnitude (.061 and .081, respectively), while the verbal coefficient was anomalously negative.

Multiple regression analyses were conducted using z-scaled data aggregated for broader groupings of departments. For the quantitative departments, regression results for the combined sample and those for engineering, math-science and economics subgroups were similar (see text, Table 8 and related discussion).

- The standard partial regression (beta) coefficients for quantitative, analytical, and verbal scores in the combined quantitative sample were

* In evaluating the validity coefficients obtained in this study for GRE analytical ability scores it is important to know that prior to October 1985 users were advised by ETS not to consider the analytical score in selective admission. Assuming that this advice was followed, only GRE verbal and quantitative scores were considered directly in admitting the ESL samples, and their relationship with FYA will tend to be reduced somewhat by the resulting restriction in range. For analytical scores, only some "incidental" restriction in range would be expected (they are positively correlated with the verbal and quantitative scores). However, attenuating effects on validity due to restriction of range in selection should tend to be greater for the verbal and quantitative scores than for the analytical scores. Validity studies based on samples tested after October 1985 will be necessary in order to evaluate the relative contribution of the three GRE measures under conditions in which all three were eligible for consideration in selective admission.

.24, .18, and -.01, respectively. (The negative beta coefficient for verbal reflects a negligible suppression effect.)

- Patterns of regression weights for the small bioscience and social science samples were consistent with the patterns of validity coefficients.

When departmentally z-scaled FYA (Zfya) was regressed on departmentally z-scaled GRE (Zgre) variables (Zq, Za, and Zv) in the combined quantitative sample, and in the engineering, math-science, and economics samples, respectively, only quantitative (Zq) and analytical (Za) scores contributed significantly to prediction of Zfya. The best-weighted composite in the combined quantitative sample was .238 Zq + .176 Za = Z'fya (predicted Zfya).

This general equation was used to compute Z'fya for students in each of the 86 quantitative departments. The average department-level validity coefficient for this standard composite tended to be higher than the average coefficient for either quantitative or analytical ability considered separately (see Table 9 and related discussion).

Effects of English Proficiency on GRE Validity*

Regression analyses based on pooled departmentally z-scaled data were conducted for subgroups of students classified according to level of English proficiency, variously defined by (a) the relative verbal performance index or RVPI, (b) the relative analytical performance index or RANPI, (c) self-reported better communication in English (SR-BCE) versus other status, and (d) TOEFL total score, respectively (see Table 11 and related discussion).**

Interpretation of results for classifications based on TOEFL total score was complicated by the fact (a) that many students did not have TOEFL scores and (b) that differences in regression associated with "availability versus nonavailability" of the score were much more pronounced than those associated with differences in TOEFL-score level among those with TOEFL.

However, on balance, controlling for level of English proficiency as defined by these variables did not appear to have a clear moderating effect on the relationship between the z-scaled FYA criterion and the z-scaled GRE scores for ESL students in any of the quantitative subgroups (see Table 11 and related discussion).

* In these and subsequent analyses, only data for students from the 86 quantitative departments were included.

** RVPI = discrepancy between observed verbal score, V, and predicted verbal score, V', where V' = .52 (GRE-Q) + 185, a regression equation based on data for a sample of U.S. examinees tested during 1981-82. RANPI = discrepancy between analytical score, A, and predicted analytical score, A', where A' = .661 (GRE-Q) + 202, based on data for the same general sample.

Parallel validity for TOEFL and GRE verbal. Regression results obtained when Zfya was regressed on Zq, Za, and Zv generally paralleled those obtained when z-scaled TOEFL total score was substituted for Zv. Neither variable contributed significantly to prediction in the primarily quantitative samples.

GRE Validity for National and Regional Subgroups

Simple correlations between a standard composite of z-scaled quantitative and analytical scores (predicted Zfya, or Z'fya, specified by the regression of Zfya on Zq and Za scores in the combined quantitative sample—that is, .238 Zq + .176 Za) were computed (a) for students from all quantitative departments classified by country of citizenship (N > 9) and by regions defined for the study, and (b) for students classified by both region and graduate major area—that is, engineering, math/science, and economics (see Table 12 and Table 13, and related discussion).

The correlation between this composite and Zfya tended to be relatively consistent for subgroups defined in terms of national origin and/or academic area. Correlations did not tend to be higher for subgroups that were homogeneous than for those that were heterogeneous with respect to national origin.

- For 20 national contingents with N > 9, the median Z'fya/Zfya coefficient was r = .36, and for nine regional contingents the median was r = .34, as compared to the general quantitative sample coefficient of r = .35.
- For 22 subgroups (by region and quantitative area) the median coefficient was r = .37.

Comparative Performance of Regional Subgroups on FYA and GRE Variables

The study was not designed specifically to assess the extent to which the average academic performance of students from different countries or groups of countries (regions) tended to be consistent with their average GRE performance. Generally speaking, an assessment along these lines is complicated by the national diversity of the foreign student population, lack of consistency across departments in the national and/or regional mix of enrolled students, and differences in the fields of study selected by various national and regional groups. Members of various national or regional subgroups may not be enrolled in departments that are comparable with respect to grading standards, degree of selectivity, and so on.

Such factors complicate interpretation of observed differences in the average within-department standing of national or regional contingents as reflected in their means on departmentally z-scaled predictors, predictive composites, and/or criterion variables.

Exploratory analyses were conducted to assess (a) differences in the average academic performance of students in several regional classifications

defined for the study and (b) the extent to which the observed differences in academic performance tended to correspond to observed differences in GRE performance. Two parallel analyses were made, one involving comparisons based on departmentally z-scaled data (indicating average within-department standing on predictor and criterion variables) and the other involving comparisons based on original FYA and GRE scores (indicating average standing on the FYA criterion and the GRE scores without regard to department of enrollment).

Results of the regression of FYA on GRE scores for students in the quantitative sample, treated without regard to their departments of enrollment, paralleled almost identically the results of the regression of Zfya on Zgre scores in pooled data for the same sample. Means of predicted FYA values (FYA') were computed using the general sample equation: **.0013 GRE-Q + .0006 GRE-A.** The corresponding predicted Zfya values (Z'fya) were specified by: **.238 Zq + .176 Za.** Thus, regional comparisons were based on two sets of mean observed criterion scores, namely, FYA and Zfya, and the corresponding mean predicted criterion scores, FYA' and Z'fya.

For the most part, patterns of findings regarding regional differences based on the two sets of data were consistent. For example, with one exception the ranking of regional groups in terms of mean FYA was consistent with ranking based on mean Zfya (see Table 14 and Table 15 and related discussion):

- At one extreme, students from Europe had mean FYAs of approximately 3.6, and they also enjoyed relatively high within-department Zfya standing, averaging 0.2+ standard deviations above the foreign ESL means for their respective departments. At the other extreme, students from Africa and the Mideast earned grades averaging approximately 3.3, and they were in departments in which they tended to perform at a lower level than other ESL students (below average by 0.2 standard deviations).

- For students from Asia, who accounted for some 68 percent of all foreign students in the sample, mean FYA was approximately 3.5 (corresponding to the grand FYA mean for all quantitative students without regard to department) and their Zfya means were approximately 0.0. By inference, Asian students tended to provide a substantial common element in the regional mix of the respective department-level samples; both the department-level (Zfya) means and the grand mean FYA reflect substantially the comparatively high FYAs earned by the Asian students.

- Effects associated with department of enrollment were illustrated in data for students from South American countries and Mexico whose FYA mean of 3.4 was lower than the general mean FYA of 3.5 but who were from departments in which they enjoyed above average relative standing (mean Zfya was +0.11).

Patterns of relationships between mean FYA versus mean FYA' were similar to those for mean Zfya versus mean Z'fya (see Table 16 and Figure 1, Section VII, and related discussion). Again, the principal departure from parallelism in the two sets of data was associated with the South American contingent.

- o European students had somewhat higher mean FYA and mean Zfya than predicted from the respective general ESL equations. Students from Africa and the Mideast, regional groups with the lowest ranking on GRE variables were also were the lowest ranked on both mean FYA' and mean Z'fya. The academic standing of Asian students tended to be consistent with expectation based on the general ESL equations.

The findings of these exploratory analyses suggest (a) that there are systematic differences by world region in the average academic performance of foreign students, and (b) that level of academic performance of regional subgroups tends to be generally consistent with expectation based on GRE scores. However, the findings should be thought of primarily as suggesting directions for research designed specifically to assess the possibility that among foreign ESL students, some national or regional subgroups may tend to perform better (less well) than other subgroups with comparable GRE scores.

Implications of Findings

The findings that have been reviewed permit relatively strong inferences regarding the relationship between GRE scores and first-year average grades for foreign ESL students in quantitative fields—students enrolled in engineering, math/science, and economics programs. GRE/FYA relationships for general department-level samples of foreign ESL students in these quantitative areas appear to be quite comparable in both level and pattern to relationships reported for samples of U.S. students in similar fields.

A standard composite of GRE quantitative and analytical scores had general validity for predicting relative within-department standing on the FYA criterion across all the quantitative areas represented: the several engineering fields, math-science fields, and economics. For students in these fields, neither GRE verbal scores nor TOEFL total scores contributed significantly to prediction when included in a battery with GRE quantitative and analytical scores.

The validity of the standard composite was not affected by introducing control for English proficiency, variously indexed, nor was validity stronger for groups that were homogeneous than for groups that were heterogeneous with respect to national origin.

In evaluating this finding, it is important to recall that the students in these quantitative departments were highly selected in terms of quantitative ability as measured by the GRE, and in terms of English proficiency as measured by the TOEFL. They were in fields emphasizing quantitative reasoning abilities and international symbols. These fields presumably require comparatively low levels of general English language verbal communication skill. The opposite may be true for fields such as, say, education.

In the limited sample of ESL students from five "verbal" departments (four education and one political science), the correlation between FYA and GRE verbal scores was stronger than that for either GRE quantitative or analytical scores. This is consistent with a priori expectation based on

validity study findings for U.S. citizens. Accordingly, it seems reasonable to hypothesize that this pattern may tend to be typical for foreign ESL students in primarily verbal fields.

Findings for the six bioscience departments were not consistent with expectation and can be interpreted only as reflecting sampling effects.

The average scores of ESL students on the verbal and analytical ability measures (382 and 486, respectively) were more than one standard deviation lower than would be expected for U.S. examinees with the same quantitative ability—the mean quantitative score for foreign ESL students was 698, only slightly higher than the quantitative mean of 687 reported by the GRE Validity Study Service (VSS) at ETS for math-science samples composed primarily of U.S. citizens (with a GRE verbal mean of 528). The mean FYA for foreign ESL students was approximately 3.5 (or B+ on the grading scale employed), as compared to a mean of 3.4 reported by the GRE VSS for U.S. math-science samples.

This particular pattern of findings suggests not only that the foreign ESL students in this sample tended to be "academically sucessful" but also that they may have tended to receive higher grades, on the average, than their U.S. counterparts with comparable scores on the GRE. Research is needed to assess the validity of this inference as well as the possibility that the average academic performance of some regional or national contingents may not be consistent with their average performance on the GRE.

With regard to the analytical measure, it is again noted that scores on this measure presumably were not used directly in selecting the samples under consideration in this study, whereas scores on the quantitative and verbal measures were used directly in selection. The contribution of the analytical measure may be overestimated somewhat in the current samples.

On balance, the study findings suggest that inferences based on GRE scores regarding the subsequent academic performance of foreign ESL applicants, especially those applying for admission to primarily quantitative departments, are likely to be as valid as those for U.S. applicants.

Section I: Background

The Graduate Record Examinations (GRE) General Test is widely used for evaluating the academic qualifications of applicants for admission to graduate programs in the United States. The General Test traditionally has provided measures of developed verbal and quantitative reasoning abilities (GRE verbal or GRE-V and GRE quantitative or GRE-Q). In 1977, the test was restructured (see Miller & Wild, 1979) to include a measure of developed analytical reasoning ability, a revised version of which was introduced in October 1981 (see, for example, ETS, 1985). There was some restructuring of the verbal and quantitative measures in terms of format and time allotments, but no change in test content was involved. Prior to the October 1985 test administration, users were advised not to consider the analytical score in admission.

The verbal test includes antonyms, analogies, sentence completion, and reading passages or reading comprehension item-types, and the quantitative test includes quantitative comparison, regular mathematics, and data interpretation item-types. These verbal and quantitative items sample two well-established ability domains. Less is known regarding the "ability domain(s)" sampled by the analytical ability measure, which includes 38 analytical reasoning (AR) and 12 logical reasoning (LR) item-types, both of which call for considerable verbal processing.

The examinee population taking the GRE General Test and the samples used in standardization and calibration (scaling) are made up predominantly of U.S. citizens toward whom the tests are oriented educationally, culturally, and linguistically. However, the test is also taken by foreign nationals. During 1981-82, for example, non-U.S. citizens representing more than 140 different countries, territories, or other geopolitical entities made up approximately 16 percent of the total GRE examinee population (Wilson, 1984a, 1984b).

There are large differences between the population of U.S. examinees and the population of foreign examinees in average performance on the verbal and analytical measures. The verbal and analytical performance of foreign examinees, largely individuals for whom English is a second language (ESL examinees), is markedly lower than that of U.S. examinees or of foreign ENL (English native language) examinees. The depressed verbal performance of foreign ESL examinees is attributable, primarily, to their less-than-native levels of proficiency in English.

However, performance on the GRE quantitative measure does not appear to vary with English language background. U.S. examinees, foreign ESL examinees, and foreign ENL examinees in the same fields of study tend to have similar quantitative means. National contingents with very depressed verbal and analytical means frequently have very high quantitative means.

In essence, it appears that for foreign ESL examinees GRE verbal test items (indeed, all English-language verbal test items) are measuring selected aspects of "developing ESL proficiency" rather than level of "developed verbal reasoning abilities," which the test measures in samples of native English speakers. Thus, population differences in level of "verbal reasoning ability" or "analytical reasoning ability" cannot be inferred from observed population differences between U.S. vs foreign ESL examinees in test performance.

At the same time, the population differences in verbal test performance represent "real" population differences in "functional ability" to perform English-language tasks such as those represented by the verbal and analytical test items under testing conditions, including time constraints.

Quantitative test items, on the other hand, appear to be measuring the same underlying abilities for foreign ESL examinees as for U.S. examinees. Available evidence suggests that U.S. and foreign examinees with similar GRE quantitative scores have similar levels of quantitative ability.

Given evidence of systematic population differences both in background and in test performance—evidence permitting the strong inference that the verbal and analytical tests are not measuring the same underlying "ability constructs" for U.S. and foreign ESL examinees (prospective students)—questions naturally arise regarding the criterion-related or predictive validity of GRE scores for members of the foreign ESL student population. For example:

- For individuals in general samples of foreign students, how valid are the GRE scores for predicting subsequent performance on some criterion of success in graduate school (say, first-year average grade, or FYA)? Are the GRE/FYA correlations (GRE/FYA validity coefficients) for general samples of foreign ESL students comparable to those typically observed for general samples of U.S. students? Do GRE/FYA validity coefficients tend to be comparable for different national contingents of foreign students? Are GRE/FYA correlations typically stronger in samples that are relatively homogeneous with respect to linguistic-cultural-educational background variables (for example, samples from specific countries or from countries judged to be similar with respect to language, culture, or educational systems) than in heterogeneous samples? Is criterion-related validity typically higher in samples of foreign ESL students with "higher levels of English proficiency" than for those with "lower levels of English proficiency?"

- How closely does relative criterion standing of various national contingents correspond to their relative standing on GRE predictive composites developed for general ESL samples? Do students with more "ESL proficiency," tend to outperform students with less "ESL proficiency"—for example, do they tend to earn higher mean FYA than would be expected for foreign ESL examinees generally, who have similar GRE scores? Within the population of foreign ESL students, do some national-linguistic subgroups tend to earn higher average grades than would be expected for foreign ESL students generally, who have similar GRE scores?

The Present Study

The present study was designed as a GRE population validity study for foreign ESL students. It was primarily concerned with obtaining and evaluating evidence bearing on the first set of questions outlined above. More specifically, the study was designed primarily to assess:

- the level and pattern of GRE/FYA correlations in general samples of

foreign ESL students in selected graduate fields,

- o the possibility of systematic differences in the level and pattern of GRE/FYA correlations for subgroups of foreign ESL students, especially subgroups differing in English language background as reflected by national-linguistic origin, and by level of performance on various English-proficiency-related test variables, including total score on the Test of English as a Foreign Language or TOEFL (ETS, 1983), and

- o whether improved prediction of FYA for foreign ESL students might be expected by considering information regarding national origin and scores on the TOEFL in conjunction with GRE scores.

The study was not designed to address questions regarding the comparative performance of various subgroups of foreign examinees. However, exploratory analyses were undertaken that permitted limited inferences in this regard.

The study was conducted by Educational Testing Service (ETS) under the auspices of the Graduate Record Examinations Board. In March 1984, 100 graduate schools that usually receive the largest number of GRE score reports from non-U.S. citizens were invited to cooperate in a study with the foregoing objectives by providing data for samples of non-U.S. students, (a) regardless of U.S. visa status, who (b) were first-time, full-time students during the academic years 1982-83 and 1983-84, respectively, (c) completed the academic year in which they initially enrolled, and (d) earned a first-year average grade.

The graduate fields (departments) targeted for study were primarily quantitative fields known to be most popular among international graduate students: engineering—chemical, civil, electrical, industrial, and mechanical; mathematics, chemistry, physics, computer science, statistics, and economics. Several fields that involve more verbal (less heavily quantitative) subject matter were also targeted: education and political science; agriculture, biology, biochemistry, and microbiology.

Over 100 departments provided name, sex, and date of birth identification for the designated entering cohorts of foreign students. This identification was used to locate records of the students in the GRE computerized history file—records containing GRE General Test scores and the corresponding test administration date(s) and other relevant information (if provided by individuals when they registered to take the GRE).

Based on the file-matching procedures, data collection rosters were prepared by ETS and sent to the departments. Departments were asked (a) to supply a first-year average grade and, if available, TOEFL total score, (b) to identify native English-speakers, and (c) to provide information regarding country of citizenship and/or U.S. vs. non-U.S. undergraduate origin (when the roster indicated that this information was not available in the GRE file for a student).

Basic requirements for inclusion of a departmental data-set in the study were that the department have a minimum of five foreign ESL students with:

c verbal, quantitative, and analytical ability scores from a restructured GRE General Test that was administered after September 1981,

o a first-year average grade, and

o data on country of citizenship.

Distribution of Students by School and Department

A total of 97 departments from 23 graduate schools met these basic requirements. The departments (fields) represented and their grouping by academic area, were as follows:

1. Engineering (chemical, civil, electrical, industrial, mechanical)
2. Math/Science (applied math, statistics, chemistry, physics, computer science)
3. Economics
4. Quantitative total (engineering + math/science + economics)
5. Bioscience (microbiology, agriculture, biochemistry, biology)
6. Social science (education, political science)

Table 1 shows the distribution of ESL students by school and graduate department. The 23 schools (identified only by a two-digit study code) ranged from the top 10 percent through the bottom 10 percent among the 100 graduate schools that typically receive the largest number of GRE score reports from non-U.S. citizens.

- o There were 86 department-level ESL samples from primarily quantitative fields, with a total of 1,213 foreign ESL students: 40 engineering samples with a total of 702 students, 36 mathematics and physical science samples (373 students), and 10 economics samples (138) students.

- o Data were available for only 6 biological science samples (55 students) and 5 social science samples (4 education samples with a total of 76 students, and 1 political science sample with 9 students).

Table 1 shows that most of the department-level samples were quite small. The actual distribution of Ns for the 97 foreign-ESL samples, shown in Table 2, points up the positively skewed nature of the sample-size distribution.

- A total of 32 samples included between 5 and 9 ESL students, 36 samples included between 10 and 14 students, 13 included between 15 and 19 students, and 16 included 20 or more students. The median N was 12.

Data were also available for a total of 42 non-U.S. English-native-language students from major English-speaking countries (31 from the quantitative fields, one from a bioscience field, and 10 from the social science fields). These data were employed in the study primarily to point up differences between foreign ENL students and foreign ESL students in patterns of average performance on the GRE General Test.

Table 1

Distribution of Foreign ESL (English Second Language) Students by School and Department

Sch	Department code: Engineering						Math/ Physical Sci							Econ		Bioscience				Educ	P.S.		
	64	65	66	67	68	Tot	54	59	62	72	76	78	Tot	84	Quan tot	07	31	34	35	Bio	85	92	Soc sci	Total
02	17	32	-	12	-	94	-	9	-	-	6	-	15	10	119	-	-	-	-	0	26	-	26	102
07	9	30	33	12	7	91	-	11	11	-	13	12	47	15	153	-	7	8	-	15	17	9	26	194
10	14	-	31	-	-	45	-	-	-	-	-	-	0	-	45	-	-	-	-	0	-	-	0	45
13	-	-	14	9	14	37	-	-	-	8	8	-	16	12	65	-	-	-	-	0	-	-	0	65
14	6	-	31	-	20	57	-	15	-	-	-	9	24	6	87	-	-	-	-	0	-	-	0	87
19	-	10	-	-	-	10	-	-	-	-	-	-	10	-	10	-	-	-	-	0	-	-	0	10
24	5	42	19	13	24	103	-	12	10	15	-	21	58	11	172	6	-	-	-	6	21	-	21	199
28	-	-	-	-	-	0	5	7	26	-	-	-	38	-	38	-	-	-	-	0	-	-	0	38
30	13	13	12	-	10	48	-	-	-	-	5	11	16	11	75	-	-	-	-	0	-	-	0	75
32	-	-	-	-	-	0	-	-	-	-	-	-	0	48	48	-	-	-	-	0	-	-	0	48
33	-	11	36	-	17	64	-	-	13	-	8	16	37	-	101	-	-	-	-	0	12	-	12	113
34	-	-	52	-	-	52	-	-	5	5	-	-	10	5	67	-	-	-	-	0	-	-	0	67
38	-	19	-	25	13	57	-	-	10	11	-	12	33	-	90	-	16	-	-	16	-	-	0	106
41	-	-	15	-	-	15	-	-	5	-	11	5	21	12	48	-	-	-	-	0	-	-	0	48
49	-	5	13	-	-	18	-	-	-	-	-	7	7	-	25	-	-	-	7	7	-	-	0	32
60	-	-	-	-	-	0	-	-	-	-	-	13	13	-	13	-	-	-	-	0	-	-	0	13
65	5	-	-	-	10	15	-	-	6	-	-	-	6	-	21	-	-	-	-	0	-	-	0	21
67	-	-	-	-	-	0	-	-	9	-	-	-	9	-	9	-	-	-	-	0	-	-	0	9
73	11	-	-	18	-	29	-	-	-	-	-	-	0	-	29	-	-	-	-	0	-	-	0	29
82	-	-	-	-	-	0	-	-	-	-	-	7	7	-	7	-	-	-	-	0	-	-	0	7
87	-	-	-		-	0	-	-	-	-	-	-	0	8	8	-	-	-	-	0	-	-	0	8
91	-	-	-	-	-	0	-	-	-	-	-	-	0	-	0	-	-	11	-	11	-	-	0	11
92	-	-	-	-	-	0	-	-	16	-	-	-	0	-	16	-	-	-	-	0	-	-	0	16
Total	80	162	256	89	115	702	5	54	111	39	51	113	373	138	1213	6	23	19	7	55	76	9	85	1353
Depts	8	8	10	6	8	40	1	5	10	4	6	10	36	10	86	1	2	2	1	6	4	1	5	97

Note. Quant(itative) total is sum of Ns for Engineering, Math & Physical Science, and Economics.

Departments and codes: Engineering--64 Chemical, 65 Civil, 66 Electrical, 67 Industrial, 68 Mechanical ; 54 Applied Math, 59 Statistics, 62 Chemistry, 72 Mathematics, 76 Physics, 78 Computer Science; 84 Economics; 07 Microbiology, 31 Agriculture, 34 Biodhemistry, 35 Biology; 85 Education, 92 Political Science.

Table 2

Distribution of 97 Study Departments by Size of ESL Sample

5	2	(1)
4*	8	(1)
4	2	(1)
3*	6	(1)
3	01123	(5)
2*	566	(3)
2	0114	(4)
1*	55556 66777 899	(13)
1	00000 01111 11111 12222 22222 33333 33344 4	(36)
0*	55555 55555 66666 77777 78888 89999 99	(32)

Note. Sample sizes are specified by combining leading digits with successive digits in the respective rows. For example, 52, 48, 42, 36, 30, 31, 31, 32, 33, and so on. Mdn N = 12.

In view of the limited number of bioscience and social science samples available, and their typically small size, the study focussed primarily on data for ESL students in the 86 engineering, math-science, and economics departments.

National Origin and English-Language Background of Students in the Sample

The sample was extremely diverse with respect to national origin. Over 95 of 140 countries (territories, protectorates, or other geopolitical entities) represented in the GRE examinee population were represented in the study sample. However, more than 90 percent of the 1,353 foreign ESL students were accounted for by 39 countries that were represented by at least five students. Table 3 lists the 39 countries in order of total representation in the study sample, and shows the distribution of student-nationals by academic areas.

- The largest contingents of students were from Asian countries. More than two-thirds of the students (67.8 percent) were from Asian countries, 10.9 percent were from Europe, 8.6 percent from the Americas (excluding Canada), 7.4 percent from the Mideast, and 4.9 percent from Africa.

English-Language Background and National Origin

The last column of the table shows the TOEFL total mean reported by ETS (1983, Table 10) for all U.S.-bound TOEFL examinees from each country, tested during 1980-82, without regard to educational level, designated as TOEFL-LEVEL to help reinforce the fact that they are not the TOEFL means of students in the sample.

- The national means (TOEFL-LEVEL values) shown in Table 3 indicate typical levels of English proficiency, as measured by the TOEFL, for contingents of prospective U.S. students from the respective countries. They may be thought of as reflecting, in part, background differences associated with national-linguistic origin (a) in the usual patterns of English language acquistion and usage (for example, type, duration, intensity, and quality of experience in the use of English as a second language), and (b) degree of overlap between native language and English. U.S.-bound students from India, Singapore, or the Philippines, for example, typically have had a "richer" English language background including more experience in using English in both general and academic settings than, say, students from Taiwan, Japan, Thailand, or the Mideast, whose native languages and English, moreover, have relatively few common elements. The degree of overlap between native language and English varies markedly and it may be seen that TOEFL-LEVEL values also tend to be higher for examinees from western European countries than for Asian or Mideastern examinees.

- Background differences other than those involving English-language usage are also partially reflected by differences in TOEFL means. In a study of TOEFL examinees tested during 1977-79 (Wilson, 1982a), moderate correlations (approximately .5) were found between TOEFL means and published indicators of national development such as literacy rate, higher education

Table 3

Distribution of Students by Country of Citizenship and Academic Area for 59 Countries Represented by Five or More Students

Country	Engin	M&PS	Econ	Quant total	Biosci	Soc Sci	All fields	TOEFL LEVEL*
Taiwan	188	94	12	294	10	16	320	(493)
India	108	63	5	176	4	1	181	(555)
Korea	68	52	26	146	7	12	165	(504)
Iran	50	10	1	61	2	1	64	(484)
Japan	26	7	7	40	2	2	44	(487)
PRepChina	17	22	2	41	2	1	44	(473)
Greece	28	8	4	40	2	1	43	(502)
Hong Kong	19	9	3	31	0	1	32	(508)
Thailand	17	2	1	20	3	9	32	(473)
Malaysia	12	4	5	21	1	5	27	(534)
Pakistan	13	6	2	21	2	0	23	(518)
Mexico	7	3	5	15	2	5	22	(514)
Chile	3	5	8	16	3	0	19	(520)
Turkey	12	1	4	17	1	0	18	(500)
W Germany	3	10	4	17	0	0	17	(576)
Egypt	7	3	3	13	2	0	15	(485)
Colombia	5	4	4	13	0	1	14	(508)
Denmark	11	2	0	13	0	0	13	(580)
Spain	4	0	9	13	0	0	13	(543)
Brazil	5	4	0	9	1	3	13	(513)
Venezuela	5	4	0	9	0	3	12	(479)
Lebanon	11	1	0	12	0	0	12	(487)
Peru	3	3	2	8	1	1	10	(513)
Indonesia	2	1	6	9	0	1	10	(479)
Israel	6	2	0	8	0	1	9	(528)
Nigeria	6	1	0	7	1	1	9	(509)
Singapore	3	1	1	5	0	4	9	(563)
Philippines	1	4	2	7	0	1	8	(547)
Algeria	5	3	0	8	0	0	8	(511)
Yugoslavia	3	5	0	8	0	0	8	(530)
Jordan	4	3	0	7	0	1	8	(458)
Cyprus	2	3	1	6	0	0	6	(482)
Ivory Coast	0	0	4	4	1	1	6	(498)
Argentina	2	2	2	6	0	0	6	(543)
Nepal	5	0	0	5	0	1	6	(523)
Italy	4	0	0	4	0	1	5	(552)
Tanzania	1	1	0	2	1	2	5	(542)
Bangladesh	2	1	2	5	0	0	5	(483)
Vietnam	3	2	0	5	0	0	5	(497)
Subtotal	671	346	125	1142	48	76	1266	
% of Total	95.6	92.8	88.7	94.1	87.3	89.4	93.6	
Total	702	373	141	1213	55	85	1353	

*TOEFL-LEVEL is the TOEFL Total mean for all TOEFL-takers from the country during the period 1980-82, as reported by ETS (1983, Table 10), not the mean for student-nationals in the present study. The grand mean for all TOEFL examinees during 1980-82 was 503, S.D. = 66.

enrollment rate, per capita expenditures for education, and so on, in analyses involving means for some 100 different countries. Students from higher TOEFL-LEVEL countries as compared to those from lower TOEFL-LEVEL countries may tend to enjoy some educational advantages as well as advantages in the acquisition of general ESL proficiency.

Mean TOEFL-LEVEL differences appear to be quite stable over time. In the study cited above, a correlation of .94 was found between means of 129 national contingents of U.S.-bound TOEFL-takers in two different testing years.

A classification of countries of citizenship that takes into account both geographic location (world region) and characteristic differences in English language background is provided in Table 4. U.S.-bound students (who take U.S. admission-related examinations) from Category I countries (Asia I, Europe I, Africa I, America I), as compared to their Category II counterparts, typically, (a) have had more extensive experience in using English as a second language, and/or either (b) earn higher average scores on the TOEFL, and on other English-language verbal measures such as GRE Verbal and GMAT Verbal, or (c) earn verbal scores that are relatively more consistent with expectation based on their quantitative scores (see, ETS, 1983; Powers, 1980; Wilson 1982a, 1982b, 1984a, 1984b, 1985, 1986b). No within-region subgrouping was judged to be feasible for the Mideastern countries, which tend to have Category II characteristics.

The last column of the table shows for each regional classification the average of the TOEFL-LEVEL values for the respective member countries (shown in Table 3) weighted according to the number of students represented in the present sample, including students from "other" countries. To reiterate, these TOEFL-LEVEL values are not the TOEFL means of students in the sample.

As a working proposition, it was assumed that the backgrounds of U.S.-bound students from Category I countries as compared to Category II countries, typically, were more conducive to development of general ESL communicative competence.

Table 4

Distribution of Foreign ESL Students by Country of Citizenship within "Regional" Classifications Based on Geographic and English-Background-Related Variables

Region/ Country*	Academic area: Engin	M&PS	Econ	Quant total	Biosci	Soc Sci	Total	TOEFL LEVEL*
Asia II	(374)	(203)	(66)	(643)	(29)	(48)	(720)	496
Taiwan	188	94	12	294	10	16	320	
Korea	68	52	26	146	7	12	165	
Japan	26	7	7	40	2	2	44	
PRepChina	17	22	2	41	2	1	44	
Hong Kong	19	9	3	31	0	1	32	
Thailand	17	2	1	20	3	9	32	
Malaysia	12	4	5	21	1	5	27	
Pakistan	13	6	2	21	2	0	23	
Indonesia	2	1	6	9	0	1	10	
Nepal	5	0	0	5	0	1	6	
Bangladesh	2	1	2	5	0	0	5	
Vietnam	3	2	0	5	0	0	5	
Other	2	3	0	5	2	0	13	
Asia I	(112)	(68)	(8)	(188)	(4)	(6)	(198)	556
India	108	63	5	176	4	1	181	
Singapore	3	1	1	5	0	4	9	
Philippines	1	4	2	7	0	1	8	
Other	0	0	0	0	0	0	0	
Mideast	(76)	(20)	(2)	(98)	(2)	(4)	(104)	483
Iran	50	10	1	61	2	1	64	
Lebanon	11	1	0	12	0	0	12	
Israel	6	2	0	8	0	1	9	
Jordan	4	3	0	7	0	1	8	
Other	5	4	1	10	0	1	11	
America II	(35)	(27)	(24)	(86)	(7)	(16)	(109)	512
Mexico	7	3	5	15	2	5	22	
Chile	3	5	8	16	3	0	19	
Colombia	5	4	4	13	0	1	14	
Brazil	5	4	0	9	1	3	13	
Venezuela	5	4	0	9	0	3	12	
Peru	3	3	2	8	1	1	10	
Argentina	2	2	2	6	0	0	6	
Other	5	2	3	10	0	3	13	

Region/ Country	Academic area: Engin	M&PS	Econ	Quant total	Biosci	Soc Sci	Total	TOEFL LEVEL*
America I	(0)	(3)	(3)	(6)	(1)	(0)	(7)	555
Europe II	(49)	(20)	(9)	(78)	(3)	(2)	(83)	507
Greece	28	8	4	40	2	1	43	
Turkey	12	1	4	17	1	0	18	
Yugoslavia	3	5	0	8	0	0	8	
Cyprus	2	3	1	6	0	0	6	
Other	4	3	0	7	0	1	8	
Europe I	(29)	(18)	(15)	(62)	(0)	(2)	(64)	561
W Germany	3	10	4	17	0	0	17	
Denmark	11	2	0	13	0	0	13	
Spain	4	0	9	13	0	0	13	
Italy	4	0	0	4	0	1	5	
Other	7	6	2	15	0	1	16	
Africa II	(16)	(9)	(8)	(33)	(6)	(3)	(42)	497
Egypt	7	3	3	13	2	0	15	
Algeria	5	3	0	8	0	0	8	
Ivory Coast	0	0	4	4	1	1	6	
Other	4	3	1	8	3	2	13	
Africa I	(10)	(5)	(2)	(17)	(3)	(4)	(24)	539
Nigeria	6	1	0	7	1	1	9	
Tanzania	1	1	0	2	1	2	5	
Other	3	3	2	8	1	1	10	
Total	702	373	141	1213	55	85	1353	510

Note. Classification of countries within several of the world regions (Asia I versus Asia II, Europe I versus Europe II, and so on) takes into account differences in English-language background associated with national-linguistic origin. Contingents of U.S.-bound students from Category I countries as compared to those from Category II countries are assumed to have backgrounds that are more conducive to the acquistion of ESL proficiency. Consider, for example, typical patterns of ESL acquisition and usage in America II (e.g., Mexico and South America) versus America I, a classification that includes Guyana, Haiti, Honduras, Jamaica, Trinidad and Tobago (none of which was represented by five or more nationals in the study sample), countries with strong English-speaking traditions.

* The regional TOEFL-LEVEL values shown in the last column are averages of the member-country TOEFL-LEVEL values weighted according to the number of students from those countries (see Table 3 for country values and definitional notes).

Section II: Description of Study Variables and Sample Performance on the Variables

GRE General Test verbal, quantitative, and analytical ability scores and a first-year average grade were available for all students. Data on undergraduate origin (U.S. versus other), age, and sex were available for most students. Several variables thought of as reflecting different aspects of "acquired ESL proficiency" were also employed:

- o total score on the Test of English as a Foreign Language (TOEFL)
- o the discrepancy between observed GRE verbal scores and the scores that would be predicted for U.S. examinees with given quantitative scores (a relative verbal performance index, or RVPI)
- o a similarly derived discrepancy index for the analytical score relative to the quantitative score (a relative analytical performance index, or RANPI),
- o self-reported English language communication status (better communication in English (BCE) than in any other language versus the opposite status).

More detailed information regarding these English-proficiency-related variables and their role(s) in the study, findings regarding the performance of the sample on these and other study variables, and the intercorrelations of study variables in the total sample sample are provided later in this section.

Scaled-scores (200 - 800) for the GRE verbal, quantitative, and analytical measures were available from testing after September 1981. The predictive properties of the traditional verbal and quantitative measures are well known. Less is known regarding patterns of predictive validity in different academic areas for the analytical ability measure that was introduced in October 1981. Positive correlations are expected a priori for GRE scores and academic criteria. Negative validity coefficients for the GRE are theoretically anomalous and, if found, usually may be explained in terms of sampling error or selection effects.

The FYA criterion was reported on the same numerical scale by all departments, namely, A = 4, B = 3, C = 2, D = 1, and F = 0. As is well known, the FYA metric is grading-context specific—that is, the "level" of academic performance associated with a given FYA cannot be be assumed to be comparable across different departments. In assessing GRE/FYA validity, therefore, attention is focussed primarily on analysis of within-department relationships.

At the same time, grades generally have the same relative significance across all grading contexts—for example, in departmentally heterogeneous samples, students with mean FYA of 3.8 may be assumed to be faring relatively better academically, on the average, within their respective departments than students with mean FYA averaging 3.0. Thus, even when considered without regard to department of enrollment (grading context), the FYA conveys useful

information regarding the academic progress and accomplishment of students. It is assumed that within-department grading standards were comparable for all students regardless of citizenship status.

Variables Related to English Proficiency

TOEFL Total Score

The TOEFL total score (reported on a scale with an effective range between 213 and 677) reflects performance on three separate measures: English vocabulary and reading comprehension (with items that are more general in nature and considerably less difficult than the vocabulary and reading comprehension items in the gre verbal test), knowledge of rules governing English language structure and written expression, and a measure of English language listening comprehension (ETS, 1983). In the present study, TOEFL total score was treated both as a potential predictor of FYA and as a basis for classifying students according to general levels of English proficiency.

TOEFL total score is moderately highly correlated with GRE verbal score in general samples of TOEFL/GRE test-takers. For a general sample of 3,808 TOEFL/GRE examinees tested during 1977-79 (Wilson, 1982b), the correlation between these two scores was approximately .70. Given this amount of overlap, the correlations of the respective measures with academic criteria might be expected to be approximately equal. There is evidence suggesting that this is a reasonable working hypothesis (see Sharon [1972] for evidence of parallel patterns of criterion-related validity for GRE-V and TOEFL; see also, Wilson [1985] for evidence of parallel levels of validity for scores on the verbal section of the Graduate Management Admission Test [GMAT] and TOEFL, respectively, in samples of foreign ESL MBA students).

TOEFL total scores were supplied for only about 60 percent of all foreign ESL students (N = 818). Only 68 of the 97 department-level samples supplied TOEFL scores for at least five students. Information was not available on departmental TOEFL requirements, bases for exemption, and so on. Several departments with relatively large numbers of students indicated that because of clerical problems they were not providing TOEFL scores. The uneven availability of TOEFL scores across departments results in some interpretive complications in analyses involving TOEFL scores.

TOEFL score estimated from GRE-Verbal score (TOEFL-EST). To provide an estimate of TOEFL performance (TOEFL-EST) for students for whom TOEFL total scores were not available, a regression equation for predicting TOEFL total score (T') from GRE verbal (V) score was derived using data for the 1977-79 GRE/TOEFL sample cited above (GRE-V, mean = 360, S.D. = 108; TOEFL Total mean = 559, S.D. = 63; r = .70): $T' = \text{TOEFL-EST(IMATED)} = .406\ (V) + 413$.

TOEFL/VERBAL. For exploratory purposes, a variable called TOEFL/VERBAL was created by imputing TOEFL-EST scores for individuals without TOEFL Total scores.

GRE Verbal Performance Relative to Quantitative Performance

The discrepancy between observed GRE-V score, V, and the predicted GRE-V score, V', that would be predicted from GRE-Q score using a regression equation based on data for U.S. examinees (called a Relative Verbal Performance Index or RVPI), may be thought of as indexing "degree of English proficiency deficit" in the verbal performance of foreign GRE examinees (Wilson, 1984a, 1986b).

For national contingents of foreign ESL examinees, mean RVPI values (a) are systematically negative, (b) are typically relatively large in absolute value, and (c) tend to vary with English-language background. For contingents of foreign ENL (English-native language) GRE examinees, on the other hand, mean RVPI values tend to vary around zero, as would be the case for samples of U.S. examinees.

The predicted GRE verbal score used in calculating the RVPI was computed using a regression equation based on data for U.S. mathematics and physical science examinees tested during 1981-82 (GRE-V, mean = 520, S.D. = 109; GRE-Q = 645, S.D. 104; estimated r = .50), described in detail elsewhere (Wilson, 1984a):

Predicted verbal = V' = .52 (Q) + 185, and
RVPI = V - V', where V = observed verbal score.

In the present study, the RVPI was treated as a potential moderator of GRE/FYA relationships.*

GRE Analytical Performance Relative to Quantitative Performance

Procedures analogous to those used in deriving the RVPI were employed to derive an index of the discrepancy between observed GRE analytical score, A, and the predicted analytical score, A' (the score that would be predicted for a U.S. examinee with a given quantitative score). This discrepancy index was called the Relative Analytical Performance Index, or RANPI. The analytical ability measure calls for relatively extensive verbal processing, albeit of a

* In a study involving GMAT verbal and quantitative scores for MBA students from 59 schools (Wilson, 1985), a comparably derived RVPI measure (reflecting GMAT-V score relative to GMAT-V score predicted from GMAT-Q using an equation for U.S. GMAT examinees) was analyzed as a potential moderator variable. When FSL MBA students were classified according to roughly the top, middle, and bottom thirds on the RVPI index, the GMAT/FYA correlations were found to be highest for examinees in the top third on the GMAT RVPI (assumed to have the least English proficiency deficit), lowest for those in the bottom third (assumed to have the greatest English proficiency deficit, and in between for students in the middle third (with medium English proficiency deficit). This finding indicated that the GMAT RVPI "moderated" GMAT/FYA relationships in samples of foreign ESL MBA students.

more specialized character than that involved in the GRE verbal test. The RANPI, which has not been used previously, was thought of as possibly indexing a somewhat different type of "English proficiency deficit" for the foreign ESL population than that indexed by the RVPI. In the present study RANPI was treated as a potential moderator of GRE/FYA relationships.

The equation specified below was used to obtain predicted analytical score (A'). The equation was based on data for the sample of 1981-82 U.S. examinees cited above, without regard to field of study: analytical mean = 520, standard deviation = 124; quantitative mean = 521, standard deviation = 132; estimated Q,A correlation = .65.

Predicted analytical = A' = .611(Q) + 202, and
RANPI = A - A', where A = observed analytical score.

Self-Reported English Language Communication Status

GRE background questions include, "Do you communicate better in English than in any other language?" For study purposes "Yes" = SR-BCE (self-reported better communication in English) status = 1; "No" and "no response" = 0. Foreign ESL examinees who report BCE status typically are from nonnative-English speaking countries with a strong English speaking tradition. They typically are not natively proficient in English but tend to be more proficient than their ESL counterparts who report better communication in a language other than English, as indicated by higher average scores on the TOEFL and the GRE Verbal Test, for example. However, the ESL-BCE examinees earn lower average verbal scores than foreign ENL (English native language) examinees (Wilson, 1982b).

Other Variables

The TOEFL total mean reported by ETS (1983) for all 1980-82 TOEFL-takers from a given country was ascribed to each student from that country in the present study sample. This variable, called TOEFL-LEVEL (see Tables 3 and 4 and related discussion), was thought of as reflecting differences in English-language and educational background associated with national-linguistic origin.

Information regarding sex, age in years (as of October 1982), and undergraduate origin (U.S. versus other location) was available for most students.

Performance on the Study Variables

Table 5 provides summary statistics (means and standard deviations) for the variables described in the preceding section for foreign ESL students, by broad academic area. For perspective, statistics are also shown for the total sample (N = 42) of foreign ENL (English native language) students.

Table 5

Summary Statistics for the ESL Sample on Selected Study Variables, by Broad Academic Areas, with Comparative Data for All ENL Students*

Variable	ESL sample				ENL sample
	Total	Quant total	Biosci	Soc Sci	Total
FYA (Mean)	3.45	3.49	3.18	3.44	3.55
(Standard deviation)	0.42	0.41	0.52	0.42	0.44
GRE Verbal	382.0	384.4	375.5	351.5	546.0
	105.6	106.1	88.1	103.3	131.4
GRE Analytical	485.6	492.1	450.5	414.5	591.7
	107.0	105.6	108.1	94.9	126.8
GRE Quantitative	684.3	698.4	595.8	541.3	684.5
	98.7	81.3	127.3	148.3	113.4
Relative Verbal Performance Index (RVPI)	-158.9	-163.8	-119.4	-114.9	5.0
	106.7	104.7	106.3	118.9	126.7
Relative Analytical Performance Index (RANPI)	-134.2	-136.2	-115.2	-117.9	-28.2
	92.4	91.7	94.2	99.0	106.7
TOEFL Total	566.9(a)	568.0(b)	549.7(c)	549.6(d)	N.A.
	44.7	44.7	43.0	40.7	
TOEFL Estimated from GRE Verbal	568.1	569.1	565.4	555.7	634.7
	42.9	43.1	35.8	42.0	53.3
TOEFLEVL (Country means ascribed to students)	510.2	510.7	504.0	506.8	N.A.
	28.7	29.0	20.4	28.2	
Self-Reported Better Communicator in English = 1	0.16	0.16	0.20	0.16	N.A.
Attended U.S. Undergraduate School = 1	0.16	0.16	0.22	0.14	0.12
Sex (M = 0, F = 1)	0.17	0.13	0.27	0.61	0.31
Age (in years, as of October 1982)	25.8	25.4	27.2	29.8	26.3
	4.0	3.7	4.5	5.3	6.4

* Means in first row for all variables, standard deviations in row 2 except for nominally coded (1,0) variables. Ns for groups were, from left to right, 1353, 1213, 55, 85, and 42, except for Age (total ESL N = 1286), Sex (N = 1307), and TOEFL Total: (a) N = 818, (b) N = 771, (c) N = 22, and (d) N = 25. N.A. indicates not applicable for English-native-language (ENL) students.

Diskette 546.84, working copy of 2, Document 2

First-Year Average Grade

For the total sample of foreign ESL students, mean FYA was 3.45; for those in quantitative fields, the mean was 3.49. For the foreign ENL students, mean FYA was 3.55. Such mean FYA levels are consistent with those reported by ETS for samples composed predominantly of U.S. citizens. The mean of 3.18 for students in bioscience departments is based on data for only 55 students from six departments.

- The GRE Validity Study Service (VSS) at ETS reported median first-year averages of 3.39 for 49 physical science departments, 3.53 for 33 bioscience departments, and 3.51 for 102 social science department-level samples, respectively (Burton & Turner, 1983). Although very few of the departments participating in this study were included among those that participated in the GRE VSS, the comparison suggests that these foreign ESL students, especially those in quantitative fields, probably earned first year grades comparable to those of their U.S. counterparts.

GRE General Test Performance

Several trends are noteworthy (see Appendix A for department-level data):

- The quantitative ability mean for all foreign ESL students and for those in quantitative departments was above the eightieth percentile for basic GRE reference groups (e.g., ETS, 1985, p. 17). For all major areas, including the social sciences, the basic pattern of abilities was the same: highest on quantitative, lowest on verbal, with analytical in between.

- Foreign ESL students and foreign ENL students had identical means on the GRE quantitative measure, but the verbal and analytical means of ESL students were much lower than those of ENL students.

- The large negative RVPI and RANPI means (-158.9 and -134.2, respectively) for foreign ESL students indicate average verbal and analytical performance markedly lower than would be expected for U.S. examinees with GRE quantitative scores averaging 684. Note that the verbal and analytical means for ENL students were quite consistent with expectation for U.S. examinees: mean RVPI = 5.0 and mean RANPI = -28.2, discrepancies that may be accounted for by sampling effects.

For perspective, the medians of the distributions of means of GRE verbal and quantitative scores for 57 physical science departments participating in the GRE VSS through June 1982 were verbal = 528 and quantitative = 687 (Burton and Turner, 1983, p. A-1) as compared to 384 and 698 for ESL students in the present study. The median analytical mean for GRE VSS participants was 588 (for scores on the pre-October 1981 analytical test only).

TOEFL Performance

About 60 percent of the total sample but proportionately fewer of those in bioscience and social science samples had TOEFL scores. The TOEFL total mean of 567 indicates that the students with TOEFL scores were highly selected in terms of the aspects of English proficiency measured by this test—the average student was at approximately the eighty-fourth percentile for all graduate-level TOEFL examinees (ETS, 1983).

The TOEFL-EST mean, 568, indicates the mean of the distribution of estimated TOEFL scores for students with the same GRE verbal mean as that of the present sample. The close agreement between the TOEFL and TOEFL-EST means indicates that students without TOEFL scores had GRE verbal scores comparable to those with TOEFL scores. This suggests that, on the average, students without TOEFL scores were roughly comparable to those with TOEFL scores in terms of aspects of "English proficiency" tapped by the GRE verbal measure. It is of incidental interest to note that the estimated TOEFL score for the ENL students was approximately 635, a value higher than the TOEFL mean attained by any contingent of U.S.-bound TOEFL-takers.

The high TOEFL mean for the foreign ESL sample reflects in part the fact that foreign ESL applicants typically are screened for "English proficiency." Presumably, those admitted are judged to have either (a) at least a minimally adequate level of proficiency in English to begin academic study full-time or part-time and/or (b) a reasonable likelihood of being able to acquire a minimally adequate level given a period of intensive ESL instruction.

However, the high TOEFL mean is also attributable in part to the fact that TOEFL examinees who take the GRE, on the average, are much more proficient in those aspects of English proficiency measured by the TOEFL than TOEFL examinees generally.

- For some 3,808 TOEFL/GRE examinees tested during 1977-79, for example, the TOEFL total mean was 559 (Wilson, 1982b) as compared to a mean of 508 for all graduate-level TOEFL examinees tested during 1980-82 (ETS, 1983). The TOEFL mean of 559 corresponds to the eightieth percentile rank in the distribution of scores for all graduate-level TOEFL examinees tested during 1980-82.

Results of surveys of institutions to determine score levels on the TOEFL that are associated with various types of admissions decisions (e.g., ETS, 1983), indicate that scores in the 550 range frequently are judged to be sufficient for unconditional admission to academic programs for ESL applicants judged to be academically qualified—that is, for beginning academic work without concomitant participation in intensive ESL instruction or other activities designed to improve English proficiency. Thus, it is possible that many if not most of the ESL students in the present sample may have surpassed a "threshold" level of proficiency required for academic functioning in an English-language environment—especially as students in fields that emphasize primarily quantitative abilities, internationally employed notations, specialized English vocabulary, and so on.

Other Variables

The TOEFL-LEVEL mean for the ESL sample was 510--close to the grand mean for all U.S.-bound TOEFL examinees as reported by ETS (1983). To reiterate, TOEFL-LEVEL was derived by ascribing to each student in the sample from a given country the TOEFL total mean of individuals from that country who took the TOEFL during 1980-82, as reported by ETS (1983).

About 16 percent of the ESL students reported BCE status (better communication in English than any other language), and the same percentage reported attending a U.S. undergraduate school.*

Some 17 percent of the students were female. Average age for all ESL students was 25.8 years; students in social sciences (primarily education) were considerably older (mean = 29.8 years), and those in biosciences were somewhat older (mean = 27.2 years) than those in the quantitative fields (mean = 25.4 years).

Intercorrelations of Selected Variables in the Total ESL Sample

Table 6 shows intercorrelations of designated study variables in the total ESL sample; intercorrelations of GRE scores and the two measures derived from GRE scores (RVPI and RANPI) are shown for the total foreign ENL sample.

- The correlation between GRE verbal and TOEFL in the sample of ESL students (.71) was almost identical with that (.70) found in the general sample of TOEFL/GRE examinees referred to earlier.

- The analytical measure and the RANPI as compared to the verbal measure and the RVPI had substantially lower correlations with TOEFL total score, TOEFL-LEVEL, and self-reported BCE status. This suggests that the type of ESL facility measured by the verbal test is not very similar to that measured by the analytical test. RVPI and RANPI were not strongly correlated (r = .41).

- Differences on other study variables between students with TOEFL scores and those without TOEFL scores are pointed up by coefficients for a variable called YES-TOEFL (1 = student had a score versus 0 = no score available). Students without TOEFL scores, on the average, were (a) quite similar to those with TOEFL scores in terms of mean score on the verbal and analytical measures (r = .00 and r = .03), (b) slightly lower, on the average, with respect to RVPI and RANPI (low negative correlations), but (c) somewhat higher in quanti- tative ability (r = .12).

- **The strongest single correlate of YES-TOEFL was location of undergraduate school (r = -.36).** Students not reporting a U.S. undergraduate school

*Both of these are underestimates of the actual percentages of students in the respective statuses since they represent the percentage of all students, including some for whom data were not available.

Table 6

Intercorrelations of Selected Study Variables in the Total ESL Sample

Total correlation matrix

Variable	V	A	Q	RVPI	RAN-PI	TOEFL total	YES TFL	TOEFL Level	SR-BCE	Age	Sex	U.S. Sch
GRE-V	—	.46	.22	.88	.39	.71	.00	.52	.26	-.18	-.07	-.03
GRE-A	.50	—	.51	.21	.83	.48	.03	.22	.03	-.35	-.05	-.04
GRE-Q	.30	.54	—	-.26	-.07	.25	.12	-.01	-.01	-.26	-.27	-.16
RVPI	.90	.26	-.15	—	.42	.60	-.06	.52	.26	-.05	.06	.05
RANPI	.39	.84	-.01	.41	—	.41	-.05	.25	.04	-.24	.12	.06
TOEFL Total						—	.00*	.50	.28	-.20	-.04	-.05
YES-TFL = 1*							—	.08	-.01	.02	-.06	-.36
TOEF-LEVEL								—	.35	-.13	-.04	-.12
SR-BCE = 1									—	-.11	-.03	.01
Age										—	-.01	-.20
Sex (M = 0, F = 1)											—	-.01
U.S. undergraduate school = 1												—

Note. N = 1353 for coefficients above diagonal except those involving Age (maximum N = 1286), Sex (N = 1307), U.S. undergraduate school (N = 1291), and TOEFL Total (N = 818). Entries below the diagonal are coefficients involving GRE scores, or variables derived using GRE scores, for the total sample of ENL (English-native-language) students (N = 42).

* YES-TFL = YES-TOEFL = TOEFL Total score available = 1; not available = 0.

were more likely to have TOEFL scores and vice versa. This is consistent with the fact that a record of successful performance in a U.S. school is frequent- ly accepted as evidence of adequate ESL proficiency—applicants with such a record may be exempted from taking TOEFL (ETS, 1983).

- It is noteworthy that ESL students who reported a U.S. undergraduate school tended to have lower GRE scores, especially GRE quantitative scores ($r = -.16$) than their counterparts who did not do so.

- Correlations in the .5 range for GRE verbal, RVPI, and TOEFL total, respectively, with TOEFL-LEVEL indicate that classification of students with respect to level of performance on these measures of verbal skills would result in substantial incidental sorting in terms of background variables that are linked to country of citizenship through the TOEFL-LEVEL score.

- The pattern of coefficients for self-reported BCE status with th verbal test measures was like that for TOEFL-LEVEL, but relationships were considerably weaker.

Section III: Study Methods and Procedures

Data were available for 97 small samples of ESL students predominantly from quantitative fields: engineering, mathematics, chemistry, physics, statistics, computer science, and economics. Complete GRE/FYA data were also available for small samples from five bioscience departments and six social science departments (five education departments and one political science department). Also available on a complete-data basis were TOEFL-LEVEL scores and scores on two nominally coded (1,0) background variables: U.S. undergraduate school versus other status, and self-reported better communication in English (SR-BCE) versus other status. TOEFL scores were available for at least five students in only 68 departments.

The present study was concerned primarily with assessing the typical levels of within-department GRE/FYA relationships for foreign ESL students generally and in various subgroups. Estimates of GRE validity coefficients based on single small samples such as those available for study would not be reliable. However, by pooling information from all the small samples, reliable estimates of the typical within-department correlation between GRE scores, FYA, and other variables can be obtained by analyzing interrelationships among departmentally standardized predictor and criterion scores in pooled samples of students from "similar" departments—that is, by analyzing pooled within-department data matrices. Coefficients in these matrices summarize basic trends in the distributions of department-level GRE/FYA coefficients.

It is convenient to standardize the predictor variables as well as the criterion variable when attention is focussed primarily on assessing trends in within-department predictor/criterion correlations (as in the present study) rather than on the development of operational predictive equations for use in particular departments. In the present study, data for U.S. students were not available for analysis.

General Methodological Rationale

Given common GRE/FYA (predictor/criterion) data sets for a relatively large number of small samples of individuals in "similar" graduate programs being offered by departments in different graduate schools, it is possible to draw meaningful inferences regarding general levels and patterns of GRE/FYA (predictor/criterion) correlation and the relative weighting of predictors by pooling data from all the "similar" settings. The estimates derived from the pooled data may be thought of as approximating population estimates.

One useful pooling procedure involves standardization of the predictor and criterion variables of interest within each department-level sample prior to pooling (see, for example, Wilson, 1979, 1982c, 1985, 1986a, 1986b). In this approach, original or raw scores on the respective variables are subjected to a z-scale transformation—raw scores are expressed as deviations from department means in department standard deviation units and are thus transformed to a common scale with mean of zero and standard deviation of unity in all samples.

- The intercorrelations among the departmentally-standardized variables for combined samples from several departments constitute a pooled, within-department correlation matrix. Validity (or other coefficients) based on pooled departmentally standardized data for several departments are equivalent to size-adjusted means of corresponding coefficients for the smaller, department-level samples. The use of size-adjusted averages of correlation coefficients to summarize results of comparable analyses that have been conducted in different samples is a well-established meta-analytic technique (e.g., Mosteller & Bush, 1954). The contribution of a particular departmental data set to pooled within-department validity estimation is a function of sample size.

- There is reason to believe that much of the variability in observed validity coefficients for common predictors and criteria across "similar" settings is due to statistical artifacts rather than setting-specific differences in "criterion content." For example, in an analysis of 726 law-school validity studies, Linn, Harnisch, & Dunbar (1981) estimated that about 70 percent of the variation in school-level validity coefficients across studies was accounted for by differences in sample standard deviations, estimated criterion reliability, and sample size, respectively. Similar findings have been reported for validity studies involving common selection tests and job performance criteria in occupational settings (for example, Pearlman, Schmidt, & Hunter, 1980). In the Cooperative Validity Studies Project (Wilson, 1979), in analyses involving over 40 departments in five disciplines the majority of department-level regression coefficients for GRE predictors were found not to differ significantly from the pooled within-department coefficients. Statistically significant deviations could be accounted for by clear outlier effects in small samples.

Thus, predictor/criterion correlation coefficients (validity coefficients) based on departmentally standardized data pooled across departments within various disciplines—that is, coefficients in pooled, within-department data matrices—may be thought of as approximating population values, around which department-level coefficients may be expected to vary, due primarily to statistical and sampling considerations (sample size, degree of selection, criterion reliability, and so on) rather than context-specific validity-related considerations such as real differences in the content of the criterion. For example, economics departments may differ with respect to the amount of course work in quantitative methodology typically required during the first year of graduate study.

Questions regarding the relative contribution of GRE scores and other variables in predictive composites may be addressed by applying multiple regression methods to pooled, within-department data matrices. Standard partial regression weights for GRE scores and other independent variables, and multiple correlation coefficients, for example, as well as simple correlations, based on departmentally standardized data for several departments, pooled within various disciplines, may be thought of as approximating population values. If one is concerned with developing regression equations applicable for small department-level samples, it has been found that when department-level regression coefficients are adjusted toward corresponding "population" values, the resulting equations generate more reliable predictions in subse-

quent samples than equations based solely on local data (see, for example, Braun and Jones, 1985).

The foregoing interpretative rationale for pooled, within-department estimates of correlation or regression coefficients rests on an assumption that the departments for which data are pooled are generally similar with respect to the nature of the academic tasks that students are required to complete. The programs of study offered by the departments for which data are pooled should require the exercise of generally similar patterns of skills, abilities, and so on--especially the types of skills and abilities represented by the predictor variables of interest.

For academic departments, a logical initial criterion for pooling is academic discipline or field. A more comprehensive pooling rationale would involve the reasonable assumption that tasks required of students by departments representing different fields or specializations within the same general academic area are generally similar in relative demand on, say, verbal as opposed to quantitative skills. Academic programs in English, history, political science, and so on, may be assumed a priori to make greater demands on verbal abilities than on mathematical or quantitative abilities; for programs in mathematics, chemistry, physics, and so on, the opposite is true.

This line of reasoning leads to the a priori expectation of higher validity for quantitative scores than for verbal scores in the primarily quantitative fields and the opposite pattern of validity for these measures in the primarily verbal fields. Observed differences in patterns of validity for verbal and quantitative scores are consistent with a priori expectation (see, for example, Willingham, 1974; Wilson, 1979, 1982c, 1986a, 1986b; Burton & Turner, 1983).* Less is known regarding patterns of predictive validity across disciplines for the analytical ability measure.

Application in the Present Study

Summary statistics (means, standard deviations, and intercorrelations) were computed for study variables within each of the 97 department-level samples of ESL students. The FYA criterion and original scores on the GRE and other continuous variables were z-scaled by department--the resulting department-level distributions had equal means (zero) and standard deviations (1.0). The original scores of foreign ENL students were z-scaled using parameters for the foreign ESL students.

* Braun and Jones (1985) reported that clustering departments empirically on the basis of patterns of sample means on the respective GRE measures provided a useful basis for aggregating data for the purpose of estimating regression coefficients for the GRE scores for members of a cluster. Clusters thus formed empirically may include samples from different disciplinary areas, but will tend to correspond basically to a priori clusters based on disciplinary affiliation that differ primarily along a verbal-relative-to-quantitative-emphasis dimension.

1. To evaluate levels and patterns of GRE/FYA correlations in general samples of ESL students, means of department-level validity coefficients (weighted by sample size, or size-adjusted, unless otherwise specified) were computed for each of the departments and classifications of departments designated below:

- o Chemical, civil, electrical, industrial, mechanical, engineering, and Engineering, total
- o Statistics, chemistry, physics, mathematics, computer science, and Mathematics and Physical Science, total
- o Economics
- o Quantitative, total (Engineering + Math/Science + Economics)
- o Biosciences (total)
- o Social sciences (total)

In view of the very limited representation of departments from bioscience and social science fields, primary emphasis was placed on analysis of data for the 86 departments from primarily quantitative fields.

2. To obtain general (population) estimates of the relative weighting of GRE scores in composites for predicting FYA, multiple regression analyses were conducted using pooled matrices of departmentally z-scaled FYA **(Zfya)** and GRE **(Zgre)** data for several groups of departments: (a) engineering (b) mathematics and physical sciences, (c) economics, (d) quantitative total, (e) biosciences, and (f) social sciences. The simple correlation between a standard composite of **Zgre** predictors, based on a regression equation developed for the total quantitative sample, and **Zfya** was computed for each department-level sample. Means of validity coefficients for the composite predictor were compared with means for the individual GRE predictors to assess the potential for incremental validity in a uniformly weighted general composite.

The TOEFL-LEVEL score was included as a supplemental predictor in certain of the multiple regression analyses to assess the possibility that this variable, which was thought of as reflecting differences in English-language background linked to country of citizenship, might have incremental validity when included with GRE scores.

3. Multiple regression analysis of departmentally z-scaled data, pooled for subgroups of students from quantitative departments, was employed to explore the possibility that GRE/FYA relationships might tend to vary across subgroups of foreign ESL students judged to differ in "level of ESL proficiency" and English-language background, as defined by:

a. score level on the Relative Verbal Performance Index (RVPI),
b. score level on the Relative Analytical Performance Index (RANPI),
c. TOEFL total-score level,

d. self-reported BCE status versus other status.

These analyses, incidentally, provided information regarding the average relative within-department standing of students in various subgroups as reflected in the means on z-scaled GRE (Zgre) scores and the z-scaled FYA (Zfya) criterion—that is, means on Zq, Za, Zv, and Zfya. Observed mean Zfya for the proficiency subgroups involved were compared with estimated or expected means (Z'fya) based on the departmentally z-scaled Zgre scores, using general regression equations based on pooled z-scaled data for all departments involved in the analysis. These comparisons provided a basis for tentative inferences regarding the comparative performance of various subgroups.

4. Are within-department predictor/criterion relationships stronger in groups that are homogeneous with respect to national origin than in samples that are heterogeneous with respect to national origin? To evaluate this question, coefficients reflecting the relationship between Zfya and a standard composite of Zq and Za scores were computed for samples of students classified by country of citizenship and world region and for samples classified by world region and type of quantitative department—engineering, math-science, economics—as well as for the combined sample of foreign ESL students from all quantitative departments.

Consistent with the primary objectives of the study, the foregoing analyses were designed to provide evidence regarding the typical levels of within-department relationships between the GRE and FYA variables for foreign ESL students and the effect on these relationships of introducing controls for "levels of English proficiency," variously defined, and/or country of citizenship. These analyses are described in detail in Sections IV through VI.

The study was not designed to address questions regarding the comparative academic performance of students from different countries or regional groups, or the extent to which level of academic performance was consistent with level of GRE performance. Limited analyses related to these questions were possible, however. These analyses and related findings are described in detail in Section VII of this report.

Section IV: GRE/FYA Relationships For Foreign ESL Students, by Academic Area

This section presents findings regarding GRE/FYA correlations for foreign ESL students in departments classified by field and in broader area classifications. Size-adjusted means of department-level GRE/FYA correlation coefficients are shown for various classifications. Results of regression analyses of z-scaled FYA (Zfya) on the z-scaled GRE (Zgre) predictors (Zq, Zq, Zv), using data aggregated for departments within broad academic areas, are presented. Also reported are trends in department-level correlations between FYA and three non-GRE variables: U.S. versus other undergraduate origin, self-reported better communication in English (BCE) versus other status, and TOEFL-LEVEL.

Interpretive Perspective

One of the principal questions implicitly at issue in this study is whether GRE/FYA correlations for samples of foreign ESL students are similar to those typically observed for samples of U.S. students. Since data for U.S. students were not collected for the departments in this study, it is useful to review briefly findings regarding typical levels and patterns of GRE validity for predicting academic criteria in samples composed exclusively or predominantly of U.S. students.

There is a substantial body of evidence regarding the relationship of scores on the well-established verbal and quantitative ability measures to performance in graduate study, typically measured by first-year average grades (for example, Willingham, 1974; Wilson, 1979, 1982c; Burton & Turner, 1983). Evidence regarding the predictive validity of the analytical ability measure introduced in October 1981, and not made operational until October 1985, is much more limited.

Evidence regarding typical levels and patterns of validity for the verbal and quantitative ability measures is based on cumulative findings for several hundred departments. Consistent with a priori expectation, GRE quantitative scores are more valid predictors than GRE verbal scores in quantitatively oriented fields such as chemistry, engineering, mathematics, and so on, while the opposite pattern typically holds for verbal fields such as English, history, sociology, political science, education, and so on.

- In two cooperative studies, each of which involved a total of 100 or more departments ranging from highly verbal to highly quantitative, median validity coefficients for verbal and quantitative scores, respectively, in the more verbal departments were .31 and .25 (Wilson, 1979) and .27 and .25 (Wilson, 1982c). For "quantitative" departments in the respective studies, typical verbal and quantitative coefficients were .20 and .31 in the earlier study and .18 and .28 in the later one. The earlier study involved scores on tests administered prior to October 1977, while the later study involved scores on the "restructured" test introduced in October 1977.

- For 118 social science departments that participated in the GRE Validity Study Service (VSS) at ETS through June 1982, the pooled within-department validity coefficients (size-adjusted averages) for GRE-V and GRE-Q were .26 and .23, respectively; for 56 mathematics and physical science departments, typical coefficients for GRE-V and GRE-Q were .12 and .27, respectively (Burton & Turner, 1983).

A comparable body of evidence is not yet available for the revised GRE analytical ability measure introduced in October 1981. Analytical scores have been found to be positively correlated with first-year graduate school grades (Kingston, 1985; Swinton, 1985) and with undergraduate grades (Wilson, 1984c, 1986a) in a variety of fields.

- Based on findings reported by Kingston (1985), for graduate departments sending post-September 1981 GRE General Test scores and FYA data to the GRE Validity Study Service (VSS) at ETS, the analytical score was more heavily weighted than the verbal score in predictive composites for graduate engineering and mathematical sciences departments.* However, questions regarding the incremental and/or differential validity of the analytical ability measure (for prediction of grades in quantitative as opposed to verbal areas of study, for example) for U.S. or other student groups remain unresolved.

GRE/FYA Validity Coefficients for Foreign ESL Samples

Table 7 shows size-adjusted means of department-level GRE/FYA correlations for foreign ESL students in designated groups of departments. For each grouping, the number of departments (samples) is shown, along with the total number of students on which the coefficient is based (see Appendix A for distributions of department-level coefficients and other department-level data).

* In evaluating Kingston's findings regarding the analytical ability measure, and in evaluating the coefficients obtained in the present study for this measure, it should be kept in mind that during the period in which the students involved were admitted to graduate school, test users were advised not to consider the analytical score in admitting students. Thus, the predictive value of the analytical score presumably was not affected by restriction of range due to direct selection, whereas restriction due to direct selection was a factor affecting the contribution of both the verbal and the quantitative scores. The fact that it was not used directly in selection theoretically should enhance the observed predictive validity of the analytical ability measure relative to that of the verbal and quantitative ability measures. Studies involving samples admitted after October 1985 will be needed to resolve questions regarding the incremental contribution of the GRE analytical ability measure to prediction under conditions in which all three GRE scores have been considered in selecting students.

Table 7

Simple Correlation of GRE General Test Scores with FYA: Weighted (Size Adjusted) Means of Department-Level Coefficients for Departments Grouped by Graduate Field and Area

Graduate Area	No. depts.	No. students	GRE-Q r	GRE-A r	GRE-V r
Engineering	**40**	**702**	**.289**	**.278**	**.114**
Chemical	8	80	.315	.362	.173
Civil	8	162	.163	.088	-.047
Electrical	10	256	.296	.340	.185
Industrial	6	89	.400	.342	.165
Mechanical	8	115	.347	.300	.104
Math/Science	**36***	**373**	**.351**	**.268**	**.023**
Statistics	5	54	.330	.425	.189
Chemistry	10	111	.353	.190	.012
Mathematics	4	39	.602	.266	-.013
Physics	6	51	.452	.452	.011
Computer Sci	10	113	.249	.220	-.013
Economics	**10**	**138**	**.313**	**.282**	**.210**
All Quantitative	**86***	**1213**	**.311**	**.275**	**.097**
Bioscience	**6**	**55**	**.061**	**.081**	**-.023**
Soc Sci	**5**	**85**	**.116**	**.184**	**.253**

Note. The coefficients shown are "size adjusted" averages of department-level coefficients (i.e., weighted according to sample size) for samples of five or more nonnative English-speaking students.

* Includes data for one applied mathematics department (N = 5).

Table Diskette 546.84 Doc. 18 page 1

- o For foreign ESL students in all subgroups of "quantitative" departments, GRE quantitative scores and GRE analytical scores were more highly correlated with FYA than were GRE verbal scores.

- o The basic pattern of GRE validity across the respective types of quantitatively oriented fields is suggested by coefficients based on the total quantitative sample of 1,213 students, namely, .311, .275, and .097 for quantitative, analytical, and verbal scores, respectively, notwithstanding the fact that coefficients for analytical scores were higher than those for quantitative scores in several fields.

- o The pattern of comparatively stronger validity of the analytical ability scores relative to the validity of the verbal scores in these ESL samples is consistent with Kingston's (1985) findings for samples from which ESL students (U.S. as well as non-U.S. citizens) were excluded.

For the small sample of 85 students from five social science departments (four from education, one from political science), verbal scores were most valid, and quantitative scores were least valid: coefficients for verbal, analytical, and quantitative scores were, respectively, .253, .184, and .116. The pattern of higher validity for verbal than for quantitative scores in this ESL sample in which education students predominated is consistent with that reported by the GRE VSS for 17 education samples (.26 and .19 for verbal and quantitative respectively—Burton & Turner, 1983).

For the six bioscience samples with a total of 55 students, coefficients for quantitative scores and analytical scores were positive but atypically low and of about the same magnitude (.061 as compared to .081), while the verbal coefficient was anomalously negative.

The GRE/FYA correlations in Table 7, except for those obtained in the limited bioscience sample, appear to be comparable to coefficients that have been found for U.S. students, as reviewed at the beginning of this section.

Multiple Regression Results for Foreign ESL Students by Academic Area

Table 8 shows selected findings of multiple regression analyses based on departmentally standardized data aggregated for broader groupings of departments: engineering, mathematics and physical sciences, economics, all quantitative, biosciences, and social sciences, respectively.

These results indicate the limited contribution of verbal scores to prediction of FYA in the primarily quantitative fields, with the possible exception of economics.

The patterns of regression coefficients tended to be quite similar for the respective quantitative areas. Coefficients based on aggregated data for the engineering, math/science, and economics departments were similar to those for the combined quantitative-department sample of 1,213 foreign ESL students.

Table 8

Regression Results for ESL Students Using Departmentally Standardized Data Pooled by Graduate Major Area

Field/Area	N	Correlation with FYA			Beta weight*			R
		Q	A	V	Q	A	V	
Engineering	702	.29	.28	.11	.21	.19	.01	.33
Math/Science	373	.35	.27	.03	.29	.17	-.08**	.38
Economics	138	.32	.28	.22	.25	.15	.12	.38
All quant	1213	.31	.27	.10	.24	.18	-.01**	.35
Bioscience	55	.06	.08	-.02	.03	.09	-.02	.10
Social Sci	85	.12	.18	.25	.05	.08	.21	.27

* Standard partial regression coefficient. Underscored coefficients are statistically significant at $p < .05$.

** Negative weight indicates suppression; note positive simple correlation.

Validity of a Standard GRE Predictive Composite for Quantitative Departments

The similarity in regression results for engineering, math-science, and economics departments suggested the potential utility of a standard GRE predictive composite including only quantitative and analytical ability scores, with weights specified by regression results for the combined sample of students from all quantitative departments.

A composite of departmentally z-scaled quantitative and analytical ability scores (Zq and Za) weighted to predict z-scaled FYA in the combined quantitative sample was computed for each student: .238 Zq + .176 Za. The simple correlation between this standard composite and the criterion was determined for each department-level sample.

Table 9 shows the means (size-adjusted) of the department-level coefficients for this standard composite (last column of table) for quantitative departments classified by field and major area. Mean coefficients for the individual GRE predictors (from Table 7) are included for perspective.

The results in Table 9 indicate that the standard composite had general validity for predicting relative within-department standing for departments in the three general quantitative areas: engineering, math/science, and economics. For these three areas, coefficients for the standard composite were somewhat higher than coefficients for the highest single predictor (typically quantitative ability). Observed differences between coefficient for the composite predictor and that for the best single predictor may be thought of as reflecting reasonable estimates of the amount of incremental validity involved in using a composite of two GRE scores. (Again, it is important to recall that the analytical score probably was not used directly in selection).

Simple Correlation of Selected Non-GRE Variables with FYA

Table 10 shows size-adjusted coefficients summarizing trends across departments in the relative academic performance of students who (a) reported BCE status (better communication in English than in any other language) versus other status and (b) reported attending a U.S. undergraduate school versus other status. **Positive coefficients indicate higher mean FYA for those with BCE status and for those reporting a U.S. undergraduate school.** Size-adjusted coefficients reflecting trends in the relationship between TOEFL-LEVEL scores and FYA are also shown in the table. The number of departments involved in the respective analyses varied; in several departments, no student reported BCE status or no student reported a U.S. undergraduate school.

The variable called TOEFL-LEVEL was essentially unrelated to FYA in the samples studied. In analyses not reported in Table 10, it was found that TOEFL-LEVEL did not have incremental validity when included in a battery with the GRE predictors.

Table 9

Validity Coefficients for a Standard Composite of GRE Quantitative and Analytical Scores versus Coefficients for Individual GRE Scores: Size-Adjusted Averages of Department-Level Coefficients for Quantitative Departments Grouped by Fields and Major Areas

Graduate Area	No. depts.	No. students	GRE-Q r	GRE-A r	GRE-V r	Composite* r
Engineering	**40**	**702**	**.289**	**.278**	**.114**	**.330**
Chemical	8	80	.315	.362	.173	.402
Civil	8	162	.163	.088	-.047	.155
Electrical	10	256	.296	.340	.185	.372
Industrial	6	89	.400	.342	.165	.430
Mechanical	8	115	.347	.300	.104	.354
Math/Science	**36****	**373**	**.351**	**.268**	**.023**	**.370**
Statistics	5	54	.330	.425	.189	.422
Chemistry	10	111	.353	.190	.012	.340
Mathematics	4	39	.602	.266	-.013	.521
Physics	6	51	.452	.452	.011	.522
Computer Sci	10	113	.249	.220	-.013	.283
Economics	**10**	**130**	**.313**	**.282**	**.210**	**.370**
All Quantitative	**86****	**1213**	**.311**	**.275**	**.097**	**.347**

Note. The coefficients shown are "size adjusted" averages of department-level coefficients (i.e., weighted according to sample size) for samples of five or more foreign English-second-language (ESL) students.

* The entries in this column are size-adjusted averages of department-level simple correlation coefficients between Z(fya) and a standard composite of departmentally z-scaled GRE quantitative, Z(q), scores and analytical, Z(a), scores, weighted according to results of the regression of Z(fya) on Z(q) and Z(a) in the combined sample of students (N = 1,213) from all quantitative departments: Predicted Z(fya) = Z'(fya) = .238 Z(q) + .176 Z(a).

** Includes one applied mathematics department (N = 5) for which data are not shown separately.

Table 10

Simple Correlation of Selected Background Variables with FYA: Size-Adjusted Means of Department-Level Coefficients for Departments Grouped by Graduate Field and Area

Field	TOEFL-LEVEL			Self-Reported BCE = 1			U.S. Undergraduate school = 1		
	No. depts.	(N)	r	No. depts.	(N)	r	No. depts.	(N)	r
Engineering	**40**	**702**	**.035**	**37**	**679**	**-.005**	**37**	**506**	**-.226**
Chemical	8	80	-.073	6	66	-.090	8	80	-.116
Civil	8	162	-.133	8	162	-.098	8	162	-.204
Electrical	10	256	.152	10	256	.048	10	256	-.236
Industrial	6	89	.111	5	80	.122	2	30	-.285
Mechanical	8	115	.028	8	115	-.138	5	75	-.300
Math/Science	**36***	**373**	**-.078**	**29***	**301**	**-.094**	**23***	**243**	**-.210**
Statistics	5	54	-.052	5	54	.042	3	36	-.030
Chemistry	10	111	-.117	8	90	-.202	5	50	-.420
Mathematics	4	39	.037	2	26	-.026	3	34	-.288
Physics	6	51	-.135	4	27	-.155	4	38	-.188
Computer Sci	10	113	-.039	9	104	-.017	7	85	-.093
Economics	**10**	**138**	**.030**	**6**	**98**	**-.207**	**6**	**99**	**-.171**
All Quantitative	**86***	**1213**	**.001**	**71**	**1078**	**-.046**	**65**	**948**	**-.219**
Bioscience	**6**	**55**	**-.017**	**4**	**41**	**.115**	**5**	**49**	**-.029**
Soc Sci	**5**	**85**	**-.017**	**5**	**85**	**.101**	**5**	**85**	**-.142**

Note. The coefficients shown are "size adjusted" averages of department-level coefficients (i.e., weighted according to sample size) for samples of five or more nonnative English-speaking students. **TOEFL-LEVEL** = TOEFL means of U.S.-bound TOEFL examinees by country ascribed to citizens of the respective countries in the present sample. **Self-reported BCE** = better communication in English than in any other language versus other status. **U.S. undergraduate school** = designated a U.S. school versus other.

* Includes one applied mathematics department (N = 5) for which data are not shown separately.

- o Size-adjusted means of TOEFL-LEVEL/FYA coefficients were .035 (all engineering samples), -.078 (mathematics and physical sciences), .030 (economics), -.001 (all quantitative), -.027 (biosciences), and -.017 (social sciences).

- o These coefficients indicate no systematic tendencies across departments for academic standing to vary with national origin as indexed by historic country-level means on the TOEFL. In evaluating this outcome, it should be recognized (a) that only a very limited number of countries could be represented in the very small department-level samples, (b) that the range of background differences indexed by this variable was correspondingly restricted, and (c) that the particular mix of countries was not constant over departments.

Coefficients for BCE status indicate that there was no systematic tendency for students reporting better communication in English to earn better grades than others. In fact, for quantitative fields, the opposite pattern tended to be slightly more prevalent as indicated by the negative coefficients. BCE coefficients were positive, but low, for biosciences and social sciences.

The size-adjusted mean correlation for U.S. versus other undergraduate school with FYA was negative for every academic classification. This indicates a relatively strong tendency across departments for foreign ESL students who reported attending a U.S. undergraduate school to earn lower grades, on the average, than their counterparts who did not do so.

- o In evaluating this finding, it is useful to recall (from Table 6) that students with U.S. undergraduate origins had lower average scores on all three GRE variables, especially on GRE quantitative, than their counterparts who completed their undergraduate studies elsewhere.

Section V: Zgre/Zfya Relationships for Subgroups Differing in Performance on Selected English-Proficiency-Related Measures: Pooled Z-scaled Data for Quantitative Departments

This section presents findings of regression analyses based on pooled z-scaled data for subgroups of foreign ESL students from quantitative departments only. Analyses were not conducted for students in bioscience and social science departments.

The analyses were concerned with trends in Zgre/Zfya relationships across subgroups differing in relative levels of "ESL proficiency" as measured by (a) the RVPI (relative verbal performance index), (b) the RANPI (relative analytical performance index), (c) self-reported better communication in English or BCE status versus other status, and (d) scores on the TOEFL. The question generally at issue was whether validity coefficients would tend to be higher for subgroups with higher levels of English proficiency than for those with lower levels, as indexed by the respective measures.

With regard to the three test variables, the comparatively low intercorrelations reported previously (Table 6) indicate that the aspects of "ESL proficiency" involved in processing the verbal content of the analytical ability items are not very similar to those involved in processing GRE verbal or TOEFL items.

Procedures Involved in Defining Subgroups

For the RVPI, the RANPI, and the TOEFL, respectively, score levels for "high," "medium," and "low" subgroups were set in such a way that if the respective score distributions were normal, about one-third of the students would be in each group. This was approximately true for RANPI and TOEFL. However, the distribution of RVPI scores was skewed positively—about 47 percent of all students were in the "low" category as compared to 23 percent in the "high" category; all department-level samples were heterogeneous with respect to scores on the RVPI and RANPI measures and some representation in each score-level was assumed for the majority of the samples.

With respect to the self-reported BCE status variable, fourteen of the 86 department-level samples included no student who reported better communication in English. This was taken into account by forming three subgroups for analysis: an "English better" subgroup and an "other language better" subgroup for students from departments with at least one BCE student, plus a third subgroup including "other language better" students from 14 departments with only non-BCE students enrolled.

TOEFL scores were unavailable for a substantial proportion of the students. Three TOEFL-score subgroups were formed for 769 students from 75 departments represented by at least five students with GRE/FYA scores, at least one of whom also had a TOEFL score. For 10 departments in which at least one but fewer than five students had TOEFL scores, the available TOEFL scores were

z-scaled using parameters for TOEFL/VERBAL (TOEFL score or TOEFL score estimated from GRE verbal score). A total of 198 students from the 75 departments had GRE and FYA data, but no TOEFL scores.

Results

Table 12 summarizes the principal findings of the regression analyses: means of **Zgre** scores and mean **Zfya** for each subroup, simple **Zgre/Zfya** correlations based on pooled z-scaled data for the subgroups, and multiple correlations for **Zgre** composites, one involving only quantitative **(Zq)** and analytical **(Za)** scores and the other including all three GRE scores **(Zq, Za, Zv)**. **Ztoefl/Zfya** coefficients as well as **Zv/Zfya** coefficients are shown for the TOEFL sample. Findings based on the total quantitative sample are also shown.

Overall, the findings do not indicate a consistent tendency for **Zgre/Zfya** correlations to be "moderated" by "level of ESL proficiency" as defined by the variables under consideration.

- o Levels of **Zgre/Zfya** relationships did not tend to vary directly with levels of proficiency defined in terms of relative verbal performance, relative analytical performance, or self-reported English communication status. In analyses involving TOEFL as the classificatory variable, only for the subgroup with TOEFL scores of 585 plus (at about the **eighty-seventh** percentile for all graduate-level TOEFL examinees) were **Zgre/Zfya** coefficients (**.37, .36, and .18 for Zq, Za, and Zv, respectively**) noticeably higher than those for the total quantitative sample (**.31, .27, and .10, respectively**); corresponding multiple correlations were **.44** as compared to **.35.**

Interpretation of correlational results for classifications based on TOEFL total score is complicated by the fact that differences in **Zgre/Zfya** correlation associated with "availability versus nonavailability of TOEFL total score" were more pronounced than those associated with differences in score level among students with TOEFL scores.

- o The multiple correlation for **Zgre** composites in the "No TOEFL" subgroup was only **R = .22,** as compared to multiple correlations of **.36, .34, and .44** for "Yes TOEFL" students in lower, medium, and higher TOEFL-score classifications, respectively.

There is no ready explanation for this unanticipated pattern of findings regarding TOEFL. It is difficult to attribute the observed differences in level of **Zgre/Zfya** correlations between the "Yes TOEFL" and "No TOEFL" subgroups to English-proficiency related factors. For example, the mean relative within-department standing of the "No TOEFL" subgroup on GRE verbal was about the same as that for students in the middle TOEFL-score range—mean **Zfya** = -0.09 as compared to -0.11). This suggests that the "No TOEFL" students (presumably screened for ESL proficiency by other means) were roughly comparable to "Yes TOEFL" students in terms of the type of "ESL proficiency" measured by GRE verbal items. The "No TOEFL" and "Yes TOEFL" subgroups were also roughly comparable with respect to relative standing on analytical ability, but the

Table 11

GRE/FYA Relationships for Subgroups Defined by Relative Standing on Selected ESL-Proficiency-Related Test or Self-Report Variables in Analyses Using Departmentally Standardized Predictor/Criterion Data Pooled Across Quantitative Departments

Variable #	(N)	Zfya++ Actual	Zfya++ (Est)	Z-scaled means Zgre-q	Z-scaled means Zgre-a	Z-scaled means Zgre-v	GRE/FYA correlation Zgre-q	GRE/FYA correlation Zgre-a	GRE/FYA correlation Zgre-v	Multiple correlation Zq,Za	Multiple correlation Zq,Za,Zv
Relative Verbal Performance Index											
RVPI high	289	0.01	(0.03)	-0.15	0.36	1.07	.28	.23	.15	.30	.30*
RVPI med	352	-0.04	(-0.07)	-0.20	-0.15	-0.02	.29	.32	.16	.36	.36
RVPI low	572	0.02	(0.01)	0.20	-0.08	-0.53	.35	.28	.11	.37	.37*
Relative Analytical Performance Index											
RANPI high	391	0.14	(0.17)	-0.02	0.98	0.46	.34	.30	.11	.36	.36
RANPI medium	499	0.02	(-0.01)	0.00	-0.08	-0.04	.30	.25	.02	.30	.30*
RANPI low	353	-0.18	(-0.15)	0.02	-0.90	-0.30	.30	.24	.10	.32	.32
Self-reported BCE Status											
English better (a)	198	-0.09	(-0.01)	-0.06	0.04	0.51	.35	.27	.10	.38	.38*
Other better (a)	883	0.02	(0.00)	0.02	-0.01	-0.11	.30	.27	.12	.34	.34
Other better (b)	132	0.00	(0.00)	0.00	0.00	0.00	.34	.33	.04	.40	.41*
TOEFL Total Score ##											
585+	252	0.09	(0.11)	0.17	0.40	0.75	.37	.36	.18(.15)	.43	.44(.44)
550-584	265	0.15	(0.02)	0.11	-0.02	-0.11	.25	.31	.12(.11)	.34	.34(.34)
Less than 550	252	-0.07	(-0.07)	-0.07	-0.31	-0.55	.36	.23	.01(.12)	.36	.36(.36)
No TOEFL	196	-0.22	(-0.08)	-0.27	-0.07	-0.09	.21	.12	.06	.22	.22*
Total	1213		0.00	0.00	0.00	0.00	.31	.27	.10	.35	.35*

++ Initial entry is the observed mean. The entry in parentheses is z-scaled mean estimated from Zgre-q and Zgre-a, using a general regression equation based on data for the total quantitative sample (N = 1,213): **Zfya(est) = .238 Zgre-q + .176 Zgre-a.**

 # Relative Verbal Performance Index: The discrepancy between the observed GRE verbal score and GRE verbal score predicted from quantitative score using a regression equation based on data for U.S. GRE examinees. **Relative Analytical Performance Index:** A comparable discrepancy index involving observed analytical ability score and analytical score predicted from quantitative score using a U.S. regression equation. **Self-Reported BCE status:** Self-reported relative communicative ability, either better in English or better in some other language; **(a) categories** include only students from departments in which at least one student reported BCE status; (b) includes data for departments in which all students reported better communication in a language other than English.

The TOEFL analysis was based on pooled data for 967 students from 75 departments with five or more students with complete GRE/FYA data, at least one of whom also had a TOEFL total score. A total of 769 students from these departments had GRE, FYA, and TOEFL scores, and 198 had GRE and FYA only. Where TOEFL N was less than 5 for a department, TOEFL was z-scaled with reference to parameters for TOEFL/VERBAL, a variable defined as either TOEFL total score (when available) or estimated TOEFL (from GRE-Verbal) for students without TOEFL. Coefficients in parentheses indicate either simple Ztoefl/Zfya correlation or the multiple correlation obtained when Ztoefl was substituted for Zgre-verbal in the analysis.

* Weight for Zgre-verbal is negative (suppression) in this composite.

"No TOEFL" subgroup was lower in terms of quantitative ability (by -0.27 within-department standard deviations, on the average, and correspondingly lower in mean FYA (by -0.22 standard units).

For the respective subgroups, as in the total quantitative sample, adding the analytical score (Za) to the GRE quantitative score (Zq) yielded some increment in validity, but including the GRE verbal score (Zv) or the TOEFL total score (Ztoefl) did not yield any increments in validity.

Comparative Performance of Subgroups

By virtue of the process of z-scaling, means of each department-level sample, and for all aggregations of data involving intact department-level samples of foreign ESL students, were 0.00. For subgroups within a department, or for aggregations of data involving selected members of departmental samples, the means of z-scaled variables indicate, in standard units, the average deviation of subgroup members from the means of their respective departments. For example, the Zfya means in the RVPI analysis were 0.01, -0.04, and 0.02 for members of the high, medium, and low RVPI subgroups, respectively. The "medium RVPI" subgroup mean (-0.04) indicates an average FYA that was .04 standard units below departmental FYA means; other z-scaled means may be interpreted in the same way.

Predicted or expected average standing on the z-scaled FYA (Zfya) criterion was computed for each subgroup based on z-scaled quantitative and z-scaled analytical scores only, using the total quantitative sample regression equation: Predicted Zfya = Z'fya = .238 Zq + .176 Za = 0.00. These estimated means are shown in parentheses following the observed mean Zfya for each subgroup.

Overall, the average relative within-department academic standing of the respective subgroups was generally consistent with expectation based on their relative standing on GRE quantitative and analytical ability.

The limited predictive role for GRE verbal is pointed up by the fact that higher and lower "ESL proficiency" groups in every analysis were sharply differentiated in terms of mean verbal score (mean Zv), but the direction of mean Zfya differences was not necessarily consistent with the direction of the verbal score differences. This pattern is epitomized by results of the RVPI analysis.

- Students in the high RVPI group and students in the low RVPI group had z-scaled FYAs averaging 0.01 and 0.02, respectively; their FYAs were typical for their respective departments. However, they differed by about 1.6 standard deviations, on the average, with respect to z-scaled GRE verbal scores: high RVPI students averaged 1.07 standard units above departmental verbal means while low RVPI students, typically, were 0.53 standard units below departmental GRE-verbal means.

Section VI: Correlation of a Standard Zgre Predictive Composite with Zfya for Foreign ESL Students Classified by National Origin and Graduate Major Area—Data for Quantitative Departments

Results reported in preceding sections indicate substantial general validity (across types of quantitative departments) for the standard composite of quantitative and analytical scores specified by the regression of Zfya on Zq and Za in the combined quantitative sample—namely, .238 Zq + .176 Za.

This section presents data on the correlation between this general predictive composite and Zfya (a) in samples of students, (N > 9 only) from all quantitative departments combined, classified by country of citizenship and by regions defined for the study, and (b) in samples classified by both region and graduate major area—that is, engineering, math-science, and economics. These classifications introduced some control for linguistic-cultural-educational background variables associated with national origin, as well as for type of quantitative major.

The analyses were concerned in part with assessing the consistency of the relationship between a standard predictive composite, weighted to maximize the multiple correlation of Zq and Za with Zfya in the general quantitative sample, across the subgroups defined wholly or in part by national origin. Also of interest was the question of whether the correlation between Z'fya (the predicted z-scaled FYA) and Zfya (the observed z-scaled value) would tend to be stronger in subgroups that were homogeneous with respect to national-linguistic-cultural background than in the heterogeneous general sample.

Table 12 shows simple correlations between Z'fya and Zfya for larger national contingents within designated regions and for all students from countries in the respective regions. These analyses are based on pooled data for all quantitative departments. Table 13 shows comparable coefficients for students classified by region and by type of quantitative major.

As noted earlier, the subgrouping within regions (for example, Europe I versus Europe II, Asia I versus Asia II) was designed to take into account differences in characteristic patterns of English language acquisition and usage and/or in the typical TOEFL scores of contingents of U.S.-bound students from the respective countries. No country from Africa I or America I was represented by as many as 10 students.

Students from Category I countries as compared to those from countries in Category II are assumed typically to have "richer English language backgrounds." The findings in Table 12 and Table 13 indicate that the introduction of control for national-linguistic origin did not have a systematic influence on the level of relationship between Z'fya and Zfya.

- Coefficients for Category I national and regional contingents in Table 12 appear to be comparable to those for Category II contingents.

Table 12

Correlation of a Composite of Z-Scaled GRE Quantitative (Zq) and Analytical (Za) Scores with Z-Scaled FYA (Zfya) for Subgroups Defined by Country and Region: All Quantitative Fields

Country/Region	N	r
France	13	.39
West Germany	17	.41
Spain	13	.11
Europe I	62	.36
Greece	40	.54
Turkey	17	-.24
Europe II	78	.42
Lebanon	12	.15
Iran	61	.28
Mideast	98	.26
Africa I	17	.65
Egypt	13	.60
Africa II	33	.48
America I	6	.29
Chile	16	.39
Colombia	13	.43
Mexico	15	.35
America II	86	.32
India	176	.37
Asia I	188	.34
Hong Kong	31	.40
Japan	40	.32
Korea	146	.50
Malaysia	21	.23
Pakistan	21	.09
Peoples' Rep	41	.49
Taiwan	294	.19
Thailand	20	.30
Asia II	643	.34
All students	1213	.35

Note. The predictive composite was specified by .238 (Zq) + .176 (Za), with weights based on the regression of Zfya on Zq and Za in the pooled matrix of departmentally z-scaled data for students from all quantitative departments. Coefficients tabled indicate the simple correlation between the predictive composite and Zfya in the subgroups designated.

Coefficients are shown by country only for countries represented by 10 or more students in the total quantitative sample; regional coefficients include data for all students from a region. Thus, for example, the total of 62 students from Europe I included 17 from West Germany, 13 from France, and 13 from Spain, plus 19 from other countries in the region. (See Table 4 for complete enumeration of countries in regions.)

Table 13

Correlation of a Composite of Z-Scaled GRE Quantitative (Zq) and Analytical (Za) Scores with Z-Scaled FYA (Zfya) for Subgroups Defined by Region and Major Graduate Area

Region	N	Engineering r	N	Math-Science r	N	Economics r	N	All Quant r
Europe I	29	.35	18	.72	15	.09	62	.36
Europe II	49	.47	20	.47	9	.07	78	.42
Mideast	76	.22	20	.45	2	-	98	.26
Africa I	10	.69	5	.79	2	-	17	.65
Africa II	16	.39	9	.64	8	.49	33	.48
America I	0	-	3	-	3	-	6	.29
America II	35	.29	27	.35	24	.30	86	.32
Asia I	112	.35	68	.29	8	.60	188	.34
Asia II	374	.31	203	.34	66	.51	643	.34
All Regions	702	.33	373	.37	138	.37	1213	.35

Note. The predictive composite was specified by .238 (Zq) + .176 (Za) = 0.0, a regression equation based on the pooled matrix of departmentally z-scaled data for 1,213 students from all quantitative fields. The coefficients tabled represent the simple correlation between this composite and Zfya for the designated subgroups.

- For 22 subgroups (by region and type of quantitative major) with coefficients in Table 13, the median coefficient was r = .37.

- For the 20 country contingents shown separately in Table 12, the median validity coefficient for the standard composite was r = .36, and for the nine regional contingents the median was r = .34, compared to the general sample coefficient of r = .35.

The results in Tables 12 and 13 indicate that the correlation between the z-scaled criterion and the standard composite of z-scaled predictors tended to be relatively consistent for subgroups defined in terms of national origin and/or academic area. Correlations did not tend to be higher for subgroups that were homogeneous than for those that were heterogeneous with respect to national origin.

Section VII: An Exploratory Analysis of the Comparative Performance of Regional Subgroups on FYA and GRE Variables— Students from All Quantitative Departments

Consistent with the primary objectives of this study, the analyses reported in the preceding sections were designed to assess the patterns and levels of correlation between the FYA criterion and the GRE predictors and the possibility that GRE/FYA relationships might be affected by introducing controls for selected English-proficiency-related test and background variables or for national origin.

The average levels of within-department GRE/FYA correlations tended to be similar for samples of foreign ESL students from different types of quantitative departments—engineering, math/science, economics. Based on analysis of departmentally z-scaled data pooled across various combinations of quantitative departments, the level of predictor/criterion correlation was not influenced by introducing controls for level of "English proficiency," variously indexed, or for national origin.

This section describes a supplementary, exploratory analysis of (a) differences in the average academic performance of students in the regional subgroups defined for the study and (b) the extent to which subgroup differences in average academic performance tended to correspond with observed differences in average GRE performance. The analysis was based on data for the combined sample of students from quantitative departments.

An analysis based on the departmentally z-scaled FYA and GRE variables (**Zfya** and **Zgre**) provided information regarding the average standing of members of regional subgroups on these variables within their respective departments of enrollment. A parallel analysis based on FYA (as computed and reported by departments) and GRE scores (on the 200-800 scale used for reporting) provided information regarding the average standing of members of the regional subgroup on these variables without regard to their departments of enrollment.

In deciding upon this approach, the following factors were taken into account:

- The regional subgroups differed markedly in size (from N = 17 for Africa I, to N = 643 for Asia II). The regional mix in many of the individual departmental samples, median N = 12, could not be representative of the regional mix in the total sample.
- Moreover, members of the respective regional subgroups were not necessarily enrolled in comparably "representative" arrays of departments with respect to (a) field and/or major area (engineering versus economics or engineering versus math/ science, for example), (b) degree of selectivity (level of GRE scores), (c) grading standards (level of grades awarded relative to level of GRE scores), and so on.

In such circumstances, observed regional-group differences in average within-department standing will reflect to some extent factors associated with the nonrandom distribution of regional members among the various departments. Accordingly, it was considered important to analyze not only the relative within-department standing of subgroups but also their relative standing without regard to their departments of enrollment. It was reasoned that parallel findings would permit stronger conclusions about regional differences in academic performance than would be possible if comparisons were based solely on z-scaled data.

The usefulness of this approach is contingent, in part, on the degree of comparability between the regression of FYA on GRE quantitative and analytical ability scaled scores for all ESL students without regard to their departments of enrollment, and the pooled within-department regression results for the same combined sample (that is, the regression of Zfya on Zgre variables).

To make a determination regarding this question, a multiple regression analysis based on the FYA as reported by departments and GRE variables on the familiar 200-800 scale was conducted for the total sample (N = 1,213) of ESL students. The results obtained closely paralleled results of the regression of Zfya on Zgre scores in this same sample.* For example:

- In the total sample analysis, simple correlations of GRE Q, A, and V with FYA were .34, .28, and .11, respectively, as compared to .31, .28, and .10 for the pooled within-department analysis.

- Only quantitative and analytical scores had significant weights when FYA was regressed on Q, A, and V in the total sample analysis; standard partial regression coefficients for Q and A were .268 and .144, respectively, in the analysis without regard to department of enrollment, as compared to .238 and .176 for Zq and Za in the pooled within-department analysis.

- For the total sample analysis, R = .36 for the Q,A composite, as compared to R = .35 for the Zq,Za composite.

For the z-scaled within-department data set, means were computed, by region, for Zfya, Zq, Za, Zv, and Z'fya, where Z'fya = predicted Zfya = .238 Zq + .176 Za.

Regional means were also computed for FYA, GRE-Q, GRE-A, GRE-V, and FYA', where FYA' = predicted FYA = .0013 (Q) + .0006 (A) + 2.2831, a regression equation based on the total-sample analysis alluded to above. These means were expressed as deviations from the grand mean for all quantitative students without regard to department of enrollment.

* Trends in findings of regression analyses conducted separately for engineering, math/science, and economics samples were similar to those reported here for the combined sample of students from quantitative departments.

Findings

Means of groups on the original FYA and GRE variables are shown in Table 14. Table 15 shows (a) these means expressed as deviations from the grand means for all students without regard to department and (b) corresponding regional means on the departmentally z-scaled variables. **In both tables, regions are listed in descending order with respect to mean FYA.** Findings for the sample of English-native-language (ENL) students are included for perspective.

With regard to differences among regional groups in "average academic performance," several trends are noteworthy:

- **The ranks of regional groups in terms of mean FYA corresponded closely to their ranks in terms of mean Zfya. Excluding America II, the ranks for ESL groups on the two indices of academic standing corresponded perfectly.**

- At one extreme, students from the two European contingents earned grades averaging about 3.6, and they were in department-level samples in which they tended to outperform other ESL students. At the other extreme, students from the two African contingents and the Mideast earned grades averaging about 3.3, and they were in department-level samples in which they tended perform at a lower level than other ESL students.

- The two large contingents from Asia (accounting for some 68 percent of all foreign students in the study sample) earned grades averaging approximately 3.5, corresponding to the grand mean for all students, and their z-scaled means were approximately 0.0. By inference, Asian students tended to provide a substantial "common element" in the regional mix of the respective department-level samples; their academic performance tended to influence both the department-level sample FYA means and the total sample FYA mean.

- Effects associated with department of enrollment are illustrated in the data for students from America II. Their FYA mean of 3.44 placed them 0.12 standard deviations below the grand mean for all students without regard to department (mean $\overline{FYA} = -0.12$), but their mean Zfya was 0.11, indicating that their FYAs tended to be higher than the general ESL means in their respective departments. Thus, by inference, these students tended to be from department-level ESL samples with comparatively low mean FYA.

It is also noteworthy that on both indices of "relative academic standing," the average standing of students from Category I countries (with ESL-conducive backgrounds) was similar to that of students from Category II countries (with ESL-resistant backgrounds). These background differences are reflected in regional means on the verbal measure and on the analytical measure: **the verbal and analytical means of Region I ESL students were higher than those for Region II ESL students.**

Data for the contingent of ENL (English-native-language students) suggest that their academic performance tended to be roughly comparable to that of the typical ESL student in their respective departments. Their mean FYA, 3.54,

Table 14

Means of Regional Groups on Regularly Scaled FYA and GRE Scores: Pooled Data for 1,213 Students in 86 Engineering, Math/Science, and Economics Departments

Region	N	FYA	GRE-Q	GRE-A	GRE-V
Europe I	62	3.63	696	569	447
Europe II	78	3.57	681	497	366
Asia II	643	3.52	721	485	352
Asia I	188	3.49	708	523	500
America II	86	3.44	628	478	404
Africa II	33	3.34	609	415	324
Mideast	98	3.32	662	466	344
Africa I	17	3.31	639	510	432
All "I"*	273	3.50	697	530	482
All "II"	840	3.51	703	483	357
ESL Total**	1213	3.49	698	492	384
Standard Deviation		(0.41)	(81)	(106)	(106)
Foreign ENL	31	3.54	726	610	539

Note. Regions are listed in descending order with respect to mean FYA.

* Includes data for six America I students.

** Includes Mideast and two students not classified by region.

Table 15

Means of Foreign Students on FYA and GRE Variables, by Region, Expressed (a) as Deviations from the Means of Their Respective Departments of Enrollment and (b) as Deviations from Grand Means for All Students: Data for All Quantitative Departments

Region	N	Z(fya) (a)	FYA (b)	Z(q) (a)	GRE-Q (b)	Z(a) (a)	GRE-A (b)	Z(v) (a)	GRE-V (b)
Europe I	62	0.27	0.14	-0.16	-0.02	0.42	0.73	0.43	0.59
Europe II	78	0.22	0.08	-0.31	-0.22	-0.02	0.05	-0.16	-0.17
Asia II	643	0.02	0.07	0.24	0.28	-0.04	-0.07	-0.25	-0.30
Asia I	188	-0.02	0.00	0.05	0.12	0.22	0.29	0.88	1.09
America II	86	0.11	-0.12	-0.42	-0.86	0.07	-0.13	0.23	0.19
Africa II	17	-0.18	-0.37	-0.72	-1.10	-0.56	-0.73	-0.52	-0.57
Mideast	98	-0.34	-0.41	-0.52	-0.44	-0.23	-0.25	-0.28	-0.38
Africa I	17	-0.47	-0.44	-0.52	-0.73	0.02	0.17	0.22	0.45
All "I"	273*	-0.00	0.02	-0.07	-0.07	0.24	0.36	0.72	0.92
All "II"	840	0.04	0.05	0.09	0.01	-0.05	-0.08	-0.20	-0.25
ESL Total	1213**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Foreign ENL	31	0.08	0.12	-0.09	0.35	1.19	1.20	0.84	1.46

Note. Regions are listed in descending order with respect to mean FYA. Column (a) entries are means on departmentally z-scaled variables, indicating average within-department standing of students from a region. For example, the average within-department FYA of students from Europe I was 0.27 standard deviations above the means of their respective departments; their average within-department standing for quantitative ability was relatively lower (mean Z(q) = -0.16), they were above average (by 0.42 standard units) for analytical ability, Z(a), and so on. Column (b) entries represent the means of groups on FYA (A = 4, B = 3, . . ., F = 0) and GRE scaled scores expressed as deviations from grand means for all students without regard to department of enrollment, in total sample standard deviation units. For example, the average FYA for Europe I students (3.63) was approximately 0.14 standard deviations (S.D. = 0.41) higher than that for all students (grand mean FYA = 3.49), their GRE-Q mean (696) was approximately 0.02 standard deviations (S.D. = 81) below the grand mean (698) for all students, and so on. For the English-native-language (ENL) students, means are scaled relative to data for ESL students. ENL students earned grades averaging about 0.12 standard deviations better than average for foreign ESL students in their respective departments, their GRE verbal scores averaged 1.46 standard deviations above department (ESL) means, and so on.

* Includes 6 America I students. ** Includes Mideast and two students without data on country of citizenship.

placed them 0.12 standard units above the grand mean for all ESL students and their mean Zfya = 0.08 (z-scaled relative to the mean for ESL students in their respective departments) indicated similar relative within-department standing. ENL students had the highest standing of any group on GRE-Q (mean = 726, 0.35 standard deviations above the grand mean). However, their Zq mean was -0.09, indicating slightly lower than average relative standing within their respective departments. By inference, the ENL students tended to be from selective departments that enrolled ESL (and other) students with very high quantitative scores.

Observed versus Predicted Criterion Standing of Regional Groups

Table 16 shows the observed means of regional groups on FYA and on Zfya and the corresponding predicted means. FYA' (predicted FYA) is based on the regression of FYA on GRE-Q and GRE-A in the total sample; Z'fya (predicted Zfya) is based on the regression of Zfya on Zq and Za in the same sample. FYA and FYA' means are expressed as deviations from the grand mean in standard deviation units (grand mean = 3.49, S.D. = 0.41).

Figure 1 points up the generally parallel nature of findings from the across-departments and the within-departments analyses. In each of the three frames, subgroups are ordered from left to right on mean FYA. Mean FYA and mean FYA', from the across-departments analysis, are shown in the top left frame, and mean Zfya and mean Zfya', from the within-departments analysis, are shown in the bottom left frame. Mean residuals (mean of "observed minus predicted performance" values) from the across-departments analysis and the within-department analysis are plotted in the top right frame.

- The profiles of mean residual values are quite similar. European students, with generally higher criterion standing, tended to have higher observed than predicted standing, and students from the Mideast and Africa II, with generally lower criterion standing, tended to have lower observed than predicted standing. There was more consistency between observed and predicted standing for the remaining groups (ENL, Asia I, and Asia II with higher observed criterion standing, and Africa II with lower criterion standing).

- Only for students from America II were across-department results appreciably different from within-department results. In the across-departments analysis, observed FYA and predicted FYA values for these students were similar—both were below the grand mean (FYA = 3.49) for all students, without regard to department of enrollment. However, this group enjoyed comparatively higher average within-department standing (mean Zfya was 0.11), while their average predicted relative within-department standing was lower (mean Z'fya was -0.09). Such a pattern of findings reflects effects associated with the particular "regional mix" and/or grading standards of the particular departments in which these students were enrolled, as indicated earlier.

Table 16

Grand Mean FYA and Mean FYA' as Compared to Within-Department Mean Zfya and Mean Z'fya for Regional Groups: Pooled Data for All Quantitative Departments*

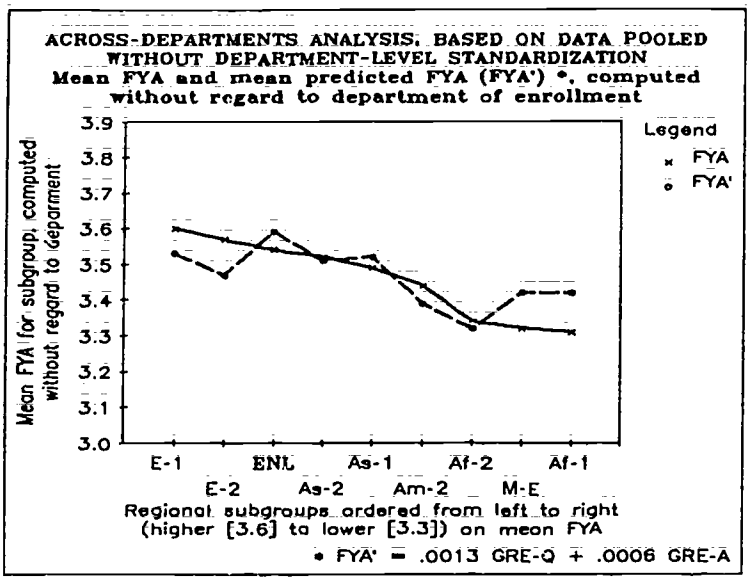
Region	N	Mean FYA	Mean FYA'	As deviations from grand mean #		Mean Zfya	Mean Z'fya
				FYA	FYA'		
Europe I	62	3.63	3.53	0.34	0.10	0.27	0.04
Europe II	78	3.57	3.47	0.20	-0.05	0.22	-0.08
Asia II	643	3.52	3.51	0.07	0.05	0.02	0.05
Asia I	188	3.49	3.52	0.00	0.07	-0.02	0.05
America II	86	3.44	3.39	-0.12	-0.24	0.11	-0.03
Africa II	33	3.34	3.32	-0.37	-0.41	-0.18	-0.27
Mideast	98	3.32	3.42	-0.41	-0.17	-0.34	-0.16
Africa I	17	3.31	3.42	-0.44	-0.17	-0.47	-0.12
All "I"	273	3.50	3.51	0.02	0.05	-0.00	0.03
All "II"	840	3.51	3.49	0.05	0.00	0.04	0.01
ESL Total	**1213****	**3.49**	**3.49**	**0.00**	**0.00**	**0.00**	**0.00**
ENL	31	3.54	3.59	0.12	0.24	0.08	0.13

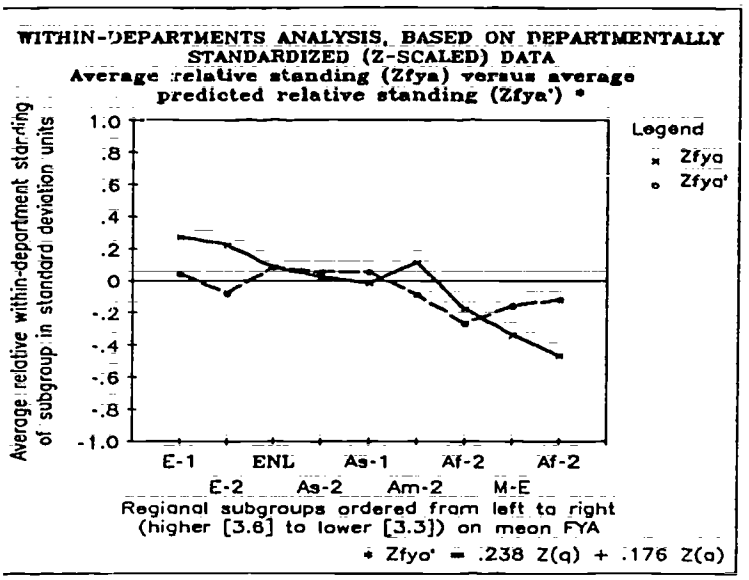
Note. Regional groups are in descending order with respect to mean FYA.

* FYA = regularly scaled FYA; FYA' = predicted FYA, using **.0013 (Q) + .0006 (A) + 2.2381**, an equation based on the regression of FYA on GRE scaled scores for the combined sample of ESL students from quantitative departments.

Zfya = departmentally z-scaled FYA; Z'fya = predicted Zfya, using **.238 Zq + .176 Za = 0.00**, an equation based on the regression of Zfya on the departmentally z-scaled GRE predictors for the same combined sample of ESL students from quantitative departments.

The observed and predicted FYA means for regional groups are expressed here as deviations from the grand FYA mean (ESL total mean = 3.49) in total sample standard deviation units (S.D. = 0.41).





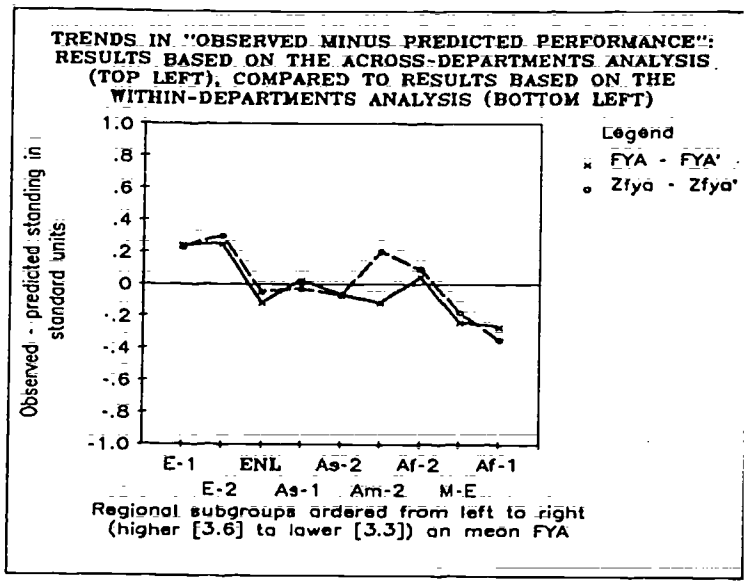


Figure 1. Findings based on analysis of data across departments versus findings based on analysis of data within departments

Generally speaking, the findings reviewed in this section indicate that there were differences among the regional subgroups in level of academic performance. There was a general tendency for subgroups with higher average criterion performance to have higher scores on GRE predictive composites. However, the findings also suggested the possibility of predictive bias for regional subgroups within the foreign ESL student population. For example, European students tended to have somewhat higher relative standing on mean FYA and mean Zfya than expected from the corresponding sets of GRE measures while the opposite was true for students from the Mideast and Africa. These particular findings should be viewed primarily as working hypotheses for future research concerned specifically with the assessment of predictive bias.

In the meantime, departments may assess the relevance of the trends revealed in this exploratory analysis by observing the typical patterns of academic performance of students from various regional subgroups such as those defined for this study.

References

Braun, H. I., & Jones, D. H. (1985). Use of empirical Bayes methods in the study of the validity of academic predictors of graduate school performance (GRE Board Professional Report GREB No. 79-13P & ETS RR-84-34). Princeton, NJ: Educational Testing Service.

Burton, N., & Turner, N. J. (1983). Effectiveness of the Graduate Record Examinations for predicting first-year grades: 1981-82 summary report of the Graduate Record Examinations Validity Study Service. Princeton, NJ: Educational Testing Service.

Educational Testing Service (1983). TOEFL Test and score manual. Princeton, NJ: Author.

Educational Testing Service (1985). Guide to the use of the Graduate Record Examinations Program, 1986-87. Princeton, NJ: Author.

Kingston, N. M. (1985). The incremental validity of the GRE Analytical measure for predicting first-year grade-point average (unpublished paper presented at the annual meeting of the American Educational Research Association).

Linn, R. L., Harnisch, D. L., & Dunbar, S. B. (1981). Validity generalization and situational specificity: An analysis of the prediction of first-year grades in law school. Applied Psychological Measurement, 5, 281-289.

Miller, R., & Wild, C. L., Eds. (1979). Restructuring the Graduate Record Examinations Aptitude Test (GRE Board Technical Report). Princeton, NJ: Educational Testing Service.

Mosteller, F. M., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), Handbook of social psychology (Vol. 1, pp. 289-334). Cambridge, MA: Addison-Wesley.

Pearlman, K., Schmidt, F. L., & Hunter, J. E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. Journal of Applied Psychology, 65, 373-406.

Powers, D. E. (1980). The relationship between scores on the Graduate Management Admission Test and the Test of English as a Foreign Language (TOEFL Research Report No. 5). Princeton, NJ: Educational Testing Service.

Shaw, A. T. (1972). English proficiency, verbal aptitude, and foreign student success in American graduate schools. Educational and Psychological Measurement, 32, 425-431.

Swinton, S. S. (1985). The predictive validity of the restructured GRE with particular attention to older students (draft project report). Princeton, NJ: Educational Testing Service.

Willingham, W. W. (1974). Predicting success in graduate education, Science, 183, 273-278.

Wilson, K. M. (1979). The validation of GRE scores as predictors of first-year performance in graduate study: Report of the GRE Cooperative Validity Studies Project (GRE Board Research Report No. 75-8R). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1982a). A comparative analysis of TOEFL examinee characteristics (TOEFL Research Report No. 11). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1982b). GMAT and GRE Aptitude Test performance in relation to primary language and scores on TOEFL (TOEFL Research Report No. 12 & ETS RR-82-28). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1982c). A study of the validity of the restructured GRE Aptitude Test for predicting first-year performance in graduate study (GRE Board Research Report No. 78-6R & ETS RR-82-34). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1984a). Foreign nationals taking the GRE General Test during 1981-82: Highlights of a study (GRE Board Research Report GREB No. 81-23aR & ETS RR-84-23). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1984b). Foreign nationals taking the GRE General Test during 1981-82: Selected characteristics and test performance (GRE Board Professional Report No. 81-23bP & ETS RR-84-39). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1984c). The relationship of GRE General Test item-type part scores to undergraduate grades (GRE Board Professional Report GREB No. 81-22P & ETS RR-84-38). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1985). Factors affecting GMAT predictive validity for foreign MBA students: An exploratory study (ETS RR-85-17). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1986a). The relationship of scores based on GRE General Test item types to undergraduate grades: An exploratory study for selected subgroups (GREB report, in press). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1986b). The GRE Subject Test performance of U.S. and Non-U.S. Examinees, 1982-84: A comparative analysis (GREB report, in press). Princeton, NJ: Educational Testing Service.

Appendix A: Summary of Department-Level Data

A-1 GRE and FYA Means and Standard Deviations, and GRE/FYA Correlations for Department-Level Samples of Foreign ESL Students, by Field

A-2 Simple Correlation of GRE Predictors with FYA for Foreign ESL Samples, by School and Department

A-3 Distributions of GRE Means for Departments, by Graduate Area

A-4 Distributions of Standard Deviations of GRE General Test Scores for Department-Level Samples in 86 Primarily Quantitative Departments, Six Bioscience, and Five Social Science Departments

A-5 Distributions of Standard Deviations of GRE General Test Scores for Department-Level Samples in 86 Primarily Quantitative Departments, By Size of Sample and without Regard to Sample Size, Respectively

A-6 Distributions of Correlations between GRE Predictors and FYA for 86 Department-Level Samples of ESL Students in Primarily Quantitative Fields

A-7 Distributions of Department-Level GRE General Test Validity Coefficients, by Size of Sample: Data of 86 Departments in Primarily Quantitative Fields

GRE and FYA Means and Standard Deviations, and GRE/FYA Correlations, for Department-Level Samples of Foreign ESL Students, by Field

DEPARTMENT/ SCHOOL NO.	N	Mean FYA	Mean GRE-V	Mean GRE-Q	Mean ANALY	Standard deviation FYA	Standard deviation GRE-V	Standard deviation GRE-Q	Standard deviation ANALY	Correlation with FYA GRE-V	Correlation with FYA GRE-Q	Correlation with FYA ANALY
CHEMICAL ENGIN												
64 2	17	3.31	408.8	716.5	503.5	0.40	125.6	64.6	98.5	0.250	-.023	0.368
64 7	9	3.59	384.4	740.0	520.0	0.28	60.0	25.5	53.4	-.136	0.555	0.689
64 10	14	3.43	400.0	697.1	497.9	0.32	136.2	86.3	115.1	0.474	0.482	0.467
64 14	6	3.67	438.3	751.7	591.7	0.15	84.2	23.2	4	0.226	0.729	0.134
64 24	5	3.66	406.0	692.0	430.0	0.31	111.5	82.3	.8	0.546	-.132	0.782
64 30	13	3.77	345.4	723.8	530.8	0.19	50.4	52.5	81.7	-.055	0.640	0.600
64 65	5	3.48	324.0	732.0	442.0	0.34	75.7	25.9	85.0	0.792	0.236	0.061
64 73	11	3.41	480.0	740.0	505.5	0.29	104.6	41.7	71.9	-.242	0.056	-.260
64 TOTAL	80	3.51	400.7	722.2	507.2	0.31	97.9	54.5	87.0	0.173	0.315	0.362
CIVIL ENGIN												
65 2	32	3.56	415.6	726.9	533.7	0.40	106.0	45.3	84.9	-.266	0.136	0.207
65 7	30	3.45	332.3	673.3	426.3	0.29	64.9	64.1	62.9	-.067	0.116	0.132
65 19	10	3.71	415.0	732.0	537.0	0.22	55.4	60.9	88.3	0.229	0.053	0.836
65 24	42	3.44	346.7	691.0	471.7	0.46	88.8	66.8	106.7	-.020	0.106	0.060
65 30	13	3.59	365.4	723.1	530.0	0.36	72.1	48.7	85.7	0.342	0.788	0.410
65 33	11	3.64	356.4	690.9	490.9	0.30	81.4	67.6	94.8	-.614	0.504	-.498
65 38	19	3.65	365.3	674.7	457.4	0.30	66.5	74.1	92.4	0.188	-.094	-.397
65 49	5	3.00	388.0	632.0	472.0	0.63	144.9	50.2	48.7	0.028	-.096	0.089
65 TOTAL	162	3.52	367.5	696.2	483.9	0.37	83.1	60.6	87.2	-0.047	0.163	0.088
ELECTRICAL ENGIN												
66 7	33	3.49	336.7	692.7	467.3	0.36	79.9	74.4	91.4	0.109	0.451	0.365
66 10	31	3.49	403.5	742.9	522.3	0.36	105.2	34.3	95.2	0.315	0.289	0.285
66 13	14	3.54	349.3	649.3	482.9	0.37	103.7	119.9	138.0	0.348	0.466	0.387
66 14	31	3.47	408.7	712.6	512.6	0.33	97.2	66.9	100.9	0.049	0.327	0.307
66 24	19	3.76	386.3	727.4	486.8	0.31	107.9	72.5	102.6	0.082	0.521	0.234
66 30	12	3.54	339.2	734.2	536.7	0.29	48.9	42.9	108.2	-.054	0.759	0.429
66 33	36	3.42	421.1	711.4	537.5	0.46	129.9	78.2	107.9	0.293	0.215	0.436
66 34	52	3.55	390.4	748.3	544.0	0.45	116.1	65.9	103.1	0.158	0.024	0.309
66 41	15	3.52	339.3	715.3	466.0	0.33	81.6	55.0	118.0	0.159	0.076	0.470
66 49	13	3.21	343.1	708.5	481.5	0.67	116.1	60.8	111.1	0.430	0.471	0.219
66 TOTAL	256	3.50	381.2	719.4	511.1	0.40	103.3	66.5	104.4	0.185	0.296	0.340
INDUSTRIAL ENGIN												
67 2	12	3.28	443.3	717.5	567.5	0.31	91.7	64.5	110.3	0.405	0.419	0.519
67 7	12	3.63	410.8	735.0	476.7	0.32	58.3	40.8	111.9	-.182	0.685	0.612
67 13	9	3.43	357.8	697.8	478.9	0.41	142.5	71.6	100.6	0.584	0.671	0.482
67 24	13	3.44	365.4	680.0	437.7	0.35	163.6	100.7	84.5	0.193	0.451	0.382
67 38	25	3.53	378.8	681.2	487.6	0.33	73.5	68.7	72.3	-.026	0.158	0.047
67 73	18	3.36	404.4	644.4	433.9	0.30	122.2	86.2	96.7	0.270	0.362	0.355
67 TOTAL	89	3.45	392.9	687.4	477.9	0.33	106.2	72.9	92.3	0.165	0.400	0.342

GRE and FYA Means and Standard Deviations, and GRE/FYA Correlations, for Department-Level Samples of Foreign ESL Students, by Field

DEPARTMENT/ SCHOOL NO.	N	Mean FYA	Mean GRE-V	Mean GRE-Q	Mean ANALY	Standard deviation FYA	Standard deviation GRE-V	Standard deviation GRE-Q	Standard deviation ANALY	Correlation with FYA GRE-V	Correlation with FYA GRE-Q	Correlation with FYA ANALY
MECHANICAL ENGIN												
68 7	7	3.77	432.9	751.4	521.4	0.31	107.7	29.7	88.6	0.533	-.597	0.457
68 [illegible]	14	3.40	387.1	726.4	535.0	0.49	90.9	75.6	108.4	0.381	0.095	-.145
68 14	20	3.67	340.0	724.5	475.5	0.30	70.0	55.4	81.3	0.266	0.612	0.588
68 24	24	3.64	337.5	706.2	455.0	0.41	82.4	95.7	97.9	-.135	0.563	0.255
68 30	10	3.67	391.0	751.0	534.0	0.31	80.8	38.4	91.2	-.618	0.200	-.304
[illegible]	17	3.40	358.8	685.9	453.5	0.39	108.7	79.1	111.9	0.202	0.004	0.375
68 38	13	3.47	360.0	693.8	476.9	0.67	55.2	52.0	67.6	-.138	0.587	0.347
68 65	10	3.66	438.0	711.0	536.0	0.36	131.2	66.2	108.3	0.535	0.725	0.763
68 TOTAL	115	3.57	368.9	714.5	488.5	0.41	87.7	67.3	94.7	0.104	0.347	0.300
ALL ENGIN	702	3.51	379.7	709.5	496.5	0.37	95.8	64.7	95.3	0.114	0.289	0.278
APPLIED MATH												
54 28	5	3.53	208.0	690.0	470.0	0.54	48.2	43.6	76.8	-.272	-.178	-.458
54 TOTAL	5	3.53	208.0	690.0	470.0	0.54	48.2	43.6	76.8	-0.272	-0.178	-0.458
STATISTICS												
59 2	9	3.72	408.9	718.3	527.8	0.31	88.8	99.7	147.4	0.151	0.458	0.474
59 7	11	3.71	284.5	702.7	462.7	0.25	49.1	66.0	86.6	-.119	0.500	-.127
59 14	15	3.48	370.0	709.3	496.0	0.41	74.4	60.5	109.3	0.507	0.275	0.638
59 24	12	3.42	337.5	670.8	458.3	0.43	113.5	59.5	60.9	0.206	-.104	0.550
59 28	7	3.39	374.3	652.9	427.1	0.35	107.4	105.3	166.6	0.036	0.761	0.563
59 TOTAL	54	3.54	352.4	693.3	477.2	0.36	84.6	73.8	107.7	0.189	0.330	0.425
CHEMISTRY												
62 7	11	3.53	350.0	710.9	503.6	0.28	52.0	59.2	86.6	-.243	0.506	0.447
62 24	10	3.[illegible]	352.0	664.0	402.0	0.61	124.6	75.0	130.1	0.424	0.666	0.668
62 28	26	3.64	355.8	707.7	455.8	0.30	78.4	67.4	88.9	0.184	0.401	-.177
62 33	13	3.17	333.8	660.8	428.5	0.47	76.7	64.4	113.3	0.071	0.365	0.542
62 34	5	3.51	378.0	692.0	514.0	0.20	130.8	49.2	71.3	-.723	-.375	-.646
62 38	10	3.69	314.0	658.0	439.0	0.29	71.2	70.5	73.7	-.111	0.654	0.323
62 41	5	3.63	388.0	830.0	472.0	0.52	63.8	124.7	166.6	0.013	-.019	-.026
62 65	6	3.42	330.0	628.3	410.0	0.26	76.7	111.4	86.3	0.708	0.430	0.492
62 67	9	3.[illegible]	427.8	705.6	500.0	0.26	115.4	57.5	74.2	-.124	0.007	0.104
62 92	16	3.49	388.1	610.0	453.1	0.47	77.0	104.9	86.7	-.084	0.288	0.205
62 TOTAL	111	3.48	360.1	671.4	455.0	0.37	83.5	76.0	94.9	0.012	0.353	0.190
MATHEMATICS												
72 13	8	3.87	356.2	703.7	463.7	0.25	89.0	61.6	168.9	-.065	0.660	0.220
72 24	15	3.32	334.7	634.0	398.0	0.47	77.1	85.8	109.0	-.098	0.550	0.234
72 34	5	3.50	372.0	752.0	528.0	0.29	137.7	42.7	99.3	0.717	0.574	0.480
72 38	11	3.55	334.5	650.0	432.7	0.38	63.1	83.1	103.9	-.190	0.644	0.246
72 TOTAL	39	3.52	343.8	667.9	437.9	0.37	83.4	74.5	118.6	-0.013	0.602	0.266

GRE and FYA Means and Standard Deviations, and GRE/FYA Correlations, for Department-Level Samples of Foreign ESL Students, by Field

DEPARTMENT/ SCHOOL NO.	N	Mean FYA	Mean GRE-V	Mean GRE-Q	Mean ANALY	Standard deviation FYA	Standard deviation GRE-V	Standard deviation GRE-Q	Standard deviation ANALY	Correlation with FYA GRE-V	Correlation with FYA GRE-Q	Correlation with FYA ANALY
PHYSICS												
76 2	6	3.75	595.0	738.3	611.7	0.28	99.1	53.4	149.3	-.339	0.539	0.341
76 7	13	3.50	417.3	723.8	496.9	0.33	112.0	68.3	127.4	0.319	0.485	0.432
76 13	8	3.86	398.7	721.2	496.2	0.27	84.9	55.1	85.2	-.667	-.102	0.337
76 30	5	3.45	480.0	720.0	620.0	0.58	192.0	107.3	107.0	0.042	0.816	0.578
76 33	8	3.67	437.5	710.0	542.5	0.32	83.3	63.9	97.8	0.025	0.450	0.632
76 41	11	3.55	403.6	719.1	532.7	0.42	75.0	84.4	129.6	0.305	0.605	0.433
76 TOTAL	51	3.62	440.4	721.6	537.3	0.36	101.6	71.1	117.2	0.011	0.452	0.452
COMPUTER												
78 7	12	3.70	503.3	758.3	604.2	0.32	148.1	83.5	81.5	-.123	0.416	0.240
78 14	9	3.79	476.7	775.6	524.4	0.23	186.8	30.9	71.6	-.052	0.540	0.055
78 24	21	3.06	333.3	690.0	469.0	0.54	75.6	82.6	74.5	0.372	0.367	0.628
78 30	11	3.50	410.9	688.2	506.4	0.60	87.3	67.6	79.5	-.392	0.306	-.262
78 33	16	3.62	458.1	710.6	553.7	0.27	104.2	66.2	86.2	0.040	0.144	0.463
78 38	12	3.22	416.7	738.3	522.5	0.45	149.1	26.6	74.6	-.225	-.006	-.287
78 41	5	3.46	526.0	746.0	666.0	0.45	135.0	27.0	106.4	-.658	-.936	0.520
78 49	7	3.27	338.6	657.1	494.3	0.54	29.7	114.8	80.0	0.132	0.155	0.445
78 80	13	3.39	468.5	720.8	576.2	0.58	155.2	50.6	65.5	-.188	0.467	-.015
78 82	7	3.70	432.9	717.1	538.6	0.16	104.8	54.7	103.0	0.540	0.356	0.234
78 TOTAL	113	3.43	427.4	717.6	536.0	0.43	115.9	63.0	79.7	-0.013	0.249	0.220
MATH/PHY SCI	373	3.50	387.8	695.3	492.4	0.39	95.5	70.5	97.4	0.023	0.351	0.268
ECONOMICS												
84 2	10	3.40	549.0	746.0	593.0	0.53	125.6	60.4	57.7	-.097	0.080	0.170
84 7	15	3.58	380.0	736.7	498.0	0.28	52.1	71.3	95.6	0.669	0.359	0.326
84 13	12	3.65	400.0	688.3	480.8	0.32	55.1	97.1	85.0	0.352	0.079	0.497
84 14	6	3.73	453.3	756.7	503.3	0.24	56.5	41.8	85.5	-.225	0.041	0.600
84 24	11	3.52	400.9	679.1	487.3	0.36	98.6	110.3	73.6	-.252	0.631	0.692
84 30	11	3.24	371.8	630.0	514.5	0.18	55.3	121.9	89.6	0.351	-.329	0.157
84 32	48	3.19	374.0	570.8	408.7	0.35	109.0	103.3	104.8	0.338	0.465	0.164
84 34	5	3.26	364.0	702.0	416.0	0.29	119.1	40.2	117.2	0.495	0.278	0.392
84 41	12	3.70	427.5	684.2	525.8	0.15	61.8	68.3	103.9	-.176	0.315	0.430
84 87	8	3.27	365.0	607.5	441.2	0.34	136.0	59.9	100.8	-.080	0.632	-.190
84 TOTAL	138	3.39	398.8	649.9	469.2	0.32	89.8	87.7	94.3	0.210	0.313	0.282
ECONOMICS	138	3.39	398.8	649.9	469.2	0.32	89.8	87.7	94.3	0.210	0.313	0.282
QUANTITATIV	1213	3.49	384.4	698.4	492.1	0.37	95.0	69.1	95.9	0.097	0.311	0.275

GRE and FYA Means and Standard Deviations, and GRE/FYA Correlations, for Department-Level Samples of Foreign ESL Students, by Field

DEPARTMENT/ SCHOOL NO.	N	Mean				Standard deviation				Correlation with FYA		
		FYA	GRE-V	GRE-Q	ANALY	FYA	GRE-V	GRE-Q	ANALY	GRE-V	GRE-Q	ANALY
MICROBIOLOGY												
7 24	6	3.43	410.0	691.7	540.0	0.46	120.5	83.8	97.4	-.751	0.369	-.569
7 TOTAL	6	3.43	410.0	691.7	540.0	0.46	120.5	83.8	97.4	-0.751	0.369	-0.569
AGRICULTURAL E												
31 7	7	3.26	332.9	514.3	417.1	0.44	39.0	174.7	112.1	-.776	0.709	0.708
31 38	16	3.18	348.1	503.7	361.9	0.48	76.0	119.0	78.8	0.512	0.083	0.603
31 TOTAL	23	3.21	343.5	507.0	378.7	0.47	64.8	136.0	89.0	0.135	0.273	0.635
BIOCHEMISTRY												
34 7	8	2.96	453.7	653.7	517.5	0.47	110.3	72.1	104.4	0.245	-.095	0.075
34 91	11	3.02	369.1	676.4	469.1	0.67	83.0	55.9	95.1	-.588	-.155	-.260
34 TOTAL	19	2.99	404.7	666.8	489.5	0.59	94.5	62.7	99.0	-0.238	-0.129	-0.119
BIOLOGY												
35 49	7	3.35	371.4	612.9	504.3	0.59	61.5	96.6	55.3	0.664	-.384	-.641
35 TOTAL	7	3.35	371.4	612.9	504.3	0.59	61.5	96.6	55.3	0.664	-0.384	-0.641
BIO SCI	55	3.19	375.5	595.8	450.5	0.52	80.7	100.0	89.1	-0.023	0.061	0.081
EDUCATION												
85 2	26	3.22	421.9	543.1	432.3	0.64	127.2	134.9	101.3	0.374	0.251	0.080
85 7	17	3.56	323.5	512.9	373.5	0.17	53.1	175.3	89.0	-.018	0.004	0.006
85 24	21	3.45	296.2	562.4	409.0	0.15	48.7	141.7	87.3	0.020	0.122	0.130
85 33	12	3.68	334.2	488.3	405.8	0.36	113.4	165.2	71.8	0.430	0.133	0.681
85 TOTAL	76	3.43	351.3	533.0	408.6	0.36	86.8	150.6	90.0	0.197	0.142	0.172
POLITICAL SCI												
92 7	9	3.57	353.3	611.1	464.4	0.17	85.4	122.1	118.6	0.727	-.097	0.282
92 TOTAL	9	3.57	353.3	611.1	464.4	0.17	85.4	122.1	118.6	0.727	-0.097	0.282
SOCIAL SCI	85	3.44	351.5	541.3	414.5	0.34	86.6	147.6	93.1	0.253	0.116	0.184
TOTAL	1353	3.48	382.0	684.3	485.6	0.37	93.9	75.3	95.4	0.102	0.288	0.262

Simple Correlation of GRE Predictors with FYA for Foreign ESL Samples, by School and Department

School	Department	N	Correlation of GRE predictor with FYA		
			GRE-V	GRE-Q	GRE-A
02	Chem Engineering	17	.25	-.02	.37
	Civil Engineering	32	-.27	.14	.21
	Indust Engineering	12	.40	.42	.52
	Statistics	9	.13	.46	.47
	Physics	6	-.34	.54	.34
	Economics	10	-.10	.08	.17
	Education	26	.17	.25	.08
07	Chem Engineering	9	-.18	.56	.69
	Civil Engineering	30	-.07	.12	.13
	Elect Engineering	33	.11	.45	.36
	Indust Engineering	12	-.18	.68	.61
	Mech Engineering	7	.53	-.60	.46
	Statistics	11	-.12	.50	-.13
	Chemistry	11	-.24	.51	.45
	Physics	13	.32	.48	.43
	Computer Sci	12	-.12	.42	.24
	Economics	15	.67	.36	.33
	Agriculture	7	-.73	.71	.71
	Biochemistry	8	.24	-.10	.08
	Education	17	-.02	.00	.01
	Political Science	9	.73	-.10	.28
10	Chem Engineering	14	.47	.48	.47
	Elect Engineering	31	.32	.29	.28
13	Elect Engineering	14	.45	.47	.39
	Indust Engineering	9	.58	.67	.48
	Mech Engineering	14	.38	.10	-.14
	Mathematics	8	-.06	.66	.22
	Physics	8	-.67	-.10	.34
	Economics	12	.35	.08	.50

GRE/FYA correlation, by school and department, continued

School	Department	N	Correlation of GRE predictor with FYA		
			GRE-V	GRE-Q	GRE-A
14	Chem Engineering	6	.23	.73	.13
	Elect Engineering	31	.05	.33	.31
	Mech Engineering	20	.27	.61	.59
	Statistics	15	.51	.28	.64
	Computer Science	9	-.05	.54	.06
	Economics	6	-.22	.04	.60
19	Elect Engineering	10	.23	.05	.84
24	Chem Engineering	5	.55	-.13	.78
	Civil Engineering	42	-.02	.11	.06
	Elect Engineering	19	.08	.52	.23
	Indust Engineering	13	.19	.45	.38
	Mech Engineering	24	-.14	.56	.26
	Statistics	12	.21	-.10	.55
	Chemistry	10	.42	.67	.67
	Mathematics	15	-.10	.55	.23
	Computer Science	21	.37	.37	.63
	Economics	11	-.25	.63	.69
	Microbiology	6	-.75	.37	-.57
	Education	21	.02	.12	.13
28	Applied Mathematics	5	-.27	-.18	-.46
	Statistics	7	.04	.76	.56
	Chemistry	26	.18	.40	-.18
30	Chem Engineering	30	-.06	.64	.60
	Civil Engineering	13	.34	.79	.41
	Elect Engineering	12	-.05	.76	.43
	Mech Engineering	10	-.62	.20	-.30
	Physics	5	.04	.82	.58
	Computer Science	12	-.33	.31	-.24
	Economics	11	.35	-.32	.16
32	Economics	48	.34	.46	.16

GRE/FYA correlations, by school and department, continued

School	Department	N	Correlation of GRE predictor with FYA		
			GRE-V	GRE-Q	GRE-A
33	Civil Engineering	11	-.61	.50	-.50
	Elect Engineering	36	.29	.22	.44
	Mech Engineering	17	.20	.00	.38
	Chemistry	13	.07	.36	.54
	Physics	8	.02	.45	.63
	Computer Science	16	.04	.14	.46
	Education	12	.43	.13	.68
34	Elect Engineering	52	.16	.02	.31
	Chemistry	5	-.72	-.38	-.65
	Mathematics	5	.72	.57	.48
	Economics	5	.50	.28	.39
38	Civil Engineering	19	.19	-.09	-.40
	Indust Engineering	25	-.03	.16	.05
	Mech Engineering	13	-.14	.59	.35
	Chemistry	10	-.11	.65	.32
	Mathematics	11	-.19	.64	.25
	Computer Science	12	-.22	-.00	-.29
	Agriculture	16	.51	.08	.60
41	Elect Engineering	15	.16	.08	.47
	Chemistry	5	.01	-.02	-.03
	Physics	11	.30	.60	.43
	Computer Science	5	-.66	-.94	.52
	Economics	12	-.18	.32	.43
49	Civil Engineering	5	.03	-.10	.09
	Elect Engineering	13	.43	.47	.22
	Computer Science	7	.13	.16	.44
	Biology	7	.66	-.38	-.64
60	Computer Science	13	-.19	.47	-.02

GRE/FYA correlations by school and department, concluded

School	Department	N	Correlation of GRE predictor with FYA		
			GRE-V	GRE-Q	GRE-A
67	Chemistry	9	-.12	.01	.10
73	Chem Engineering	11	-.24	.66	-.26
	Indust Engineering	18	.27	.36	.36
82	Computer Science	7	.54	.36	.23
87	Economics	8	-.08	.63	-.19
91	Biology	11	-.59	-.16	-.26
92	Chemistry	16	-.08	.29	.21

A-3

Distributions of GRE Means for Departments, by Graduate Area

Score category	Engineer/ Math/Sci			Economics			Bioscience			Soc Science			All Fields		
	Q	A	V	Q	A	V	Q	A	V	Q	A	V	Q	A	V
760+	1												1		
740-759	10			2									12		
720-739	15			1									16		
700-719	20			1									21		
680-699	14			2			1						17		
660-679	5	1		1			1						7	1	
640-659	6	-		-			1						7		
620-639	4	1		1			-			1			5	1	
600-619	1	2		1			1			-			4	2	
580-599		1	1	-	1		-			1			2	2	2
560-579		2	-	1	-		-			1			2	2	-
540-559		3	-		-	1	-	1		1			1	4	1
520-539		18	1		1	-	-	-		-			-	19	1
500-519		7	1		2	-	2	2		1			3	11	1
480-499		10	2		3	-		-		1			1	13	2
460-479		13	2		-	-		1			1			15	2
440-459		7	2		1	1		-	1		-			8	4
420-439		8	7		-	1		-	-		1	1		9	9
400-419		2	14		2	2		1	1		2	-		7	17
380-399		1	9			1		-	-		-	-		1	10
360-379			9			4		1	2		1	-		2	15
340-359			12						1			1			14
320-339			13						1			2			16
300-319			1									-			1
280-299			2									1			3
No. depts.		76			10						5			97	
Median	708	488	382	690	493	400	630	490	370	550	414	334	701	490	379
No. students		1075			138			55			85			1353	
Mean	705	495	382	650	469	399	596	450	376	541	414	352	684	486	382
General samples*															
Non-U.S.	652	458	364	585	443	387	543	428	382	517	420	391			
U.S.	645	582	520	592	569	498	533	530	498	483	496	485			

Note. Department means are based on data for a minimum of five students identified as nonnative English-speakers. Q = GRE quantitative, A = GRE analytical, and V = GRE verbal. Only students with GRE scores earned after 9/81 were included in department samples.

* Means for non-U.S. examinees and U.S. examinees tested between October 1981 and September 1982, classified by intended field of study.

Distributions of Standard Deviations of GRE General Test Scores for Department-Level Samples of Foreign ESL Students in 86 Primarily Quantitative Departments, Six Bioscience, and Five Social Science Departments

	GRE-Verbal			GRE-Quantitative			GRE-Analytical		
19	2			19			19		
18	7			18			18		
17	-			17		(5) (5)	17		
16	4			16		(-) (5)	16	179	
15	5			15		(-) (-)	15	-	
14	989			14		(-) (2)	14	79	
13	01568 16			13		(-) (5)	13	008	
12	5606	(0)*	(7)**	12	502	(-) (2)	12	7	
11	25924 626	(0)	(3)	11	150	(9)	11	70123 528	(2) (9)
10	57845 45689 9	(-)	(-)	10	05713 5	(-)	10	11367 44888 13357 89	(4) (1)
9	91297	(-)	(-)	9	76	(7)	9	08901 51256 7889	(5) (-)
8	34599 11557 0229	(3)	(5)	8	23446 366	(4)	8	05566 90224 56778 15679	(-) (97)
7	67127 04467 78	(6)	(-)	7	2D561 24489	(2)	7	12457 24452 4	(79) (2)
6	04235 6	(2)	(-)	6	02400 11446 68888 05667 79	(-)	6	1682	(5)
5	60255 552	(-)	(3)	5	03558 12294 557	(6)	5	38	
4	899	(-)	(9)	4	02349 12395		4	9	
3	0	(9)		3	0184		3		
2				2	36677		2		
Median	89	80	35		66	90 142		95	87 89
General sample#	119	116	123		102	131 142		112	118 117

Note. Read sample standard deviations by combining the initial digit(s) with sucessive subsequent digit(s) within each row. For example, from the last row in the distribution of sample standard deviations for GRE quantitative, for samples from departments without regard to size, it may be determined that there were five samples with standard deviations between 23 and 27, inclusive (that is, standard deviations of 23, 26, 26, 27, and 27).

* First column of parenthetical entries represents the distribution of standard deviations for the six bioscience samples. For example, one sample with verbal standard deviation of 120, another 110, and so on.

** Second column of parenthetical entries represents the distribution of standard deviations for five social science samples (four from education and one from political science): one science sample with a standard deviation of 127, and so on.

These are standard deviations of scores for a general sample of foreign GRE examinees tested during 1981-82, classified by intended graduate major as math/science, bioscience, and social science, respectively (Wilson, 1984, Table 9).

Distributions of Standard Deviations of GRE General Test Scores for Department-Level Samples of Foreign ESL Students in 86 Primarily Quantitative Departments, by Size of Sample and Without Regard to Sample Size, Respectively

	N = 5-9			N = 10-14			N = 15+		
	GRE V	GRE-Q	GRE-A	GRE-V	GRE-Q	GRE-A	GRE-V	GRE-Q	GRE-A
19	2								
18	7								
17	-								
16	-		179	4					
15	-		-	5					
14	9		79	89					
13	01568		-	16		008			
12	-	5	-	56	02	7	06		
11	259	15	7	246	0	01235	26		28
10	578	057	11367	45	1	44888	456889	35	1335789
9	9	-	089	129	7	015	7	6	12567889
8	34599	2	055669	11557	3446	022456778	0229	366	15679
7	67	2	12457	127	056	2445	0446778	124489	24
6	04	024	-	23	001144668888	168	56	056679	2
5	6	03558	3	025555	1229	8	2	4557	
4	8	02349	9	99	1239		2	457	
3	0	01			8			4	
2		3667			7				
Mdn Depts	105 (27)	53 (27)	89 (27)	81 (34)	62 (34)	86 (34)	82 (25)	71 (25)	97 (25)

GRE-Verbal Mdn = 89

19	2
18	7
17	-
16	4
15	5
14	989
13	01568 16
12	5606
11	25924 626
10	57845 45689 9
9	91297
8	34599 11557 0229
7	67127 04467 78
6	04235 6
5	60255 552
4	899
3	0
2	

GRE-Quantitative Mdn = 66

19	
18	
17	
16	
15	
14	
13	
12	502
11	150
10	05713 5
9	76
8	23446 366
7	20561 24489
6	02400 11446 68888 05667 79
5	03558 12294 557
4	02349 12395
3	0184
2	36677

GRE-Analytical* Mdn = 95

19	
18	
17	
16	179
15	-
14	79
13	008
12	7
11	70123 528
10	11367 44888 13357 89
9	08901 51256 7889
8	05566 90224 56778 15679
7	12457 24452 4
6	1682
5	38
4	9
3	
2	

Note. Read sample standard deviations by combining the initial digit(s) with successive subsequent digit(s) within each row. For example, from the last row in the distribution of sample standard deviations for GRE quantitative, for samples from departments without regard to size, it may be determined that there were five samples with standard deviations between 23 and 27, inclusive (that is, standard deviations of 23, 26, 26, 27, and 27).

Distributions of Correlations between GRE Predictors and FYA for 86 Department-Level Samples of ESL Students in Primarily Quantitative Fields

	GRE Verbal vs FYA		GRE Quantitative vs FYA		GRE Analytical vs FYA
.8		.8	2	.8	4
.7	29	.7	23669	.7	68
.6	7	.6	01334456778	.6	001334799
.5	0134458	.5	00124456679	.5	0245689
.4	0237	.4	02235556677788	.4	1333344566777889
.3	0224455578	.3	12366667	.3	11234456697889
.2	011335779	.2	0248899	.2	01223334568
.1	13366899	.1	0124466	.1	033667
.0	123444587	.0	012456888	.0	56669
-.0	235566788	-.0	0229	-.0	23
-.1	0012224488999	-.1	00038	-.1	3489
-.2	2244577	-.2	-	-.2	669
-.3	49	-.3	38	-.3	0
-.4	-	-.4	-	-.4	06
-.5	-	-.5	-	-.5	0
-.6	1267	-.6	0	-.6	5
-.7	2	-.7	-	-.7	
-.8		-.8	-	-.8	
-.9		-.9	4	-.9	
Mdn*	.06		.36		.35
Wtd mean**	.10		.31		.27

Note. Department-level coefficients are specified by combining the initial digit (with decimal) in each row with successive digits in the same row. In the distributions of correlations between GRE verbal scores and first-year average, for example, coefficients were .72, .79, .67, .50, and so on. Data are for engineering, math and physical science, and economics samples.

* Median of distribution of 86 department-level coefficients.

** Size-adjusted averages of department-level coefficients. These coefficients indicate the relationship between departmentally standardized (z-scaled) predictor and criterion variables for 1,216 ESL students in the 86 quantitative departments.

Distributions of Department-Level GRE General Test Validity Coefficients for Foreign ESL Students, by Size of Sample: Data for 86 Departments in Primarily Quantitative Fields

	GRE Verbal vs FYA			GRE Quantitative vs FYA			GRE Analytical vs FYA		
	N < 10	N= 10-14	N = 15 plus	N < 10	N = 10-14	N = 15 plus	N < 10	N = 10-14	N = 15+
.8				2			4		
.7	29			36	269		8	6	
.6	-		7	367	0344578	1	039	0179	34
.5	03458	4	1	4467	0019	256	68	0245	9
.4	-	0237	-	356	22577788	050	467889	1333357	467
.3	-	0245558	247	6	126	3667	449	2589	1136678
.2	13	13	05779	48	0	2899	23	245	013368
.1	33	9	16689	6	0	12446	03	67	36
.0	12344	7	458	14	5688	028	669	-	56
-.0	568	56	2378	2	0	29	3	2	-
-.1	29	012248899	04	0038	0		9	34	8
-.2	27	2445	7	-	-		-	669	-
-.3	4	9		8	3		-	0	-
-.4	-	-		-			6	-	0
-.5	-	-		-			-	0	
-.6	67	12		0			5		
-.7	2			-					
-.8				-					
-.9				4					
Mdn	.03	-.05	.16	.36	.47	.28	.39	.40	.31
Depts.	(27)	(34)	(25)	(27)	(34)	(25)	(27)	(34)	(25)

Note. Department-level coefficients are specified by combining the initial digit (with decimal) in each row with successive digits in the same row. In the distributions of correlations between GRE verbal scores and first-year average, for example, coefficients of .72, .79, .50, .53, .54, .55, .58, and so on, were obtained for departments with fewer than 10 (between 5 and 9) ESL students; GRE quantitative coefficients for samples of fewer than 10 students were .82, .73, .76, .63, .66, .67, 54, .54, .56, .57, and so on.