DOCUMENT RESUME

ED 381 559 TM 022 895

AUTHOR Wilson, Kenneth M.; Tillberg, Rebecca
TITLE An Assessment of Selected Validity-Related Properties of a Shortened Version of the Secondary Level English Proficiency Test and Locally Developed Writing Tests in the LACCD Context.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RR-94-54
PUB DATE Nov 94
NOTE 116p.
PUB TYPE Reports - Research/Technical (143) -- Statistical Data (110) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Academic Achievement; *College Students; Community Colleges; Correlation; Educational Assessment; *English (Second Language); Scores; Scoring; *Student Placement; Test Construction; Test Format; *Test Use; Test Validity; Two Year Colleges; *Writing Tests
IDENTIFIERS *Los Angeles Community College District CA; *Secondary Level English Proficiency Test

ABSTRACT

A study was undertaken to obtain direct empirical evidence for the validity and usefulness of English as a second language (ESL) placement of a shortened version of the Secondary Level English Proficiency (SLEP) Test being used in the Los Angeles Community College District (LACCD) and locally developed and scored writing tests. The LACCD provided scores on the shortened version of the SLEP and the writing samples, grades in English, and background data for over 10,000 students. Patterns of performance and concurrent relationships among scores on components of the LACCD placement battery were noted. Observed levels of correlation among scores and student performance in courses (grades earned) by course and college, and in various subgroups, were analyzed. Also considered was the extent to which observed relationships were influenced by nonvalidity-related factors. Findings provided direct evidence that the shortened SLEP and locally developed writing tests were providing valid information regarding ESL proficiency, and that they appear to be an effective basis for placing students within practical time constraints. One exhibit, 22 tables, and 6 figures present study data. Three appendixes contain an additional nine tables of supplemental information. (Contains 23 references.) (Author/SLD)

RESEARCH REPORT

AN ASSESSMENT OF SELECTED VALIDITY-RELATED PROPERTIES OF A SHORTENED VERSION OF THE SECONDARY LEVEL ENGLISH PROFICIENCY TEST AND LOCALLY DEVELOPED WRITING TESTS IN THE LACCD CONTEXT

Kenneth M. Wilson
Educational Testing Service

with the collaboration of
Rebecca Tillberg
Los Angeles Community College District

Educational Testing Service
Princeton, New Jersey
November 1994

An Assessment of Selected Validity-Related Properties of A Shortened Version of the Secondary Level English Proficiency Test and Locally Developed Writing Tests in the LACCD Context

Kenneth M. Wilson
Educational Testing Service

with the collaboration of
Rebecca Tillberg
Los Angeles Community College District

Educational Testing Service

November 1994

ABSTRACT

This is the report of a study that was underaken to obtain direct empirical evidence regarding aspects of the validity and usefulness for ESL placement of (a) a shortened version of the Secondary Level English Proficiency (SLEP) Test, being used for ESL placement by colleges in the Los Angeles Community College District (LACCD), and (b) locally developed and scored writing tests. The LACCD Central Office provided scores on the shortened version of the SLEP and the writing samples, grades in ESL courses, and background data (gender, language, educational status, and so on) for over 10,000 students. This report documents and evaluates

- patterns of performance on the shortened SLEP and the writing tests in the general ESL population and in selected demographic subpopulations,

- concurrent relationships among scores on the components of the LACCD placement battery,

- observed levels of correlation between scores on the shortened SLEP test, the writing test, and the placement composite, on the one hand, and student performance in ESL courses, as indexed by grade earned (a grade on the "A-F" scale, or a Pass/Fail grade), on the other, by course and by college, and in various subgroups (e.g., gender, educational level, age, language), and

- the extent to which observed relationships in placed samples are influenced by non-validity-related factors (for example, differential restriction of range on the tests that were used to place students, sample size, and type of grading system).

The findings of this collaborative undertaking provide direct empirical evidence that the shortened SLEP and locally developed writing tests are providing valid information regarding related aspects of ESL proficiency in the demographically diverse ESL student population being served by the LACCD; and the findings logically extend available evidence supportive of the validity of the Secondary Level English Proficiency Test for ESL assessment purposes in the LACCD and elsewhere. Based on study findings, the local writing tests and the shortened SLEP appear to be providing an effective basis for placing students, within time constraints that appear to be considered necessary, from an administrative perspective. Research is needed to address questions regarding the extent to which use of the full-length version of the SLEP would enhance the overall validity of placement. Some pertinent lines of inquiry are suggested.

EXECUTIVE SUMMARY

Study Context

To facilitate placement of students for instruction in English as a second language (ESL) in courses differentiated according to seven ESL proficiency levels, colleges in the Los Angeles Community College District (LACCD) consider scores on a shortened version of the Secondary Level English Proficiency (SLEP) test (e.g., ETS, 1988, 1991), and locally developed writing tests--writing samples, involving locally selected topics and rated locally, but elicited under standard time limits.

The SLEP was selected by the LACCD ESL Committee, as the standardized-test component of the placement battery, after a rigorous, comparative review of

(a) the content, psychometric properties and costs of nine commerically available ESL proficiency tests (see Appendix A.2 for data on the tests that were considered), and

(b) empirical findings (Butler, 1989) indicating that SLEP items were at a psychometrically appropriate level of difficulty for LACCD ESL students, and that SLEP scores were positively related to independently established ESL placement levels.[1] The review process and the considerations involved in arriving at decisions as to how and to what extent the SLEP would be modified for use in the LACCD, are described in detail by Butler (1989).

Objectives of the Present Study

The work reported herein was undertaken with the encouragement and support of the SLEP Testing Program at Educational Testing Service, and the LACCD central office, represented throughout by Ms. Rebecca Tillberg, to obtain direct empirical evidence regarding aspects of the validity and usefulness for ESL placement of (a) the shortened version of the SLEP used in the LACCD, and (b) the locally developed and scored writing tests. More specifically, the principal objectives of the study were

(a) to document and evaluate patterns of performance on the shortened SLEP and the writing tests in the general ESL population and in selected demographic subpopulations,

---

[1] Butler noted that both the SLEP and the Pre-TOEFL, an easier version of the familiar Test of English as a Foreign Language, were attractive because of the communicative orientation of the item types, "consistent with the communicative focus of the ESL curriculum." Cost considerations favored the SLEP (Butler, 1989: pp. 4-5).

(b) to investigate levels and patterns of concurrent relationships among scores on the components of the LACCD placement battery,

(c) to obtain base-line empirical evidence regarding observed levels of correlation between scores on the shortened SLEP test, the writing test, and the placement composite, on the one hand, and student performance in ESL courses, as indexed by grade earned (a grade on the "A-F" scale, or a Pass/Fail grade), on the other, by course and by college, and in various subgroups (e.g., gender, educational level, age, language), and

(d) to provide an analytic assessment of the extent to which observed relationships in placed samples are influenced by non-validity-related factors (for example, differential restriction of range on the tests that were used to place students, sample size, and type of grading system).

Given the data at hand, only the validity properties of the shortened version of the SLEP and the writing test are illuminated directly by this investigation.

- However, the full-length version of the SLEP can be expected to exhibit in the LACCD context, levels of concurrent, discriminant, and criterion-related validity equal to or greater than the levels that may be found to obtain for the shortened version of the SLEP that is currently being used in the LACCD.

Thus, an important incidental objective of the study was to extend available evidence of the validity of the full-length SLEP test.

General Analytical Rationale

A general outline of elements in the overall analytical rationale that guided inquiry involving the respective objectives is provided below.

Concurrent and Discriminant Validity

The placement tests (shortened LC, shortened RC, Essay) are designed to measure listening, reading, and writing skills, respectively, thought of as distinguishable but related aspects of general ESL proficiency.

- Moderate to moderately high levels of intercorrelation among such measures would constitute evidence of concurrent validity.

• If the shortened SLEP reading score is found to be more closely correlated with Essay rating than is the shortened SLEP listening score, this would constitute evidence of discriminant validity for the shortened version of the SLEP, namely, evidence that the LC and RC sections are measuring psychometrically distinguishable aspects of proficiency.

Implications of Concurrent Validity Findings for Reliability

It was not feasible to address directly questions regarding the reliability of either the shortened SLEP or the local writing tests in the specific samples under consideration.

• Internal consistency reliability estimates for item-type subsections of the full-length SLEP are available for samples involved in test development--.89 and .81, respectively, for the two item-type subsections included in the shortened LC section, and .71 and .89, respectively, for the two reading item-type sections in the LACCD (shortened) version.

• Based on available estimates and by inference from decrease in length, it is assumed that the shortened SLEP sections are somewhat less reliable than their full-length counterparts: internal consistency estimates of .94 for LC and .93 for RC (.96 for total).

• Inter-rater reliabilities for the writing test are assumed to be lower than those for the shortened SLEP, and perhaps not consistent across rating sites, but no data on inter-rater reliability are available for citation.

However, it is important to keep in mind that useful general, reliability-related inferences can be drawn from observed correlations among variables.

• Because the level of correlation between tests is limited by their respective reliabilities, if either the shortened SLEP or the writing test has very low reliability, generally speaking, the scores involved cannot correlate very highly with each other or with any other study variable. Hence, moderate to moderately strong concurrent validity coefficients, for example, are indirectly indicative of "useful" levels of reliability.

Problems Specific to the Writing Tests

The possibility of systematic differences across colleges with

respect to "rater bias" is inherent in any context involving the generation of ratings in multiple sites, even though a common scale is employed. Butler (1989) called attention to anomalous differences, by college, in the average essay ratings of students at the same ESL placement level.

- Such differences may reflect substantive differences in the average level "writing ability" of the students being assessed on the respective campuses, but systematic differences in rating standards may also be involved. SLEP (objective test) scores, of course, are free of this type of bias.

- The presence of statistically significant differences by college in level of Essay rating relative to SLEP performance, would constitute a necessary condition for inferring differences in rating standards.

- Differences in all correlational findings involving writing samples may to some extent be associated with differences across colleges in rater-reliability, topics selected, and so on.

Need to Investigate Properties of Placement Composites

Results of the foregoing lines of inquiry bear directly on the validity properties of the shortened SLEP and the local writing tests. In practice, the information provided by the three measures (two SLEP scores and the Essay rating) is combined using nominal weights for percent-right transformations of the raw scores on the shortened LC and RC sections, respectively, and the Essay. The actual contribution of two or more variables to a composite does not necessarily correspond to the nominal weights because of differences in score variabilities.

- In samples from City, Harbor, Mission, Trade, and Valley, some two-thirds of the variance in the placement composite should be associated with SLEP performance, and one-third with essay performance (nominal weights for Essay, LC and RC were, respectively, .34, .33, and .33, in samples under consideration).

- In samples from East and Pierce, some 60 percent of composite variance should be associated with essay performance, and only 40 percent with SLEP performance (nominal weights for Essay, LC and RC were .60, .20, and .20, respectively).

Regression-based estimates of the relative contribution of the test scores to composites were obtained for each college, and compared with the expected relative contribution, based on the nominal weights.

Considerations in Assessing Criterion-Related Validity

In the LACCD context, all tests used are expected to exhibit "reasonable" levels of correlation with appropriate criteria of performance in the courses of study to which the testing is related--average criterion-related validity coefficients in the neighborhood of .35 have been suggested as representing a reasonable standard (California Community Colleges, 1992: p. 20).

Although a "standard" is proposed, it is specifically recognized that

- "the magnitude of the correlation may vary as a function of the degree to which a tests was used to place students in the course under investigation and/or the variation in grading standards across classrooms" (California Community Colleges, 1992: p. 20).

Empirical evidence based on LACCD samples, regarding typical levels and ranges of test-criterion correlation--either for current placement tests or similar tests--does not appear to be available. Accordingly, analysis of test/grade correlations in the present study was designed primarily to

- generate base-line, empirical findings regarding levels and patterns of test/grade correlation for the shortened SLEP and the writing test in the LACCD context, and

- assess the extent to which effects associated with use of the tests in placement, influence both levels and patterns of observed test/grade correlations in ways that limit generalization from the observed correlations regarding test validity in the general ESL population being assessed.

Study Sample, Data, and General Procedure

The foregoing objectives were pursued, using data provided by the LACCD Central Office, for more than 10,000 students who were assessed for placement in Spring 1991, Fall 1991 and Spring 1992, respectively. Data were available for City, East, Harbor, Mission, Pierce, Trade, Valley and West--all LACCD colleges except Southwest.

Student records contained scores on the placement tests, scores on composites computed for placement on the respective campuses, data needed to classify students according to specific ESL courses and corresponding subject areas, official end-of-course grades (e.g., A-F, Pass/Fail, Withdrew), and self-reported background data: gender, educational status (e.g., less than high school graduate, high school graduate, college graduate), ethnicity

(national origin and ethnic group), date of birth, and native language (11 categories, of which nine were specific languages).

- Information provided in college catalogs (especially from City) helped guide classification of courses by ESL level.

- Analyses concerned with objectives other than those requiring the assessment of test/grade relationships, were conducted using data for the total (unrestricted, pre-placement) samples from the respective colleges.

- Only students with clearly quantifiable grades (that is, a grade of A, B, C, D, or F, or a Pass/Fail grade), in courses represented by at least 10 students with a grade and scores on the placement tests, were included in analyses concerned with criterion-related validity issues (test/grade correlations).

Organization of the Main Report

Section I of the study report provides a general overview of the process through which the foregoing placement tests were identified and selected, (b) generally reviews evidence bearing on the validity of the full-length SLEP in various use contexts, including the LACCD), (c) reviews the considerations involved in decisions to use only a portion of the full-length SLEP test and in selecting particular SLEP items for inclusion in the shortened version, and (d) provides a brief overview of study objectives and the data provided by the LACCD central office.

Section II provides a general description of the performance of the general ESL student population and several demographic subpopulations on the shortened SLEP and the writing tests.

Section III reports findings regarding concurrent- and discriminant-validity properties of the shortened SLEP and the locally developed writing tests, along with findings bearing on the comparability of the writing tests across colleges. Results of analyses of the relative contribution of SLEP scores and writing test scores in placement composites are also reported in Section III, along with illustrative evidence of the extent to which use of tests in placement introduces range-restriction-related effects that complicate validity-related interpretations of correlation coefficients computed within samples of placed students.

Section IV reports and evaluates findings regarding test/grade correlations by course, by college, and for subgroups defined by gender, language, educational status, and age. The extent to which the observed coefficients are influenced by nonvalidity-related variables (sample standard deviations, sample size, and type of

grading system) is demonstrated.

Section V provides a general review and evaluation of the findings.

Overview of Study Findings

The ESL population being assessed with the shortened SLEP is extremely heterogeneous with respect to age, educational level, ethnicity, and language background.

- The average student is about 29 years of age; based on data of record on date of birth, some students were less than 13 years old and others were over 65 years of age when assessed in the LACCD.

- The majority of students hold secondary-school diplomas earned in schools located outside the United States, but the sample includes individuals who have not completed secondary school--some having enrolled in adult education courses, or classes for special students--as well as individuals with bachelor's or higher degrees.

- The largest linguistic subpopulation is made up of native speakers of Spanish (about 44 percent of the total); none of the other directly identified language groups accounts for much more than 10 percent (11 percent reported Armenian, and less than 1 percent reported Filipino).

The shortened SLEP sections are at psychometrically efficient levels of difficulty for the general LACCD ESL population as well as for subgroups that differ markedly with respect to age, educational level, language background, and self-reported ethnic group membership.

- Results of an analysis of trends in mean scores on the shortened SLEP and the Essay, respectively, across subgroups classified by educational level indicate that the two means covaried directly. Similar patterns of covariation are discernible for the two test means across other demographic subgroups--that is, demographic subgroups with higher (lower) means on the SLEP tend to have higher (lower) means on the writing test.

Concurrent relationships between shortened SLEP scores and scores on the writing tests were moderate to strong in each college setting:

- within-college correlations between scores on the locally developed writing tests and scores on the

shortened SLEP centered at about .60, and varied between roughly .5 and .7 across colleges. These correlations were observed in general (pre-placement) samples.

- The relationship between the shortened SLEP reading comprehension score and Essay rating typically was somewhat stronger than that observed for the shortened SLEP listening comprehension score with Essay--consistent with plausible expectation that reading and writing skills, should tend to be somewhat more closely related than are listening and writing skills.

- At the same time, coefficients for total score on the shortened SLEP test (LC + RC) were somewhat larger than those for RC alone, indicating that each of the two sections may be providing some unique information regarding aspects of the ability being measured by the writing test--a face valid measure of ability to write comprehensibly in English.

- Levels of conconcurrent relationship observed for college-level samples were also observed in the results of analyses conducted by college, illustratively, for Spanish-speaking and Korean-speaking subgroups, respectively.

The foregoing findings provide strong empirical evidence of concurrent- and discriminant-validity properties for the shortened SLEP and the writing tests.

In this connection, it is considered particularly noteworthy that the levels of concurrent relationship observed in the LACCD context, between scores on the shortened version of the SLEP and scores on local writing tests (with attendant differences in topic, rating procedures, and so on), are comparable to levels that have been found to obtain between scores on the Test of English as a Foreign Language and the Test of Written English (writing samples elicited under standard conditions, scored under controlled conditions by at least two raters),

(a) in samples tested in developmental research (Carlson, Bridgeman, Camp, and Waanders, 1985) involving the TOEFL and prototypical versions of the TWE, and

(b) in samples taking both tests under fully operational conditions (see ETS, 1992b). **Among other things, such findings attest to the power of direct observation of pertinent linguistic behavior in ESL proficiency assessment.**

Analyses of levels and patterns of correlation of the placement tests with course grades, involving 59 course-level samples, were complicated due to the need to assess the influence

of nonvalidity-related factors (for example, restriction of range effects) on both levels and patterns of observed test/grade coefficients. The generalizations outlined below reflect findings reviewed in detail and evaluated in Section 4.

- Scores on the shortened SLEP and the writing test, respectively, and as combined for placement purposes in different colleges, tend to be positively correlated with course grade, but there is substantial variability across courses in level of relationship. Distributions of observed course-level coefficients for the respective tests tend to center at about the .2 level, with substantial variability (ranging upward from negative values to values above .5).

- Regression results indicated substantial agreement between actual and nominal weights for the placement tests in composites.

- Essay scores received relatively more weight than did SLEP scores in placement at East and Pierce, and in samples from those colleges, average coefficients were higher for the SLEP than for the Essay. SLEP was weighted relatively more heavily than Essay in other college samples, and in these samples, Essay coefficients typically were higher than SLEP coefficients. Such findings are interpretable as reflecting differential restriction of range effects.

- Differences across 59 course-level samples, with respect to level of test/grade coefficients for the placement tests (that is, SLEP/grade coefficients, Essay/grade coefficients, and Composite/grade coefficients, respectively), were found be associated relatively strongly with nonvalidity-related sample characteristics: differences in the corresponding sample standard deviations, type of grading system employed (A/F versus Pass/Fail), and sample size.

The respective test/grade coefficients for the 59 course-level samples, treated as dependent variables, were regressed on the five nonvalidity-related variables noted above, in order to assess their relative contribution when treated as a battery of predictors.

- Sample standard deviation for the shortened SLEP was the most highly weighted variable not only in analyses involving course-level SLEP/grade coefficients but also in analyses involving Composite/grade coefficients.

- The most important predictor of level of Essay/grade correlation was **type of grading system**; SLEP standard deviation was the second largest predictor.

• The Composite standard deviation did not contribute to level of Essay/grade correlation in the stepwise regressions, after other variables were entered.

Regression equations developed in the foregoing analysis provided a basis for estimating test/grade coefficients (SLEP/grade, Essay/grade, and Composite/grade coefficients, respectively), assuming no selection on the variables involved. Values for the total LACCD (general) sample standard deviations were substituted for the restricted values in the respective equations.

Estimated coefficients and the corresponding observed coefficients are shown below.

Test/grade correlations

	Short SLEP	Essay	Composite
A/F sys (est.)	.47	.35	.50
A/F sys (obs.)	.17	.20	.22
P/F sys (est.)	.43	.21	.42
P/F sys (obs.)	.18	.07	.19

Generally speaking, the findings involving test/grade relationships that have been reviewed should be thought of primarily as a basis for empirical assessment of the extent to which observed course-level test/grade coefficients generally are influenced by nonvalidity-related factors, and are thereby limited as bases for inferences regarding "levels or patterns of validity" for the tests involved when applied in the general population under consideration.

In analyses for subgroups, positive test/grade correlations varying around the observed values noted above were found to obtain in samples aggregated by gender, age, educational level, and language, respectively.

Regarding the Writing Tests

• Pooled within-school SLEP/Essay coefficients typically were larger than those observed for the combined sample, the latter being attenuated by inconsistent ordering of school means on the respective tests--an anomalous finding, in view of the systematic, positive within-school relationships observed.

• College differences in "rating standards" are suggested by the results of analyses of differences across colleges with respect to mean discrepancies between observed Essay rating, and Essay rating estimated from SLEP performance,

using a total-sample regression equation. Mean discrepancies (residuals) were substantially more pronounced in analyses by college than in analyses for subgroups by gender, language, educational level, and age, respectively, aggregated across colleges.

- In view of apparent differences across sites in "rating standards," pending resolution of questions regarding comparability of Essay scores across colleges, future investigations of subgroup differences in the relative development of Essay-assessed "writing ability" and SLEP-assessed skills, need to be conducted at the college level.

These and related issues constitute meaningful topics for further inquiry. The findings suggest the need for systematic study of the rating process on each campus, including the collection of data needed to evaluate degree of agreement among raters with respect to level as well as rank-order.

Concluding Observations

The findings of this collaborative undertaking provide direct empirical evidence that the shortened SLEP and locally developed writing tests, are providing valid information regarding related aspects of ESL proficiency in the demographically diverse ESL student population being served by the LACCD; and the findings logically extend available evidence supportive of the validity of the Secondary Level English Proficiency Test for ESL assessment purposes in the LACCD and elsewhere.

Considered in light of the rigorous screening and evaluation of available psychometric options that resulted in SLEP's initial selection by the ESL Committee, and other evidence of SLEP's validity cited herein--including the generally positive findings reported by Butler (1989), regarding SLEP's appropriateness for ESL assessment purposes in the LACCD--the findings that have been reviewed herein both support and extend the working proposition advanced at the outset, namely, that

- the SLEP is a valid test of psychometrically distinct, albeit closely related, aspects of acquired proficiency in English as a second language: the ability to comprehend utterances in English, and the ability to read and comprehend the substance of material written in English,

- the SLEP can be expected to provide reliable and valid information regarding these abilities, when used for ESL assessment purposes by secondary schools, colleges, and other institutions, and

- given the demographic diversity of the LACCD samples involved in this study, the findings extent available evidence indicating that the SLEP can be used validly with local ESL student populations that are quite heterogeneous with respect to age, educational level, language background, and national origin.

The local writing tests and the shortened SLEP appear to be providing an effective basis for placing students, within time constraints that appear to be considered necessary, from an administrative perspective.

As noted by Butler (1989), some cost in terms of diminished reliability and validity undoubtedly is entailed by the use of only a selected portion of the test. Available evidence bearing on this important question is limited, but consistent with the validity-cost assumption.

- In a cooperative study currently in progress (Wilson, 1993), involving large samples of Japanese-speaking students being assessed for ESL placement at Temple University-Japan, LACCD-parallel subscores were computed post hoc from data available for the full-length SLEP. Correlations between these scores and ratings of speaking ability and writing ability, respectively, were approximately .05 correlation points lower than correlations for corresponding full-length scores, that centered around .6.

- Rudmann (1991), at Irvine Valley College, reports observed correlations centering around .40 between grades in ESL courses and scores on the full-length SLEP.

Analyses involving the full-length SLEP test clearly would be needed in order to address questions of comparative validity: comparative validity of the current shortened, LACCD version of the SLEP, and the full-length test, and/or questions as to the comparative validity of the particular SLEP item types selected versus those not selected.[2]

---

[2] In connection with the latter issue, it is noteworthy that one LACCD college (West) opted to use only the shortened SLEP reading section, with the writing test, for placement. More generally, in reporting on the basic SLEP validation study, Stansfield (1984) observed that users might elect to test either listening comprehension or reading comprehension in some circumstances (e.g., listening comprehension for level-placement decisions involving courses emphasizing the development of conversational skills). An evaluation of experience at West, where only a score on items from the SLEP "reading comprehension" ability domain is used, would be useful.

One empirical approach to such an assessment that offers a variety of pertinent data-generating possibilities, would involve administering the full-length SLEP (Form 1 or one of the other available equated forms) as an exit test, to students who were tested for placement with the shortened version.

- The resulting item-level data for both tests would provide a basis for assessing average (net) change in performance on the particular item-types now being used, and make it possible to evaluate the relative validity of all eight SLEP item types.

- Moreover, end-of-course distributions of scores, needed to establish measured levels of proficiency associated with particular ESL instructional sequences, could be generated.

Generally speaking, the use of more extensive and/or comprehensive testing procedures in ESL placement can reasonably be expected to be accompanied by benefits attendant upon reduced incidence of perceived misplacement: for example, reduction in the educational and administrative costs associated with course changes, improved satisfaction with results of the placement process on the part of both teachers and students, and improved retention of students in the ESL program.

A formal evaluation of the ESL placement process would provide empirical evidence that is pertinent to cost/benefit analysis, as well as to an overall evaluation of the ESL placement program in the LACCD--for example, assessing the extent to which teachers and students, respectively, are satisfied with current placement procedures and/or results, collecting and evaluating data on the incidence of shifts in courses.

Acknowledgments

The work reported herein represents the results of a cooperative research undertaking, involving the SLEP Testing Program at Educational Testing Service and the Los Angeles Community College District, in a joint effort to obtain empirical evidence needed to address questions of mutual interest regarding the validity of tests, including a shortened version of the SLEP test, being used for ESL placement in the LACCD.

The SLEP Testing Program was represented by Ms. Stella Cowell, Program Director, who provided encouragement and support that was basic to the successful completion of the work. Ms. Rebecca Tillberg, Research Analyst, who represented the Los Angeles Community College District, actively collaborated in the joint enterprise by providing data files, detailed information regarding the context of the study, and insights regarding the functioning of ESL placement programs in the LACCD, in a variety of personal communications during the course of the study.

Brent Bridgeman and Spencer Swinton provided helpful reviews of drafts of the report.

These contributions are acknowledged with appreciation. Only the writer, however, is responsible for the contents of the report.

TABLE OF CONTENTS

SECTION I. CONTEXT OF THE STUDY

The nine colleges that make up the Los Angeles Community College District (LACCD) offer a wide range of courses in English as a second language (ESL) for nonnative-English speaking students and prospective students. The colleges are designated herein as City, East, Harbor, Mission, Pierce, Southwest, Trade, Valley, and West, respectively. The programs offered by these colleges conform generally to a curricular model that was developed by the LACCD ESL Committee during 1987-88, calling for placement of students in one of seven proficiency levels, based on results of placement testing and other pertinent considerations (Butler, 1989). As described by Butler, the model did not incorporate the use of a standard exit test--for example, alternate forms of the placement test.

Since 1988, placement of students according to ESL levels has been guided by scores on a shortened version of the Secondary Level English Proficiency (SLEP) test (e.g., ETS, 1988, 1991), and locally developed writing tests--writing samples, involving locally selected topics and rated locally, but elicited under standard time limits.

The work reported herein was undertaken, using data provided by the LACCD central office, to obtain empirical evidence regarding aspects of the validity and usefulness for ESL placement of (a) the shortened version of the SLEP used in the LACCD--and logically to extend the body of available empirical evidence bearing on the validity of the full-length SLEP test--and (b) the locally developed and scored writing tests.[1]

This section (a) provides a description of the process through which the foregoing placement tests were identified and selected, (b) generally reviews evidence bearing on the validity of the full-length SLEP in various use contexts, including the LACCD), (c) reviews the onsiderations involved in decisions to use only a portion of the full-length SLEP test and in selecting particular SLEP items for inclusion in the shortened version, and (d) provides a brief overview of study objectives and the data provided by the LACCD central office.

Section II provides a general description of the performance of the general ESL student population and several demographic subpopulations on the shortened SLEP and the writing tests. Section III reports findings regarding concurrent- and discriminant-validity properties of the shortened SLEP and the locally developed writing tests, along with findings bearing on the

---

[1] The work was undertaken with the encouragement and support of the SLEP Testing Program at Educational Testing Service, and the LACCD central office, represented throughout by Ms. Rebecca Tillberg.

comparability of the writing tests across colleges. Results of analyses of the relative contribution of SLEP scores and writing test scores in placement composites are also reported in Section III, along with illustrative evidence of the extent to which use of tests in placement introduces range-restriction-related effects that complicate validity-related interpretations of correlation coefficients computed within samples of placed students.

Section IV reports and evaluates findings regarding test/grade correlations by course, by college, and for subgroups defined by gender, language, educational status, and age. The extent to which the observed coefficients are influenced by non-validity related variables (sample standard deviations, sample size, and so on) is demonstrated. Section V provides a general review and evaluation of the findings.

Placement Testing in the LACCD

Placement testing in the LACCD involves the collection of a writing sample and administration of a shortened version of the Secondary Level English Proficiency (SLEP) test--a standardized, multiple-choice test with sections measuring the ability to comprehend utterances in English (listening comprehension) and the ability to comprehend written material in English (reading comprehension), respectively, in samples of nonnative-English speakers (for example, Educational Testing Service [ETS], 1988, 1991). Brief descriptions of the writing test and the full-length SLEP, the process through which the SLEP was originally selected, and the procedures involved in shortening the SLEP, are provided below.

The Writing Sample. Generally speaking, writing samples have face validity as measures of ability to write comprehensibly in English. Problems involved in the use of writing samples have to do with the subjective nature of the process through which the samples are scored, difficulty in obtaining a representative sample of writing skills, and so on.[2] For example, the rank-ordering of samples according to a particular rating scale may differ from rater to rater; and raters giving similar rank-orders to a set of samples may differ systematically with respect to scale-level assignments; performance may not be consistent across topics.

In the LACCD context, writing samples are collected and scored locally; essay topics are also locally developed. However, samples are scored according to a common, holistic rating scale, involving eight defined levels ranging from 0 thru 7 (see Appendix A.1).

---

[2] For a comprehensive review of problems involved in assessing "writing ability" in samples of ESL speakers, and the complex procedures required to deal with these problems, see Carlson, Bridgeman, Camp, & Waanders (1985).

According to Butler (1989: p. 7), " . . . the scale was developed by Dr. Minette Lanier, Pierce College, and Dr. Lloyd Thomas of the LACCD District Office, with suggestions made by members of the ESL Committee."

Studies designed to assess the reliability and validity of writing samples do not appear to have been conducted in the LACCD context. However, Butler reported " . . .a wide range of (average) performance in writing for the same (ESL placement) level at different colleges" (1989: p. 18), citing relatively marked differences between average essay ratings for Pierce (typically about "5" on the holistic scale) and Mission (typically about "3" on the scale). Differences in average ratings across "scoring sites," may reflect substantive differences in writing skills, but site differences in "rating standards" cannot be ruled out as contributing factors. Raters at some sites may tend to rate given samples higher(lower) than do raters at other sites.

The SLEP. The SLEP was developed by the Educational Testing Service for use by secondary-level institutions in screening and/or placing nonnative-English speaking applicants, to be tested in ETS-operated centers worldwide. When the latter practice was discontinued, the SLEP was made available to qualified users for local administration and scoring.

The basic validation study (Stansfield, 1984) that accompanied introduction of the SLEP, provided evidence that SLEP scores were related positively to ESL placement classification, years of study of English as a second language, time spent in the U.S., and other criteria, in samples from over 50 U.S. secondary schools. Moderate to relatively high concurrent relationships have been reported (e.g., ETS, 1991) for SLEP scores with scores on the Test of English as a Foreign Language [TOEFL] (e.g., ETS, 1992a), used primarily with college-level ESL students, and the Maculaities Assessment Program (Maculaities, 1982), used primarily with pre-college-level students (see Von Schilling, 1988).

The SLEP is currently being used worldwide, in a variety of assessment contexts, with both traditional and nontraditional students, not only in secondary-school settings but also in post-secondary settings. Approximately one-third of the respondents to a survey of SLEP users (Wilson, 1993) reported use of the SLEP with college-level students, or students in adult education classes: to assess readiness to undertake English-medium academic instruction, for placement in ESL courses, for course or program evaluation, admission screening, and so on.

Face validity considerations, empirical evidence of SLEP's psychometric properties, such as that cited illustratively above, and positive reports of SLEP's validity-related properties from ESL professionals using the SLEP in secondary and postsecondary settings, constitute what appears to be a conceptually persuasive

basis for asserting as a working proposition, that

> • the SLEP is a valid test of psychometrically distinct, albeit closely related, aspects of acquired proficiency in English as a second language, namely, the ability to comprehend utterances in English, and the ability to read and comprehend the substance of material written in English, and can be expected to provide reliable and valid information regarding these abilities in samples of literate ESL speakers at relatively diverse educational levels.

The SLEP was selected as the standardized-test component of the placement battery, based on a comprehensive, comparative review by the ESL Committee of the content, psychometric properties, and costs of nine commercially available ESL proficiency tests (see Appendix A.2 for data on the tests that were considered).[3] The review process and the considerations involved in arriving at decisions as to how and to what extent the SLEP would be modified for use in the LACCD, are described in detail elsewhere (Butler, 1989). As noted specifically from time to time, descriptions of the process herein are based on Butler's detailed report.

Shortening the SLEP. The SLEP includes a total of 150 multiple-choice questions in two sections--listening comprehension and reading comprehension--of 75 items each. The time required for the entire test is approximately 85 minutes: just under 40 minutes, paced by recorded prompts, for the listening section and 45 minutes for the reading section. Reusable test booklets are available for three equated forms of the SLEP (Forms 1, 2, and 3). To date, only Form 1 has been used in the LACCD (Tillberg, personal communication, 1992).

Eight different item types are represented in the SLEP, four in the listening section and four in the reading section, as follows: Listening section (Single Pictures, Dictation, Map, Conversations), and Reading section (Cartoons, Four Pictures, Cloze, and Reading Passage). Illustrative items are provided in Appendix A.3.

The decision to use a shortened version of the SLEP test was dictated primarily by time considerations--the total amount of time required for both the SLEP and a writing sample was deemed to be

---

[3] Butler noted that both the SLEP and the Pre-TOEFL, an easier version of the familiar Test of English as a Foreign Language, were attractive because of the communicative orientation of the item types, "consistent with the communicative focus of the ESL curriculum." Cost considerations favored the SLEP (Butler, 1989: pp. 4-5).

excessive.[4] To obtain empirical evidence regarding the performance of LACCD ESL-students on the SLEP, and to evaluate performance on subscores based on SLEP item types, Butler (1989) analyzed item-type and section-level scores for the entire test, in Spring 1988, for samples from the nine LACCD campuses.

Information regarding characteristics of the LACCD samples involved (e.g., age, educational status, gender, language group) was not reported. However, it is pertinent to note that the distributions of percent-right scores for the respective test-sections were judged by Butler (1989: p. 25) to indicate that the SLEP items were not inappropriately "easy" or "difficult" for LACCD samples in various proficiency levels.

Data shown in Table 1.1 indicate that average percent-right scores for LACCD samples studied by Butler were generally comparable to those reported by ETS (internal memoranda) for the samples of international students involved in the development of different forms of the SLEP. Table 1.1 includes average percent-right scores for a sample of native-English speakers (G7-12 samples from several Florida schools, tested by Holloway [1982]). Also shown in the table are estimated internal consistency reliability coefficients for the respective SLEP item-type subsections, from unpublished internal ETS test analyses (see Attachment A).

- o SLEP items are very easy for native-English speakers, and judging from the LACCD data, for example, they become increasingly more difficult as the independently assessed level of English proficiency of the test-taking sample decreases--suggesting SLEP's validity as a measure of proficiency in English as a second language generally, as well as in samples from the LACCD context.

Selecting item types for the shortened test. As reported by Butler (1989), decisions regarding item types to be included in the shortened version were guided by feedback from the respective campuses regarding the perceived appropriateness of item content, and the extent to which scores based on the respective item types discriminated among the several proficiency levels represented in the sample; also the progression in difficulty for the four reading item types in the LACCD sample--see data for last four item types in Table 1.1, for all the samples involved. After evaluating all considerations, the ESL Comittee decide to use a version of the SLEP that included four of the eight item-type sections, namely, Single Picture and Dictation item types from the Listening section,

---

[4] As reported by Butler (1989: p. 10), ". . . word came back from every campus that the counselors would be extremely reluctant to agree to what they considered to be a lengthy (2 hour) testing period" (p. 10).

Table 1.1. Comparative Performance of LACCD Samples and Designated E1 and E2 Samples on SLEP Item-Type Subsections: Mean Percent Right Scores

SECTION/	Samples*					
	E1	E2a	E2b	E2c Min	E2c Max	Estimated reliability
LISTENING						
Single picture	97	85	81	(53	83)	.811
Dictation	93	90	84	(48	87)	.889
Map	97	74	69	(44	76)	.721
Extended conversations	93	68	62	(35	69)	.840
READING						
Cartoon	94	89	87	(63	91)	.742
Four pictures	91	78	84	(55	82)	.707
Cloze	81	62	62	(29	68)	.891
Literary passage	64	48	56	(6	43)	.685

Note. E1 is used to denote English as the first or primary language; E2 is used to denote English as a second language.

* E1 data are from Holloway (1984), total sample (G7-12) of students from selected Florida schools. E2a and E2b data are from ETS unpublished internal analyses for international-student samples in 1980 (E2a) and 1988 (E2b) used in test development and equating. The E2a data are for Form 1 of the SLEP, the form being used in the LACCD. E2c data are from Butler (1989), for samples of ESL students from LACCD colleges. The minimum (Min) percent-right scores are for ESL students in the most elementary of seven ESL levels defined for the system; the maximum (Max) percent-right means are for those at the highest level.

** From unpublished internal test analyses (ETS, 1980).

and Four Picture and Cloze items from the Reading section. Time limits and item composition of the shortened SLEP and the writing sample are indicated in Table 1.2. Butler (1989) called attention to the potential loss of placement-related information associated with the elimination of a substantial number of SLEP items, and noted that " . . . it may be necessary to reevaluate the choice of SLEP Test subsections" (p. 2).

Table 1.2 Components of the Recommended Placement Battery

Test	No. items	Time (mins.)
Writing Sample	1	30
Short SLEP		
Listening	45	20
One Picture	25	
Dictation	20	
Reading		35
Four Pictures	15	
Cloze	40	
Total (short SLEP)	100	85

Note. See Appendix A.3 for illustrative items.

Guidelines were developed for using the tests in placement, including (a) use of a "standard" placement composite, and (b) recommeded composite-score ranges for proficiency levels, to facilitate uniform placement practices across colleges.

The recommended standard placement composite was to be derived using "nominal" weights of .34(Essay %rite), .33 (LC %rite), and .33 (RC %rite).[5] Score ranges for seven proficiency levels were suggested. It was recommended that "additional information such as length of time in the United States, number of years of instruction in English, educational goals, etc., should also be taken into consideration in determining where a student can best benefit from language instruction" (Butler, 1989, p. 38).

Colleges were free to modify the weighting procedure that was suggested. And, based on information provided by the LACCD office (Tillberg, personal communication, 1992), two colleges (East and Pierce) subsequently decided to use nominal weights of .60, .20, and .20, respectively, for percent-right scores on Essay, Listening, and Reading. A third college (West) initially used the standard (.34,.33,.33) weighting for components of the full

[5] The effective (as opposed to nominal) weighting of two or more variables in a composite will vary as a function of the dispersions and intercorrelations of the variables involved. "If we really want to weight tests in a battery equally we should apply to each one a weight inversely proportional to its standard deviation. Without information as to the validities of the tests and of their intercorrelations, that would be a reasonable thing to do" (Guilford, 1965, p. 424). Such information was not available for the LACCD sample studied by Butler (1989).

placement battery, but subsequently decided to use only the Reading score and the Essay score in the placement composite, with nominal weights of .60 for percent-right scores on Essay and .40 for percent-right scores on Reading.

The Present Study

According to information provided by the LACCD District Office (Tillberg, 1992, personal communication), informal feedback from member colleges suggests that those concerned with ESL placement, by and large, have been reasonably satisfied with the placement procedures that include scores on the test battery described generally above. Of course, the extent to which the positive reaction indicated by this informal feedback is "representative," can only be determined by a formal assessment of degree of satisfaction with the placement procedures--among students as well as among teachers.[6] Such an assessment is beyond the scope of the present inquiry, which is concerned with assessing selected validity-related properties of the current components of the placement battery--that is, the shortened SLEP and the writing test--in samples of ESL students from eight LACCD colleges: City, East, Harbor, Mission, Pierce, Trade, Valley, and West.

More specifically, the study was undertaken with the encouragement and support of the SLEP Testing Program at Educational Testing Service, and the collaboration of the LACCD District Office, with the following objectives:

(a) to examine patterns of performance on the shortened SLEP and the writing tests in the general ESL population and in selected demographic subpopulations;

(b) to investigate levels and patterns of concurrent relationships among scores on the components of the LACCD placement battery;

(c) to obtain base-line empirical evidence regarding observed levels of correlation between scores on the shortened SLEP test, the writing test, and the placement composite, on the one hand, and student performance in ESL courses, as indexed by grade earned (a grade on the "A-F" scale, or a Pass/Fail grade), on the other, by course and by college, and in various subgroups (e.g., gender, educational level, age, language), and

---

[6] To the extent that shifts in course-placement based on teacher (and/or student) perceptions of "misplacement" early in a term are permitted, incidence of change in courses would provide a pertinent element in an overall assessment of the adequacy of placement.

(d) to provide an analytic assessment of the extent to which observed relationships in placed samples are influenced by non-validity-related factors, especially differential restriction of range on the tests that were used to place students.

Given the data at hand, only the validity properties of the shortened version of the SLEP can be illuminated directly by this investigation. However, on logical grounds, the full-length version of the SLEP can be expected to exhibit, in the LACCD context, levels of concurrent, discriminant, and criterion-related validity equal to or greater than the levels that may be found to obtain for the shortened version of the SLEP that is currently being used in the LACCD. Accordingly a more general objective of the study is to extend evidence of the validity of the full-length version of the test.

Questions of Concurrent and Discriminant Validity

The placement tests (LC, RC, Essay) are designed to measure listening, reading, and writing skills, respectively, thought of as distinguishable but related aspects of general ESL proficiency.

> If moderate levels of intercorrelation among the measures are found to obtain, this would constitute evidence of their concurrent validity.
>
> If the shortened SLEP reading score is found to be more closely correlated with Essay rating than is the shortened SLEP listening score, this would constitute evidence of discriminant validity for the shortened version of the SLEP, namely, evidence that the LC and RC sections are measuring distinguishably different aspects of proficiency.

It was not feasible to address directly questions regarding the reliability of either the shortened SLEP or the local writing tests. Internal consistency reliability estimates for item-type subsections of the SLEP were reported above (see Table 1.1), for the SLEP form currently in use in the LACCD.

> Based on available estimates and by inference from decrease in length, it is assumed that the shortened SLEP is not as reliable as is the full-length version: internal consistency estimates are .939 for LC, .930 for RC, and .962 for the full length total score, for Form 1.
>
> Inter-rater reliabilities for the writing test are assumed to be lower than those for the shortened SLEP, and perhaps not consistent across rating sites, but as indicated, above, no data on inter-rater reliability are available for citation.

However, it is important to keep in mind that useful general, reliability-related inferences can be drawn from observed correlations among variables. For example, because the level of correlation between tests is limited by their respective reliabilities, if either the shortened SLEP or the writing test has very low reliability in samples such as those being assessed in the LACCD, the scores involved cannot correlate very highly with each other or with any other study variable.

The possibility of systematic differences across colleges with respect to "rater bias" is, of course, inherent in any context involving the generation of ratings in multiple sites, even though a common scale is employed, and Butler (1989) called attention to anomalous differences, by college, in the average essay ratings of students at the same ESL placement level. As noted earlier, such average differences by college may reflect substantive differences in the average level "writing ability" of the students being assessed on the respective campuses but systematic differences in rating standards may also be involved. SLEP (objective-test) scores, of course, are free of this type of bias. Accordingly, an analysis was made of differences between observed essay ratings and estimated ratings based on SLEP scores.

> The presence of statistically significant differences by college in level of Essay rating relative to SLEP performance, constitutes a necessary condition for inferring differences in rating standards.

Questions regarding possible effects associated with differences in essay reliability across colleges could not be addressed. However, differences by college in all correlational findings involving writing samples may to some extent be associated with differences in reliability or content (e.g., topics selected), and so on.

The foregoing lines of inquiry were concerned with validity properties of the shortened SLEP and the local writing tests. In practice, the information provided by the three measures is combined using nominal weights for percent-right LC, RC, and Essay scores. Based on the differences in patterns of nominal weighting, noted above, certain patterns of outcomes are expected to obtain:

> In samples from City, Harbor, Mission, Trade, and Valley, some two-thirds of the variance in the placement composite should be associated with SLEP performance, and one-third with essay performance.
>
> In samples from East and Pierce, on the other hand, some 60 percent of composite variance should be associated with essay performance, and only 40 percent with SLEP performance. At West, as indicated above, for samples tested after spring 1991, the composite should be

expected to reflect differences in essay performance (60 percent) more than differences in reading ability only (40 percent).

Regression-based estimates of the relative contribution of the test scores to composites were obtained for each college, and compared with the expected relative contribution, based on the nominal weights.

Assessing Criterion-Related Validity

In the LACCD context, all tests used are expected to exhibit "reasonable" levels of correlation with appropriate criteria of performance in the courses of study to which the testing is related--average criterion-related validity coefficients in the neighborhood of .35 have been suggested as representing a reasonable standard (California Community Colleges, 1992: p. 20).

Although the foregoing "standard" is proposed, it is specifically recognized that "the magnitude of the correlation may vary as a function of the degree to which a test was used to place students in the course under investigation and/or the variation in grading standards across classrooms" (California Community Colleges, 1992: p. 20). Empirical evidence regarding typical levels and ranges of test-criterion correlation--either for current placement tests or similar tests--in the LACCD context does not appear to be available.

Accordingly, analysis of test/grade correlations in the present study was designed primarily to

(a) generate base-line, empirical findings regarding levels and patterns of test/grade correlation for the shortened SLEP and the writing test in the LACCD context, and

(b) assess the extent to which effects associated with use of the tests in placement, influence both levels and patterns of observed test/grade correlations in ways that limit generalization regarding test validity from observed within-course correlations.

Data and General Study Procedures

A data file containing over 10,000 records was provided by the LACCD District Office.[7] The students involved were assessed for

---

[7] The file contained records for an undetermined number of students who earned grades in more than one course. These records were treated as "unique" for study purposes. However, the number of students assessed is, to some extent, less than the number of records.

placement in Spring 1991, Fall 1991, and Spring 1992, respectively.

The following data elements were included in the file:

A. test and demographic data

1. scores on the placement battery,

2. the composite score that was used to guide placement,

3. self-reported background data (birthdate, language, educational status, ethnicity);

B. data on student performance, namely,

1. course grades (grades, either A-F or pass/fail) or some other performance-related index (e.g., incomplete, withdrawn); and

C. course identification data, including

1. codes identifying course in which a student was most recently enrolled (not necessarily the same as the course in which the student was originally placed), by subject areas designated differentially by college:

a. Developmental Communication, English, and Speech Communication [hereafter, Speech] at City, East, Harbor, Pierce, and Valley), and

b. English as a Second Language [ESL 1 through ESL 6] at Mission, Trade, and West).

Course codes for Mission, Trade, and West (ESL ["400"] courses), directly indicated the proficiency levels associated with the correspondingly numbered courses. However, for the other colleges this was not the case.

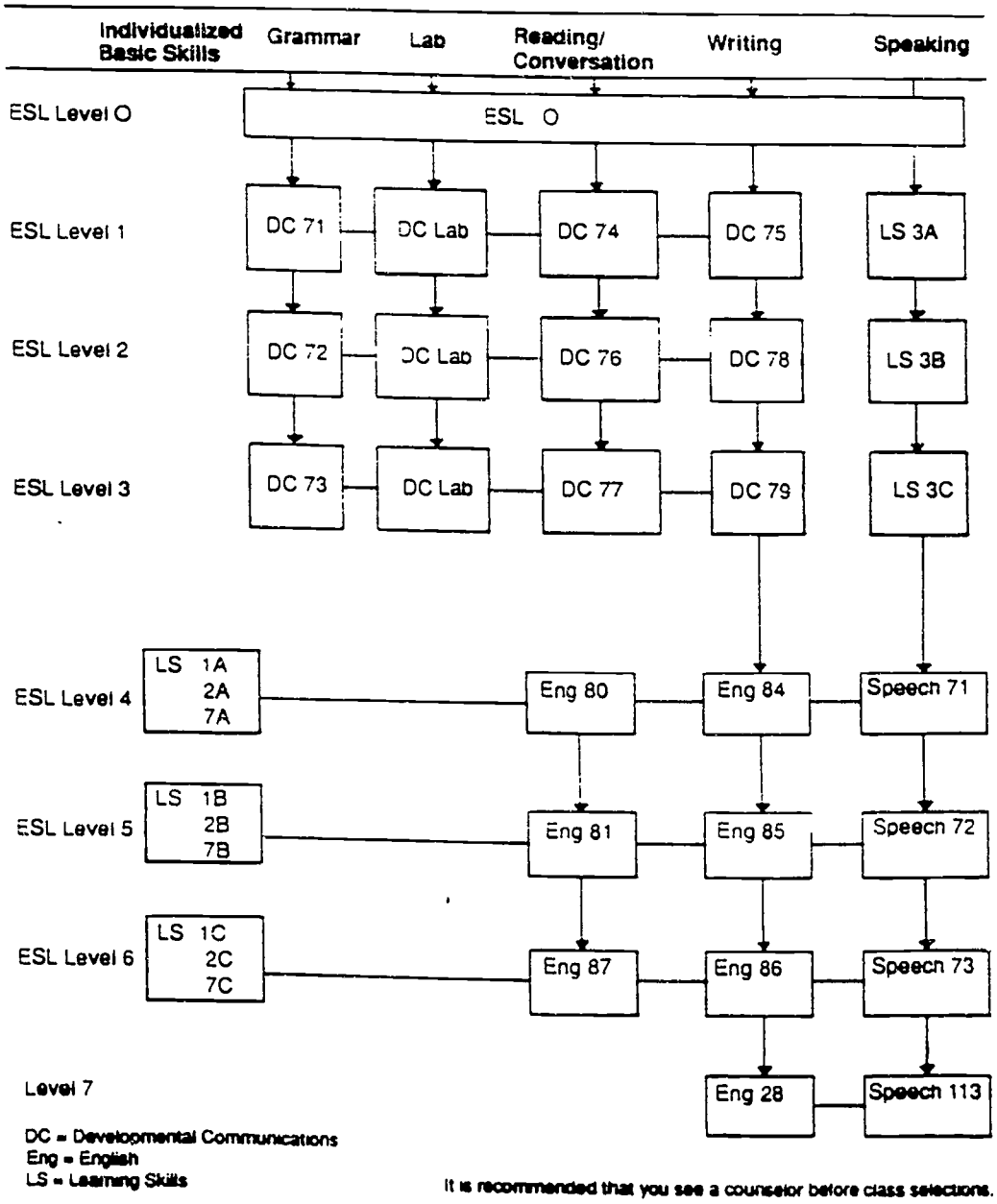
Level-classification of courses for City, East, Harbor, Pierce, and Valley Colleges, respectively, was guided primarily by the "English As A Second Language Flow-chart" shown as Exhibit A Exhibit A. English As A Second Language Flow Chart (City College) (from a recent edition of the City College catalog). Differences in ESL curricular-emphasis (grammar, reading/conversation, writing, or speaking) are noted in the chart for differently numbered courses within each level.

Section II provides a description of levels and patterns of performance on the shortened SLEP and the writing tests in the general LACCD population and selected subpopulations; Section III provides evidence regarding concurrent- and discriminant-validity properties of the tests involved; Section IV reviews findings

Exhibit A. English As A Second Language Flow Chart (City College)

ENGLISH AS A SECOND LANGUAGE FLOWCHART

ESL



regarding test/grade correlations; and findings are reviewed and evaluated in Section V.

In presenting and discussing findings and procedures, an effort is made to limit the amount of supporting detail considered directly in the text; such detail is provided in designated appendices.

SECTION II. PERFORMANCE OF ESL STUDENTS ON THE SHORTENED SLEP AND THE WRITING TESTS

This section reports the results of analyses that provide an overview of performance on the placement tests for ESL students generally; also performance of students classified by college and by selected demographic characteristics (gender, educational status, age, ethnic group, and language).

Attention is directed first to descriptive statistics for the total sample. Differences in performance on the study variables for students attending the respective colleges are then considered. Finally, descriptive statistics are provided for ESL students classified by gender, educational status, age category, ethnic group identify, and language.

The Total Sample

Descriptive statistics are shown in Table 2.1 for designated study variables, based on data for total LACCD sample. Means and standard deviations, and minimum and maximum values, shown for each variable, are based on the total number of cases with observations on the corresponding variable. In some situations, especially at West, only the shortened SLEP reading section was administered. It can be seen that some 6,500 (about 60 percent) of the originally assessed students were sufficiently "persistent" in the ESL program, to earn a grade (on either the A-F or Pass/Fail scales) in an ESL course.

Average percent-right scores for the SLEP are shown to provide a basis for inferences regarding the difficulty level of the respective sections and the combined score for the sample as a whole. The mean percent right score for the essay is shown for perspective only--inferences from percent-right scores regarding "appropriate level of difficulty" for the essay cannot be thought of as comparable to corresponding inferences from percent-right score on a standardized test such as the shortened SLEP.

> The average percent right score of approximately 68 for the shortened LC score indicates that this section is of about "middle difficulty" for the sample, and the corresponding mean of about 50 percent for the RC section indicates that the section is of somewhat greater than middle difficulty for the sample. The total shortened SLEP mean (LC + RC) of 58 translates directly into a percent-right equivalent--indicating that the items selected for the shortened SLEP are generally just below middle difficulty (approximately 67 percent for four-option, rights-scored multiple-choice tests), a difficulty level that is psychometrically efficient for measurement purposes.

Table 2.1. Descriptive Statistics for Study Variables, All LACCD (Spring 1991 through Spring 1992)

Variable*	Mean	Std Dev	Minimum	Maximum	Valid N
LC	30.45	9.88	1	45	9935
RC	27.24	10.75	1	55	10695
SLEP	57.96	18.98	3	100	9858
ESSAY	3.19	1.46	0	8	10786
LCPCT	67.66	21.96	2.2	100	9935
RCPCT	49.54	19.55	1.8	100	10695
ESSPCT	39.83	18.20	.0	100	10786
GRADEAF	2.63	1.19	0	4	5055
PASSFAIL	.68	.47	0	1	1526
AGE	29.87	9.01	10.3	83.5	10789

Note. See Appendix B.1 for full distributions of LC, RC, SLEP (shortened version), and Essay scores; see Appendix B.2, for descriptives statistics by college.

* LC Listening Comprehension, short SLEP
RC Reading Comprehension, short SLEP
SLEP LC + RC, short SLEP (100 items)
Essay Writing sample
LCPCT Percent-right conversion of LC
RCPCT Percent-right conversion of RC
ESSPCT Percent-right conversion of Essay
GRADEAF Grade on the A to F scale (4,0)
PASSFAIL Grade on the pass/fail (1,0) scale

A similar conclusion was reached by Butler (1989: p. 25), based on evidence indicating that none of the independently defined proficency-level subgroups studied . . . " 'topped out' or 'bottomed out'" (p. 25) on either the listening comprehension or the reading comprehension sections as a whole, or on the respective item-type subsections. Full distributions of scores on the shortened SLEP sections and the writing sample are shown in Appendix B.1.

Differences By College

Table 2.2 shows means and standard deviations of distributions of scores on the shortened SLEP and the Essay rating, respectively, by college. Differences in average performance are portrayed graphically in Figure 2.1. Note that mean Essay rating has been multiplied by 10, and that the eight colleges are arrayed in descending order (left to right) with respect to mean score on the shortened SLEP (LC+RC).

Butler (1989: p. 18-19) noted that there were differences by college in average essay rating for students at the same, independently defined levels of proficiency. It can be seen in Figure 2.1, that

(a) differences among colleges with respect to mean Essay rating are much more pronounced than are differences with respect to mean SLEP performance, and

(b) mean essay rating does not vary systematically with mean SLEP score across colleges--for example, three of the four highest means for the Essay are registered by the four colleges with the lowest-scoring students on the shortened SLEP.

As will be seen in Section III, scores on the shortened SLEP and Essay rating covary quite systematically within the respective colleges. Lack of corresponding covariation among the college means on these variables, suggests that raters at different colleges tend to have different "rating standards." Such differences introduce interpretive complications in analyses of average performance on the Essay relative to performance on the SLEP when data are aggregated across colleges.

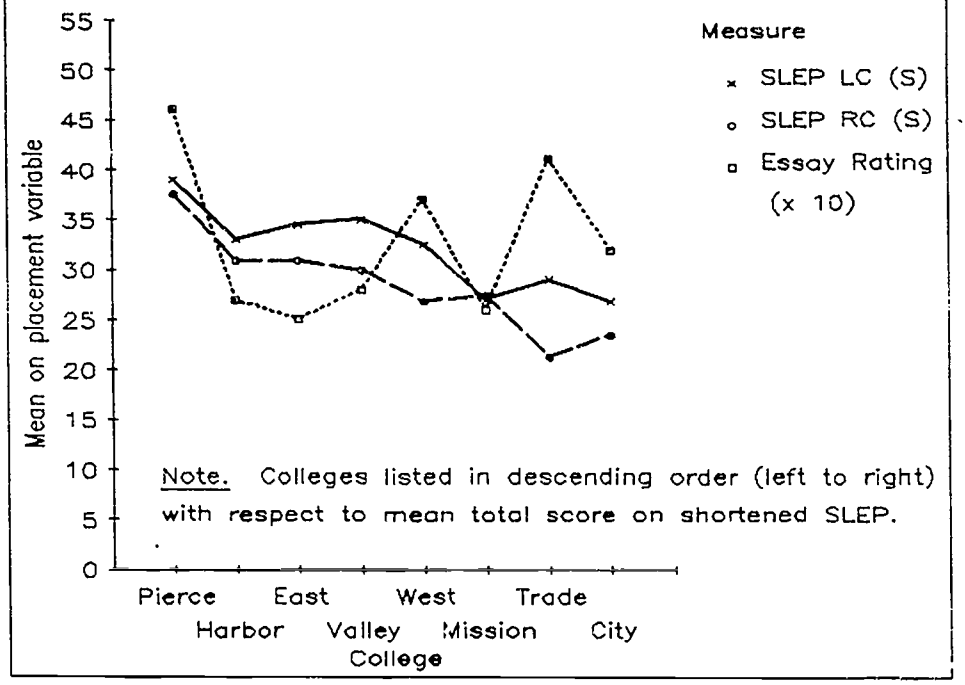
Performance by Demographic Categories

In the study (Butler, 1989) that was conducted as part of the process through which the SLEP was selected as the standardized-test component of the LACCD placement battery, demographic characteristics were not considered. Butler noted, simply (p. 5) that " . . . many of the students who enroll in LACCD classes have recently come from high schools . . .." Table 2.2 and Figure 2.1 show differences by college in performance on the basic study variables.

Table 2.2. Summary Statistics By College

College Code		N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
City	1	3872	27	10	24	10	51	19	3.2	1.5
East	2	942	35	7	31	10	65	15	2.5	.9
Harbor	3	431	35	7	33	9	68	15	2.9	1.2
Mission	4	1717	27	11	28	11	56	20	2.7	1.4
Pierce	5	493	39	5	38	9	77	13	4.7	1.3
Trade	6	477	30	7	23	8	53	13	4.5	1.4
Valley	7	1616	35	7	30	10	65	15	2.9	1.3
West	8	1241	32	9	27	10	59	18	3.7	1.2
Total		9855	31	10	27	11	58	19	3.1	1.5

Figure 2.1. Performance of students on placement variables, by college



Note. Colleges listed in descending order (left to right) with respect to mean total score on shortened SLEP.

Data shown in Tables 2.3 through 2.7 indicate that the ESL population being assessed with the shortened SLEP is extremely heterogeneous with respect to age, educational level, ethnicity, and language background. Mean scores on the shortened SLEP and the writing sample are provided for subgroups classified by gender (Table 2.3), educational status (Table 2.4), age group (Table 2.5), ethnic group (Table 2.6) and language background (Table 2.7).

Figure 2.2 highlights diversity in educational level, and diversity in language background.

o The majority of students hold secondary-school diplomas earned in schools located outside the United States, but the sample includes individuals who have not completed secondary school--some enrolled in adult education courses, some as special students--as well as individuals with bachelor's or higher degrees.

o The largest linguistic subpopulation is made up of native-speakers of Spanish (about 44 percent of the total); none of the other directly identified language groups accounts for much more than 10 percent (11 percent reported Armenian, and less than 1 percent reported Filipino).

Lack of systematic covariation across colleges between means for the writing sample and means for the SLEP has already been noted (see Figure 2.1 and related discussion, above), suggesting that site-differences in rating standards may be involved. Depite such local differences, Figure 2.3 shows a systematic pattern of covariation between shortened SLEP means and Essay means across subgroups classified by educational level (see Table 2.3 for detail).

• The subgroups are arrayed (left to right) in descending order with respect to mean score on the shortened SLEP. Note that the mean Essay rating has been multiplied (arbitrarily) by 20. The pattern shown in Figure 2.3 indicates that average level of performance on the Essay tends to decrease as average level of SLEP performance decreases.

• The consistent relationship indicated in Figure 2.3 suggests that subgroup members at "higher-rating" colleges were balanced by members of the same subgroups at "lower-rating" colleges, hence the theoretically consistent set of findings for the variables involved in a general sample <u>despite</u> local differences in rating standards. Generally similar trends (positive covariation between SLEP mean and Essay mean) can be discerned by inspecting the means reported in the various tables.

Table 2.3. Summary Statistics by Gender

Gender	Code	N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
Male	1	4250	32	9	29	11	61	18	3.3	1.4
Female	2	5605	29	10	26	11	55	19	3.1	1.5
Total		9855	31	10	27	11	58	19	3.1	1.5

Table 2.4. Summary Statistics By Reported Educational Status

Educational status	Code	N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
Missing	0	1061	27	11	26	11	54	20	2.8	1.4
Special student	1	805	34	9	31	11	65	18	3.5	1.5
Adult school< HS	2	134	28	10	24	10	52	18	2.9	1.5
Adult school < HS	3	926	27	10	24	10	51	17	2.9	1.4
Unenrolled < HS	4	846	29	11	26	11	55	20	2.9	1.5
GED/HS equiv	5	259	33	9	32	11	65	18	3.5	1.4
CA HS certificate	6	31	33	11	33	11	67	21	3.4	1.4
Foreign HS dipl	7	5804	31	10	27	10	57	18	3.2	1.4
Associate degree	8	194	34	9	31	10	65	18	3.6	1.4
Bachelor's/plus	9	729	34	8	33	10	68	16	3.7	1.4
Total		9855	31	10	27	11	58	19	3.1	1.5

Table 2.5. Summary Statistics by Age Group

Age group	Code	N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
Below 22.5 yrs		2187	34	8	30	10	64	16	3.4	1.4
22.5 - 26.4 yrs		1922	32	9	28	10	60	18	3.2	1.5
26.5 - 30.4 yrs		1780	31	10	27	11	58	19	3.2	1.5
30.5 - 36.4 yrs		1968	30	10	26	11	56	19	3.0	1.4
36.5 yrs & older		1998	27	11	25	12	52	20	2.9	1.5
Total		9855	31	10	27	11	58	19	3.1	1.5

Table 2.6. Summary Statistics by Reported Ethnic Group Membership

Ethnic group	Code	N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
Missing	0	209	31	10	29	11	61	19	2.8	1.5
Chinese	10	635	35	7	31	9	66	15	3.1	1.3
Japanese	11	617	37	5	32	7	69	11	3.8	1.2
Korean	12	869	34	6	30	8	64	13	3.6	1.4
Laotian	13	31	25	9	23	7	48	15	2.7	1.4
Cambodian	14	15	26	12	26	11	51	22	3.5	1.8
Vietnamese	15	321	31	8	28	10	59	16	3.4	1.5
Indian(Sub)	16	70	32	10	31	13	63	21	3.8	1.4
Other Asian	19	697	32	9	27	10	59	17	3.2	1.4
Black	20	124	28	7	25	8	52	14	3.6	1.4
Filipino	30	91	39	4	39	7	78	10	4.0	1.4
Mexican*	40	2022	29	10	28	11	57	19	2.9	1.5
Central Amer	41	1599	28	11	25	11	54	20	3.2	1.5
South Amer	42	654	31	10	28	11	59	20	3.3	1.4
Other Hispanic	49	818	28	11	26	11	55	20	3.2	1.5
Caucasian	50	1758	28	11	23	11	51	20	3.0	1.5
American Indian	60	3	40	1	27	9	66	10	3.7	.6
Other Pacific	79	56	38	5	33	10	71	14	4.1	1.4
Other Non-White	80	87	35	9	26	10	62	18	3.3	1.3
Decline	90	113	35	8	30	11	65	18	3.4	1.5
Total		9855	31	10	27	11	58	19	3.1	1.5

Table 2.7. Summary Statistics by Language

Language	Code	N	Shortened SLEP LC M	LC Sd	RC M	RC Sd	SLEP M	SLEP Sd	ESSAY M	ESSAY Sd
Missing	0	607	32	9	29	11	61	19	3.3	1.6
English	1	192	33	9	31	11	65	18	3.3	1.5
Armenian	2	1216	25	10	19	9	44	17	2.6	1.4
Chinese	3	607	34	7	31	9	65	15	3.1	1.3
Farsi	4	252	36	7	28	10	64	15	3.3	1.3
Filipino	5	80	38	6	38	8	76	11	3.9	1.4
Japanese	6	604	37	5	32	7	69	10	3.8	1.2
Korean	7	858	34	6	30	8	64	13	3.6	1.4
Russian	8	354	32	9	26	10	58	17	3.3	1.4
Spanish	9	4722	29	10	27	11	56	20	3.1	1.5
Vietnamese	10	297	31	8	28	10	59	15	3.4	1.5
Other	11	1000	35	8	30	11	66	17	3.6	1.5
Total		9855	31	10	27	11	58	19	3.1	1.5

Figure 2.2. Distribution of sample by educational status and by language background

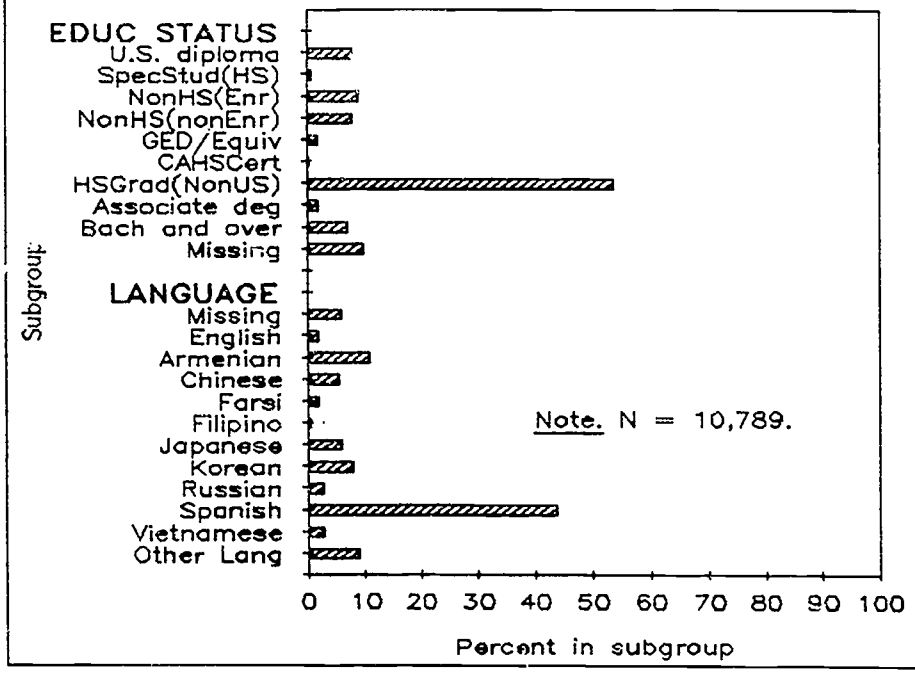
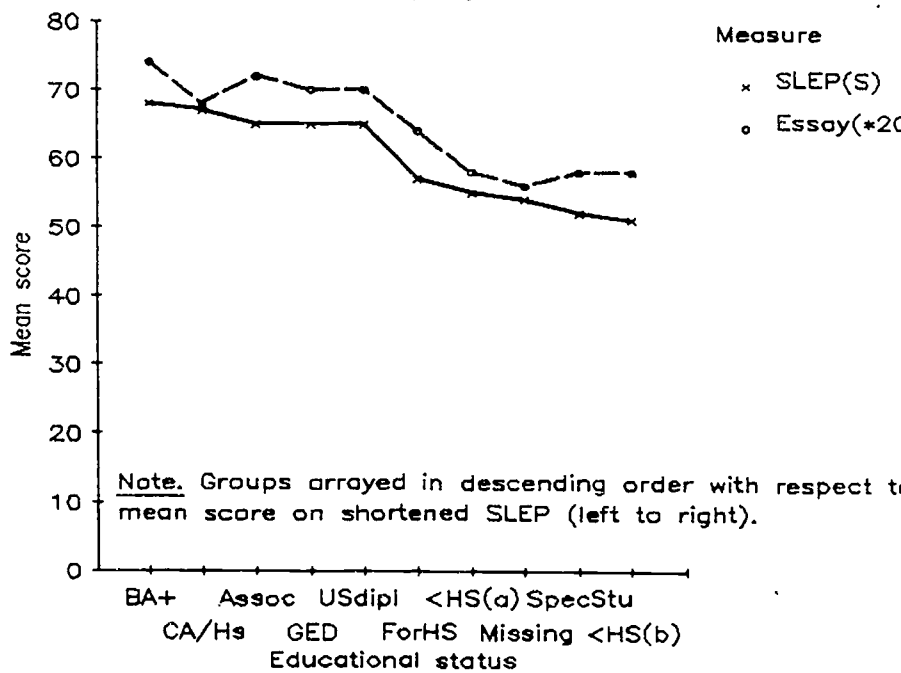


Figure 2.3. Trends in means on the Essay and the shortened SLEP, respectively, by educational status



The findings summarized above indicate that

(a) LACCD colleges are serving an ESL student population that is remarkably heterogeneous with respect to age, educational status, ethnicity, and native language,

(b) the shortened SLEP sections appear to be at psychometrically efficient levels of difficulty for the general LACCD ESL population as well as for subgroups that differ markedly with respect to age, educational level, language background, and self-reported ethnic group membership, and

(c) subgroup means on the shortened SLEP and the Essay, respectively, exhibit a systematic, positive relationship--true for subgroups classified by educational level, and for other demographic subgroups as well.

Related Findings

Some related findings are provided for additional perspective in Appendix B.2 (test performance by ESL placement level and by college), Appendix B.3 (test performance by language group and by college), and Appendix B.4 (average differences in test performance by gender, age, and educational level, by college).

Regarding the latter, it is useful to note the general pattern of findings:

- females tend to earn lower scores on the placement variables than do males,
- older students tend to earn lower scores than do younger students, and
- students with less formal education tend to earn lower scores than do students with more formal education.[8]

The findings summarized in this section suggest that both the shortened SLEP and the local writing tests are providing valid information regarding individual and group differences in the LACCD context, a conclusion that is reinforced by findings in subsequent sections of this report.

---

[8] It is noteworthy that the general negative relationship between age and test performance is stronger for listening comprehension than for either reading comprehension or writing ability (essay rating). Pursuit of these and other theoretically and pragmatically interesting questions is beyond the scope of the present study.

SECTION III. CONCURRENT- AND DISCRIMINANT-VALIDITY PROPERTIES OF THE SHORTENED SLEP AND THE WRITING TESTS

SLEP listening comprehension items have face validity as measures of the ability to comprehend utterances in English, and SLEP reading comprehension items are face valid measures of the ability to comprehend the meaning of material that is written in English (see Appendix A.3). Writing samples have general face validity as measures of ability to write comprehensibly using English.

Based on previous research, moderate to relatively strong levels of correlation have been found to obtain among both direct and indirect measures of the four basic language macroskills--including speaking, not here under consideration--in samples of educated ESL users/learners (e.g., Pike, 1979; Oller, 1983: passim; Hale, 1986; Wilson, 1989), and in samples studying languages other than English, as foreign languages (see especially, Carroll, 1967).

In a comprehensive study (Carlson, Bridgeman, Camp, and Waanders, 1985) associated with development of the Test of Written English (TWE), holistic scores for ESL-writing samples were found to be somewhat more closely related to scores on the nonlistening portions of the TOEFL (Structure and Written Expression [SWE] and Reading Comprehension and Vocabulary [RCV]), respectively, than to the TOEFL Listening Comprehension score.

> With some variation in detail (by essay topic) the SWE and RCV scores correlated approximately .6 with essay rating while TOEFL LC scores correlated at approximately the .5 level. Inter-rater reliabilities centering around .70 were reported (as estimates of reliability " . . . if only the scores from one judge are to be used operationally" (p. 57).

The SLEP Manual (e.g., ETS, 1991) reports evidence indicating that scores on the SLEP test are relatively strongly related to scores on the TOEFL.

> More specifically, SLEP Listening Comprehension (LC) score correlated more highly with TOEFL LC score (.74) than with either the TOEFL Structure and Written Expression (SWE) score or the TOEFL Vocabulary and Reading Comprehension (RC) score; SLEP Reading Comprehension (RC) score correlated equally highly with TOEFL LC (r = .80) and TOEFL RC (r = .79); the correlation between the two total scores in this sample was .82, slightly lower than that between SLEP RC and TOEFL Total (.85). These relationships obtained in samples of ESL students in intensive ESL programs on four U. S. college campuses.

Because of the strong relationship between the two tests, it is a reasonable inference that the SLEP test shares to some extent TOEFL's validity-related properties--properties that have been extensively investigated (see, for example, ETS, 1992a). The findings reported below support such an inference.

Findings Regarding Concurrent and Discriminant Validity

Concurrent correlations of scores on the shortened SLEP test with Essay scores, computed by college, are shown in Table 3.1. Standard deviations of scores on the shortened SLEP test and the writing samples involved are also shown. Several aspects of the findings are noteworthy:

(1) Concurrent relationships were moderate to strong in each setting.

(2) The RC/Essay relationship typically was stronger than the LC/Essay relationship, consistent with expectation--that is, reading and writing skills should tend to be more close-ly related than listening and writing skills. At the same time, coefficients for total score on the shortened SLEP test (LC + RC) were somewhat larger than those for RC alone, indicating that the items in the shortened LC section are providing some unique information regarding aspects of the ability being measured by the writing test.[9]

(3) The pooled, within-school coefficients typically were larger than the "total sample" coefficients, the latter being attenuated by lack of systematic covariation in the corresponding means across schools--assumed to reflect non-validity-related differences in "rating standards."

Concurrent relationships were stronger and the patterning of relationships--RC/Essay coefficients higher than LC/Essay coefficients--was more consistent with expectation, in samples assessed at City, Mission, West, and Pierce, than in samples assessed at East, Harbor, Valley or Trade.

---

[9] It is worth noting in this connection that Dictation items included in the shortened LC section are highly integrative, in that in order to identify the written sentence that matches the spoken stimulus, examinees must read four sentences. See illustrative items in Appendix A.3. This particular issue is beyond the scope of the present study, but is a fruitful one for further investigation.

Table 3.1. Concurrent Correlations of Scores on the Shortened SLEP with Essay Score, By College

College	N	Correlation with essay LC	RC	LC+RC
City	3,323	.64	.67	.70
East	951	.45	.51	.53
Harbor	514	.49	.48	.51
Mission	1,965	.56	.58	.60
Pierce	511	.48	.57	.58
Trade	648	.40	.37	.44
Valley	1,641	.43	.47	.51
West	1,271	.53	.57	.59
TOTAL	10,590	.49	.49	.53
WITHIN*		.55	.57	.61

	SLEP r	Standard deviation** LC	RC	SLEP	Essay
Mission	.59	11	11	20	1.4
City	.70	10	10	19	1.5
West	.59	9	10	18	1.2
Valley	.51	7	10	15	1.3
East	.53	7	10	15	.9
Harbor	.51	7	9	15	1.2
Pierce	.58	5	9	13	1.3
Trade	.44	7	8	13	1.4
TOTAL		10	11	19	1.5

Note. These are "missing data" correlations: each coefficient reported is based on data for all cases with observations on the pair of variables involved.

* The coefficients in this row are, in effect, size-adjusted means of the coefficients for the schools.

** Colleges are listed in descending order with respect to standard deviation of scores of the shortened SLEP test (LC + RC).

Three of the four highest (lowest) SLEP/Essay coefficients were for samples from one of the four highest-ranking (lowest-ranking) colleges with respect to SLEP standard deviation.

The general pattern of differences by college in levels of concurrent relationship is discernible in the results of analyses conducted by college for Spanish-speaking and Korean-speaking subgroups, respectively, reported in Table 3.2.

By and large, the levels of concurrent relationships observed in general samples by college tend to be present in college-level samples for the two language groups involved.

The findings reported above attest to the validity of both the shortened SLEP and the local writing tests.

It is particularly noteworthy in this connection that the levels of concurrent relationship observed in the LACCD context, between scores on a shortened version of the SLEP and scores on local writing tests--with attendant differences in topic, rating procedures, and so on--parallel levels that have been found to obtain between scores on the Test of English as a Foreign Language and the Test of Written Expression (writing samples elicited under standard conditions, scored under controlled conditions by at least two raters, and so on) in (more highly selected)

(a) samples tested in developmental research (Carlson, Bridgeman, Camp, and Waanders, 1985) involving the TOEFL and prototypical versions of the TWE, and

(b) samples taking both tests under fully operational conditions (see ETS, 1992b).[10]

Such findings attest to the validity and relevance for placement of both the SLEP and the local writing tests. The general issue of differences across colleges in "rating standards" is addressed in the following section.

---

[10] TOEFL/TWE correlations have been reported (e.g., ETS, 1992b: p. 13), by world region, for large, relatively highly selected ESL-samples taking both tests in operational administrations. Weighted averages of regional TOEFL/TWE correlations (corrected for unreliability of TOEFL scores), for almost 90,000 examinees tested in May 1991, are as follows: TOEFL Listening Comprehension (.59), TOEFL Structure and Written Expression (.61), TOEFL Reading Comprehension and Vocabulary (.59), and TOEFL Total (.63).

Table 3.2. Concurrent Correlations of Shortened SLEP Scores with Essay Rating for Two Language Groups, By College

Language /College	N	Correlation with Essay rating Shortened SLEP Test LC	RC	SLEP
Spanish				
City	1,416	.68	.70	.73
Mission	1,299	.51	.54	.57
Pierce	77	.59	.62	.66
West	146	.62	.61	.65
East	327	.49	.53	.55
Harbor	128	.44	.32	.41
Trade	322	.29	.27	.32
Valley	554	.46	.49	.53
Korean				
City	570	.63	.69	.74
Mission	21	.57	.65	.68
Pierce	39	.70	.69	.73
West	36	.56	.54	.62
East	15	.44	.78	.75
Harbor	15	.48	.36	.41
Trade	24	.54	.30	.43
Valley	98	.42	.55	.57

Note. Concurrent correlations tend to be higher in general samples at City, Mission, Pierce, and West, than at the other four colleges.

Differences in Rating Standards by College

Explication of the differences in levels and patterns of concurrent- and discriminant-validity by college is beyond the scope of this inquiry--such differences plausibly are due in part to effects associated with differences in dispersion by college inSLEP scores, differences in the reliability of ratings, and other factors. Data needed to pursue these and related issues directly are not available. However, the analyses reported below shed further light on the question of systematic differences by college in "rating standards," suggested by the lack of systematic covariation across colleges between mean writing score and mean score on the shortened SLEP.

A systematic assessment of differences in Essay performance relative to shortened SLEP performance was made, by analyzing average discrepancies between observed Essay rating and Essay rating estimated from SLEP scores, using a regression equation derived for the total sample (see Appendix B.5 for detailed results of the regression analysis). Mean residuals were analyzed by college, and for subgroups by gender, language, educational level, and age, respectively.

Salient results of these analyses are summarized in Figures 3.1 (college differences in mean residuals) and Figure 3.2 (other subgroup differences, plus college differences from Figure 3.1 for "same-scale" perspective). Positive discrepancies (residuals) indicate high Essay ratings relative to SLEP scores, and negative residuals indicate the opposite.

It is apparent that differences in mean residuals are much greater in the analysis by college, than in analyses involving the respective subgroups (for which data are aggregated across colleges). Of course, the observed mean residuals for the demographic subgroups may to some extent be artifactually inflated or deflated by college-level effects--for example, differential concentration of subgroups in higher- or lower-rating colleges. In the circumstances, interpretive inferences need to be drawn with caution.[11]

---

[11] In view of apparent differences across sites in "rating standards," pending resolution of questions regarding comparability of Essay scores across colleges, future investigations of subgroup differences in the relative development of essay-assessed "writing ability" and SLEP-assessed skills, need to be conducted at the college level. These and related issues are beyond the scope of the present study, but they constitute meaningful topics for further inquiry. More generally, the findings suggest the need for systematic study of the rating process, including the collection of data needed to evaluate degree of agreement among raters with respect to level as well as rank-order.

Figure 3.1. Mean residuals (Essay - Essay estimated from shortened SLEP scores) by college

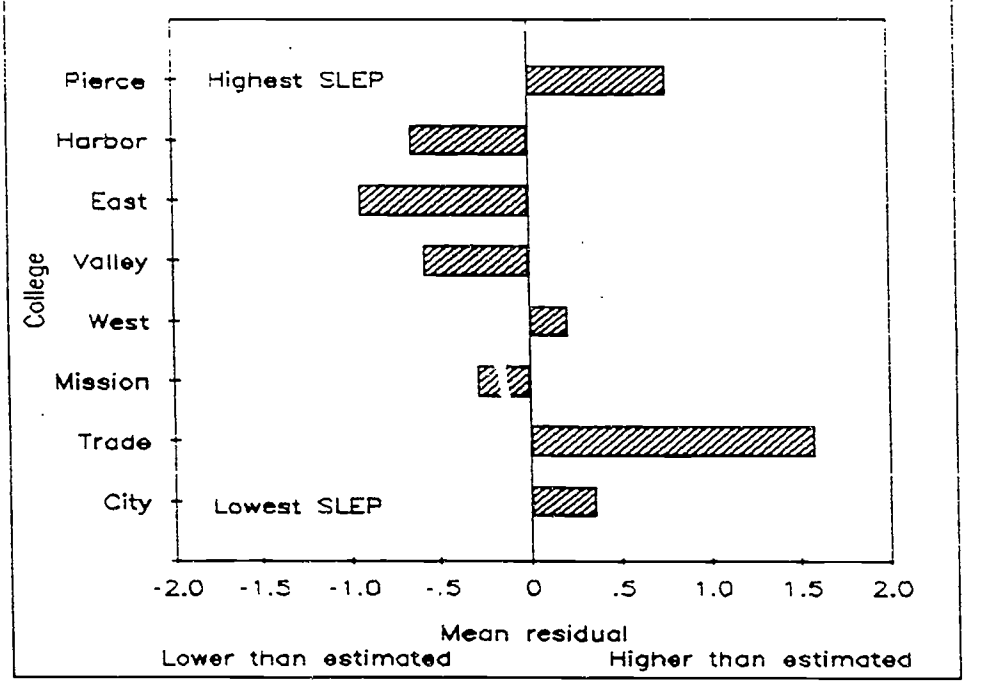
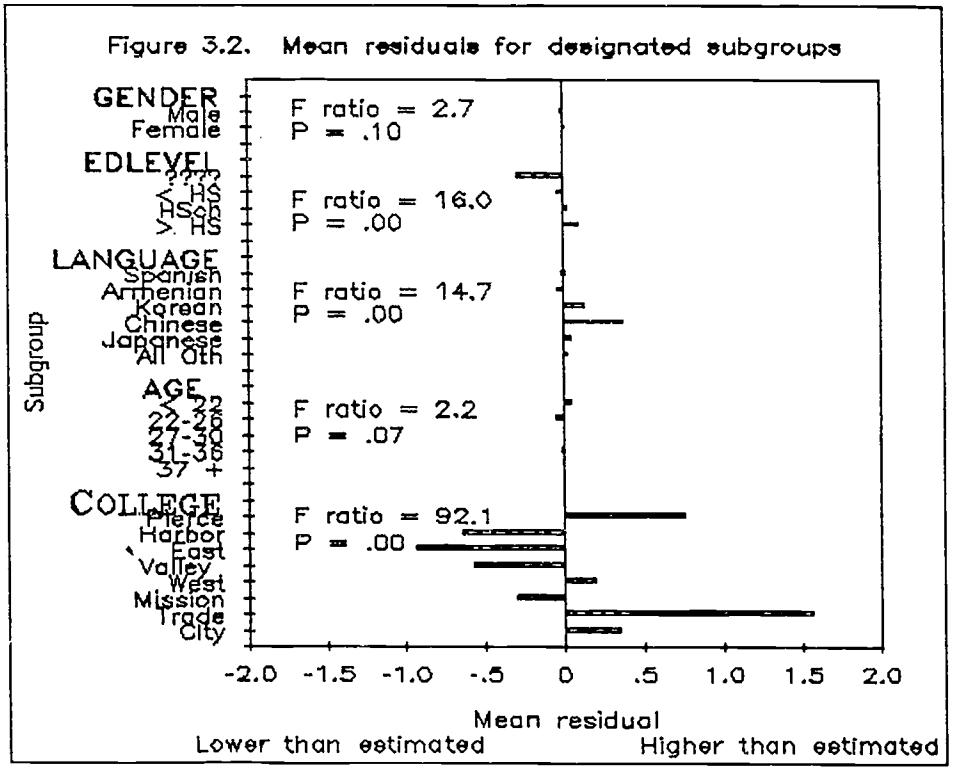


Figure 3.2. Mean residuals for designated subgroups



With these limitations in mind, it is of interest to note that mean residuals were larger in analyses by college than in the other analyses. Generally speaking,

- mean residuals by gender and by age group were extremely small;

- the pattern of residuals by educational level indicates simply that more highly educated students tend to earn somewhat higher Essay ratings, on the average, than expected from their shortened SLEP scores, and that the opposite was true for less highly educated students; and

- Chinese-speakers and Korean-speakers, albeit to a lesser extent, earned somewhat higher than expected Essay scores.

Again, these analyses do not take into account possible differences in concentration of subgroups by college.

Related Analyses: Weighting of Tests in the Placement Composite

In practice, shortened SLEP scores and Essay ratings are combined for placement purposes. As indicated earlier, raw scores on the shortened LC and RC sections of the SLEP test and the writing sample, respectively, are converted to "percent right" equivalents, and then combined using predetermined nominal weights, as indicated below:

College	Nominal weight		
	Essay	LC	RC
City, Harbor, Mission, Trade, Valley, West[12]	.34	.33	.33
East, Pierce	.60	.20	.20

Based on these nominal weights, the placement of students at the first group of colleges is expected to be influenced more by their SLEP scores than by the essay rating, while the opposite is true for the placement of students at the latter two colleges. However, the actual contribution of variables in a composite (weighted sum of scores on the variables) may differ from that indicated by the nominal weights involved (e.g., Guilford, 1965: pp. 424-425).

---

[12] The pattern indicated here is applicable only to about one-third of the sample from West. Only the Essay rating and the shortened SLEP test reading score are currently used, weighted 60/40. For present purposes, only the sample with the "34/33/33" composite is under consideration.

The actual composites computed on the respective campuses were included in the LACCD file. To assess the actual relative weighting of the three test scores in these composites, analyses were made of the "structure" of the composites used on the respective campuses--that is, the composite for a given campus was regressed on the three scores involved. Of interest are findings regarding the standard partial regression (beta) weights for the tests.

Salient findings are summarized in Table 3.3.

- Scores on the shortened SLEP test contributed roughly two-thirds of the variance when nominal wrights of ".34/.33/.33" were employed; Essay scores contributed more than did SLEP scores when nominal ".60/.20/.20" weights (or the variation employed st West) were used.

It is apparent that the actual weighting of the tests conforms relatively closely to the nominally specified weighting, with some variation in detail from college to college. Knowledge of the relative contribution of the variables in the placement composites used is important for evaluating observed correlations between course grade and scores on the placement variables.

More specifically, for example, restriction of range effects are likely to be more pronounced for SLEP scores than for Essay scores when placement is guided by the ".34/.33/.33" pattern of weights, and the opposite is likely to be true for situations in which the composite reflects relatively more Essay variance than SLEP variance. More direct evidence at the degree of curtailment due to use of tests in placement, and the attendant distorting effects on observed correlations is provided below.

Range Restriction in Placed Samples

One consequence of using test scores to place students is that the range of talent within the resulting samples is automatically restricted. The extent to which placement score distributions are restricted in placed samples is suggested in Figure 3.3, which shows distributions of standardized scores on a linear composite of percent-right scores on shortened SLEP LC and RC, and Essay, respectively; weighted to reflect their contribution to differences among the seven placement-level subgroups.[13]

---

[13] The method of multiple discriminant analysis (e.g., SPSS, 1990: p. 127-140) was used to assess differences among the seven groups with respect to the three test variables. One linear discriminant function accounted for 98 percent of observed differences among these groups; the percentage contribution of the respective tests was estimated as 35-35-32 for Essay, RC, and LC, respectively, consistent with the regression outcomes reported above.

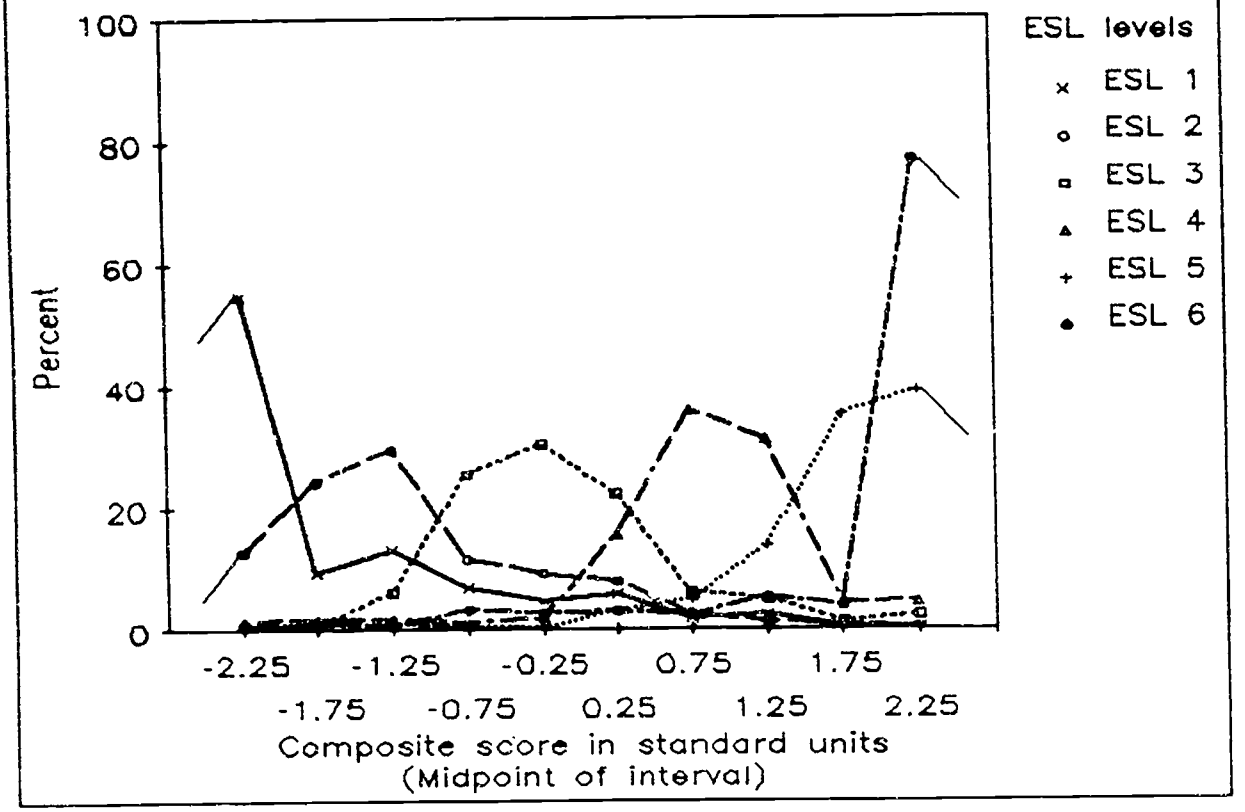
Table 3.3. Effective Contribution of Components in the LACCD Placement Battery, By Institution

College*	N	Beta weights			Percentage contribution				
		LC	RC	Essay	LC	RC	Essay	SLEP	Essay
City	3,833	.43	.35	.35	38	31	31	69	31
Harbor	431	.40	.43	.41	32	35	33	67	33
Mission	1,642	.46	.42	.27	40	37	23	77	37
Trade	474	.35	.41	.59	26	31	43	57	43
Valley	1,616	.38	.46	.41	30	37	33	67	33
West(a)	380	.40	.39	.37	34	34	32	68	32
East	941	.27	.30	.62	23	26	52	49	52
Pierce	490	.19	.25	.71	16	22	62	38	62
West(b)	858	--	.53	.61	--	46	54	46	54

Note: The composite reported by each school was regressed on LC, RC, and Essay. The "beta weights" are the resulting standard partial regression weights. Percentage contribution is the percent conversion of the ratios of the respective beta weights to the sum of all three weights.

* Colleges are grouped by known differences in formulae for weighting the placement variables. The first five colleges used the standard weighting (.33*LC + .33*RC + .34*Essay). East and Pierce used the 60/20/20 pattern. West(b) represents use of RC and Essay only, weighted .40(RC) and .60(Essay); West(a) reflects use of the full battery, with standard weighting (Spring 1991 only).

Figure 3.3. Relative frequency distributions for a composite of LCpct, RCpct, and ESSAYpct scores, weighted according to contribution to discrimination of subgroups classifed by ESL levels at City: Levels 1 thru 6 only



The foregoing analyses were based on data for City College only, and distributions are shown for students in the first six ESL placement levels only.[14] The distributions in Figure 3.3 overlap to some extent, indicating that these level classifications reflect factors in addition to test performance: initial consideration of background factors, shifts in course assignments, and so on. However, for present purposes, the important point is that use of the tests in placement generates course-level samples that are relatively homogeneous with respect to measured proficieny, **with attendant distorting effects on observed correlations among the variables involved.**

Such effects are clearly illustrated in Table 3.4, which shows intercorrelations of test scores in a restricted (English 85) at City College, and the corresponding intercorrelations in the total City sample. It is evident that the correlations observed within the restricted sample bear little resemblance to those observed in the unrestricted (preplaced) general sample.

Implications for assessing criterion-related validity. By extension, observed correlations between scores on the placement battery, on the one hand, and course grade, on the other, in restricted course-level samples (e.g., students taking English 85) are likely to be similarly limited as bases for inferences regarding the validity of the tests as measures of aspects of English proficiency in the ESL population under consideration.

Limitations of test/criterion correlations in highly selected samples were clearly illustrated by the findings of a classic empirical study reported by Thorndike (1949: p. 170-171), in which aspirants for pilot training were admitted without regard to their performance on the selection tests. After criterion data became available for all members of the unrestricted, experimental sample, validity coefficients were computed for (a) the unrestricted experimental group and (b) a restricted subgroup made up only of individuals who would have qualified based on their test performance. Validity coefficients for the various selection tests in the unrestricted and restricted samples, respectively, were as shown in Table 3.5.

- Coefficients for the "restricted" group clearly are lower than those for the "unrestricted" group.

---

[14] Students in Level 7 courses (largely Speech Communication) had lower average scores on the placement composite than did students at Level 6. This was true at City and at the other colleges as well (see Appendix B.2, for example). Reasons for this are not clear. In any event, data for the first six levels are sufficient to make the point at issue here.

Table 3.4. Intercorrelations of Placement Variables in Restricted and Unrestricted Samples: Illustrative Data, City College

Variable	ESSAY	LC	RC	SLEP	PlcComp
Restricted					
English 85 (N = 269)					
Mean	4.6	36.1	33.6	69.7	66.1
Sd	.8	4.0	5.3	6.4	4.8
ESSAY	1.0000	-.18	-.11	-.20	.53
LC	.62	1.00	-.08	.57	.44
RC	.66	.73	1.00	.78	.54
SLEP	.69	.93	.93	1.00	.72
PlcComposite	.85	.90	.89	.97	1.00
Mean	3.2	27.0	23.6	50.6	47.7
Sd	1.5	10.1	10.2	18.9	17.6
All students (N = 2,389)					
Unrestricted					

Note. Coefficients above the diagonal are for the English 85 sample, those below the diagonal are for the total sample. Underlining indicates spuriously high correlation (part-whole relationship).

Table 3.5. Illustrative Effects of Selection on Observed Validity Coefficients (from Thorndike, 1949: p. 171)

	Validity coefficients	
	Total group* (N = 1,036)	Qualified group** (N = 136)
Composite	.64	.18
Mechanical Principles Test	.44	.03
General Information Test	.46	.20
Complex Coordination Test	.40	-.03
Instrument Comprehension Test	.45	.27
Arithmetic Reasoning Test	.27	.18
Finger Dexterity Test	.18	.00

* Tested but admitted without regard to test score (unrestricted).

** Only those who would have been accepted if test scores had been considered (restricted).

- The relative size of the coefficients in the restricted group does not provide a basis for inferences regarding the relative validity of the tests when used with the entire applicant population.

As will be seen in the following section, both levels and patterns of test/grade correlations reflect range-restriction effects such as those considered illustratively above.

SECTION IV. OBSERVED CORRELATIONS OF PLACEMENT TESTS WITH COURSE GRADE

Findings reviewed in the preceding section indicate that within-college correlations between scores on the locally developed writing tests and scores on the shortened SLEP tend to center at about .60, varying between roughly .5 and .7 across colleges. These correlations were in general (pre-placement) samples from the LACCD population of ESL students.

> • The findings constitute novel empirical evidence of the validity of the locally developed writing tests; they attest directly to the validity of the shortened SLEP, and logically extend evidence of the validity of the full-length SLEP, as measures of relatively closely related but distinguishable aspects of proficiency in English as a second language.[15]

General evidence regarding the structure of human abilities indicates that a positive correlational manifold can be expected to obtain among measures of ability generally. This applies to measures of second-language macroskills (e.g., Hale, 1986; Oller, 1983; Carroll, 1967): measures of different aspects of English proficiency can be expected not only to be positively correlated, but also to exhibit "moderate" to "moderately high" levels of intercorrelation (e.g., .50 to .70) in samples from general ESL populations, such as that being served in the LACCD.

Grades earned by ESL students in courses of ESL instruction are widely accepted as validly indexing individual differences in "demonstrated competence" in the performance of varied tasks in proficiency-related domains that constitute foci of instruction--for example, reading and vocabulary, conversation, grammar, writing, or a combination of the foregoing. Accordingly, based on evidence and lines of reasoning developed above, scores on the shortened SLEP and the writing test can be expected to correlate positively with course grades.

However, when test/grade correlations are involved, only accumulated empirical evidence can resolve questions as to the levels at which specified tests (or types of tests) typically correlate--hence, can be expected to be correlated--with course

---

[15] By inference from the observed levels of correlation, writing samples are being rated on the respective campuses with a "workable" degree of rank-order reliability. Internal consistency estimates available for test-development samples for SLEP Form 1, reported in Section I, are reported below for current reference: for item-type sections included in shortened SLEP, estimated reliabilities were .81 for One Picture, .89 for Dictation, .89 for Cloze, and .71 for Four Picture items.

grades in various use contexts. Moreover, as indicated in the preceding section, observed correlations among test variables in restricted samples are limited as bases for inferences regarding the correlations among the variables in general samples.

- As shown in the preceding section, illustratively, observed correlations between scores on the writing samples and scores on the shortened SLEP in English 85, a placed sample, bear no resemblance to those observed in the general population involved. Similar distorting effects can be expected, within placed samples, for observed correlations of the placement tests with other measures, including course grades.

This section reports on analyses designed to obtain and evaluate empirical findings regarding

- levels and patterns of observed test/grade correlations in a representative array of ESL courses;

- the extent to which both levels and patterns of coefficients for the placement variables are affected by nonvalidity-related statistical artifacts (e.g., sample standard deviations for the placement tests, sample size, differences in relative weighting of SLEP and Essay in placement composites); and

- degree of consistency in observed levels of test/grade correlations across subgroups (e.g., gender, age, educational level, language);

The findings indicate that scores on the shortened SLEP and the writing test, respectively, and as combined for placement purposes in different colleges, tend to be positively correlated with course grades, and that this is true for data aggregated by college, gender, age, language, and so on; also that both level and pat-terning of correlations for placement tests with grade in ESL courses are affected in predictable ways by differential restriction of range on the placement tests.

Findings are presented and discussed following a brief overview of methodology and procedure.

General Methodology and Procedure

Analyses of test/grade correlations were conducted in a reduced sample made up students with "complete-data" records--that is, records with nonzero scores on the shortened SLEP, an essay rating, a score on the placement composite, and either a grade on the A-F scales or a pass/fail grade. Approximately 60 percent of the originally assessed population was included in the reduced sample.

Some 160 "unit" samples were represented in the complete-data sample--that is, samples representing courses in which students were actually graded (e.g., students earning a grade in English 85, City College, Spring 1991, constituted a unit sample, those earning grades in English 85 in Fall 1991 constituted a second unit sample, and so on.

In several of the unit samples--all were courses graded on the pass/fail system--there was no variability in the grade distribution; either all students passed or all students failed. These unit samples were excluded from the analysis, as were unit samples with fewer than 10 cases. The foregoing process resulted in the identification of 140 unit samples, including some 5,635 students. When students were classified within colleges by course, ignoring distinctions between terms of enrollment, the result was a total of 59 course-level samples--data for students in the respective courses, aggregated across terms of enrollment.

To permit assessment of test-criterion relationships in the the 59 course-level samples described generally above, while avoiding interpretive complications associated with possible differences in grading standards within and across colleges, all variables were "standardized" using unit-sample parameters (summarized in Appendix C, for 140 samples involved).

More specifically, distributions of scores on all variables (including grades on the five-point A-F scale [4,3,2,1,0] or the pass/fail [1,0] scale) were z-scaled, using unit-sample parameters--that is, scores on each variable were expressed as deviations from the unit-sample mean on the variable, and then divided by the corresponding standard deviation. Variables were thus converted to a common scale, with zero mean and unit standard deviation (0.0, 1.0).

From a methodological perspective, it is useful to note that when variables have been z-scaled as indicated above, coefficients computed using aggregated data for two or more intact unit samples are comparable to size-adjusted averages of corresponding unit-level coefficients.

Unless otherwise noted, the correlation coefficients reported below reflect relationships between the criterion variable (z-scaled grade), and z-scaled scores on the placement variables (shortened LC, shortened RC, shortened SLEP, Essay, and/or Comp [the composite of percent-right scores for LC, RC, and Essay, that was used to guide placement decisions]).

Analytical Procedures

Test/grade coefficients were first computed for each of the 59 course-level samples. Because scores on the shortened SLEP LC and RC sections were not used separately in placement, attention was focussed primarily on coefficients for the total score on the shortened SLEP (LC+RC).

1. To provide perspective regarding the level and range of test/grade correlations, and the extent to which level of association tended to be infuenced by range restriction and related factors, distributions of coefficients were developed for course-level samples classified according to corresponding test standard deviations.

2. To assess expected differential restriction of range effects, an analysis was made of distributions of coefficients for the shortened SLEP and the Essay, respectively, for samples from East and Pierce, and from City, respectively.

> Coefficients for SLEP were expected to be higher than those for Essay at East and Pierce (where SLEP was less heavily weighted than the Essay in placement), whereas the opposite was expected to be true for samples from City (and other colleges where Esssay was less heavily weighted than was SLEP).

3. A third line of inquiry bearing on the influence of statistical artifacts on observed levels of correlation involved regressing the observed coefficients on the corresponding sample standard deviations and sample size.

All of the foregoing analyses involved observed coefficients and corresponding sample statistics for the 59 course-level samples.

4. To assess general consistency of findings regarding test/grade correlations across diverse samples, correlational analyses involving z-scaled scores on all variables were conducted for subgroups classified by college, gender, age, educational level, language, and so on.

Finally, general estimates of levels of correlation for unrestricted samples are provided, based on the regression of course-level SLEP/grade, Essay/grade, and Composite/grade correlations on the corresponding standard deviations, sample size, and type of grading system (A/F vs. P/F: nominally coded "1" for AF system and "0" for P/F system). These estimates are intended primarily to give added emphasis to the fact that restriction-of-range effects distort observed correlations involving the test variables, and limit generalizations about the validity or relative validity of the tests involved as applied in general populations.

Findings

Results of the several lines of inquiry involving course-level coefficients are presented first; trends in test/criterion correlation in more general samples are then examined.

Analyses Involving Course-Level Coefficients

Table 4.1 reports observed course-level test/grade correlations for scores on the shortened SLEP (LC, RC, and SLEP [LC+RC], the writing sample (Essay), and composites (Comp) that were used in placement, respectively; also corresponding course-level standard deviations, and data on sample size for the 59 samples involved. Table 4.1 summarizes data only for the variables directly involved in the analyses reported below, namely, coefficients and sample statistics for SLEP (LC+RC), Essay, and Comp, respectively (see Appendix C for data on all variables).

Table 4.2 shows stem-leaf plots of test/grade coefficients for SLEP (LC+RC), Essay, and Composite (composites used in placement), for samples classified by relative degree of restriction on SLEP, as indicated by SLEP standard deviation (SD). Three SD-ranges were defined arbitrarily to include about one-third of the samples in each range. Note that SLEP standard deviation in the total sample is approximately 19.

> Level of correlation decreases as SLEP standard deviation decreases--true for the Essay and the composite, as well as for the SLEP itself.

In evaluating the median coefficients for SLEP and Essay, respectively, keep in mind that placement was dependent more heavily on SLEP performance than on Essay performance in the great majority of samples under consideration--Essay was more heavily weighted in placement than was SLEP in decisions affected placement only at East and Pierce (11 of the 59 samples involved).

Differential restriction of range effects are clearly discernible in the results of analyses summarized in Table 4.3. Stem-leaf plots are shown for 22 samples from City (where SLEP received greater weight in placement) and 11 samples from East and Pierce (where the opposite obtained).

> As expected, Essay/Grade coefficients tended to be higher at City than at East/Pierce, and SLEP/Grade coefficients tended to be higher at East/Pierce than at City.

Related analyses. The foregoing analyses involving simple cross-tabulations of the basic data under consideration indicate that

Table 4.1. Relationship of Designated Placement Variables to Course Grade in Course-Level Samples, and Standard Deviations of the Variables in the Corresponding Samples, By College and Subject

COLL	SUBJ	COUR	n	Short SLEP r	ESSAY r	COMP r	Short SLEP Sd	Essay Sd	Comp Sd
City	263	61	27	.41	.43	.48	11.4	1.1	10.8
1	263	63	14	.23	.32	.44	5.6	.7	3.8
1	263	71	230	.31	.31	.38	13.2	1.1	11.1
1	263	72	181	.12	.13	.17	11.3	1.1	9.2
1	263	73	105	.09	.32	.24	8.2	.9	7.3
1	263	74	182	.26	.33	.35	14.7	1.2	12.5
1	263	75	183	.13	.25	.20	14.6	1.1	12.3
1	263	76	117	.05	.33	.23	9.3	1.0	7.8
1	263	77	102	-.02	.39	.24	8.0	.9	5.7
1	263	78	83	.14	.17	.21	10.4	1.1	8.9
1	263	79	122	-.05	.34	.15	8.0	.9	6.7
1	399	80	15	.35	-.16	.25	6.3	.7	3.2
1	399	81	47	.13	.26	.24	9.0	.9	7.4
1	399	83	37	-.02	.30	.14	10.3	.9	8.9
1	399	84	212	.07	.10	.12	8.2	.8	5.9
1	399	85	269	.01	.12	.09	6.3	.8	4.7
1	399	86	118	.04	.27	.18	6.9	.8	5.6
1	399	87	16	.26	.00	.25	7.7	.6	4.3
1	901	72	78	.55	.35	.52	13.6	1.2	11.9
1	901	73	67	.32	.25	.38	12.9	1.1	11.5
1	901	113	46	.35	.37	.41	11.7	1.0	10.8
Harb	263	76	54	.34	.29	.39	14.8	1.2	12.3
3	399	84	59	.46	.32	.54	9.9	1.1	8.3
3	399	85	71	.06	-.02	-.02	7.7	.9	5.7
3	399	86	29	.20	.15	.18	9.3	1.3	9.6
Miss	400	1	116	.24	.01	.21	12.5	.3	8.8
4	400	2	92	.16	.07	.17	15.6	.6	10.9
4	400	3	139	.26	-.04	.23	12.2	.9	8.1
4	400	4	169	.03	-.08	.00	11.3	.7	7.8
4	400	5	139	.22	-.08	.21	10.4	1.0	6.9
4	400	6	79	.45	.17	.36	13.1	1.5	11.8

COLL	SUBJ	COUR	n	Short SLEP r	ESSAY r	COMP r	Short SLEP Sd	Essay Sd	Comp Sd
Trad	400	2	15	-.08	.54	.18	11.4	.8	7.7
6	400	3	72	.20	.26	.28	11.3	1.0	9.6
6	400	4	62	.30	.21	.29	11.2	1.1	9.2
6	400	5	21	.24	.17	.39	6.6	1.5	6.2
Vall	399	83	95	.00	.19	.11	9.2	.8	6.9
7	399	84	171	-.03	.20	.09	8.0	.8	5.7
7	399	85	165	.16	.12	.20	9.3	1.0	7.1
7	399	86	121	.13	.02	.11	9.8	1.1	8.1
7	399	87	26	.10	.16	.15	13.3	1.0	11.6
7	901	64	244	.09	.10	.12	11.2	.9	8.9
7	901	111	65	.33	.06	.26	12.2	1.2	11.1
7	901	113	85	.10	.16	.15	8.2	.9	5.5
West	400	3	109	.10	.20	.17	12.0	1.0	9.2
8	400	4	51	.12	-.06	.07	6.9	.7	5.1
8	400	5	71	-.13	-.08	-.15	8.3	1.0	7.0
8	400	6	31	-.06	.30	.19	6.4	1.0	5.5
East	263	76	52	.23	.33	.35	11.6	.7	7.7
2	399	83	54	.20	.40	.39	13.4	.7	8.0
2	399	84	202	.26	.21	.33	11.4	.7	7.0
2	399	85	104	.30	-.19	-.02	7.8	.7	5.7
2	399	86	74	.01	-.22	-.20	9.5	.9	8.4
2	901	113	133	.39	.34	.41	15.6	.9	11.4
Pier	399	84	95	.35	.17	.40	10.7	.4	4.8
5	399	85	93	.34	.25	.42	9.6	.6	5.8
5	399	86	31	.11	.13	.06	6.6	.7	5.9
5	399	87	34	.42	.16	.25	9.8	.9	8.7
5	901	113	78	.27	.14	.19	14.8	1.6	17.1

Table 4.2. Distributions of Course-Level Correlation Coefficients for Samples Classified According to Degree of Restriction on Shortened SLEP

	SLEP standard deviation 11.5 +		9.3 - 11.4		Less than 9.3	
SLEP/Grade coefficients						
Stem	Leaf		Stem	Leaf	Stem	Leaf
.5	5		.5	--	.5	--
.4	15		.4	26	.4	--
.3	123459		.3	045	.3	05
.2	034667		.2	00268	.2	346
.1	0036		.1	2346	.1	0123
.0	--		.0	1359	.0	014679
-.0	--		-.0	28	-.0	2356
-.1	--		-.1	3	-.1	3
-.2	--		-.2	--	-.2	--
Mdn	.27			.15		.08
Essay/Grade coefficients						
.5	2		.5	--	.5	--
.4	118		.4	--	.4	--
.3	5568899		.3	023	.3	0229
.2	0136		.2	1156	.2	0467
.1	5779		.1	0235677	.1	023679
.0	--		.0	2	.0	0
-.0	--		-.0	88	-.0	268
-.1	--		-.1	--	-.1	69
-.2	--		-.2	2	-.2	--
Mdn	.35			.17		.165
Composite/Grade coefficients						
.5	2		.5	2	.5	--
.4	118		.4	002	.4	4
.3	5568899		.3	3	.3	9
.2	0136		.2	0113589	.2	4455
.1	5579		.1	124788	.1	12589
.0	--		.0	0	.0	799
-.0	--		-.0	--	-.0	226
-.1	--		-.1	--	-.1	5
-.2	--		-.2	0	-.2	--
Mdn	.35			.21		.125
N	19			20		20

Note. To read coefficients, combine "stem" and "leaf" values. For example, under Composite/Grade, .52, .41, .41, .48 and so on.

Table 4.3. Plots of Course-Level Test/Grade Correlation Coefficients for Samples from City, and East and Pierce, Respectively

CITY	SLEP		Essay		Composite
Stem	Leaf	Stem	Leaf	Stem	Leaf
.5	5	.5	--	.5	2
.4	1	.4	34	.4	0148
.3	1255	.3	0122334579	.3	588
.2	3668	.2	5567	.2	01344455
.1	2334	.1	0237	.1	24578
.0	14579	.0	0	.0	9
-.0	225	-.0	--	-.0	--
-.1	--	-.1	6	-.1	--
-.2	--	-.2	--	-.2	--
No. samples = 22					
Mdn r	.135		.305		.240

EAST, PIERCE					
Stem	Leaf	Stem	Leaf	Stem	Leaf
.5	--	.5	--	.5	--
.4	2	.4	0	.4	012
.3	0459	.3	34	.3	359
.2	0367	.2	15	.2	5
.1	1	.1	3467	.1	9
.0	1	.0	--	.0	6
-.0	--	-.0	--	-.0	2
-.1	--	-.1	9	-.1	--
-.2	--	-.2	2	-.2	0
No. samples = 11					
Mdn r	.270		.191		.290

Note. To read coefficients, combine entry in "stem" column with successive entries in "leaf" columns. For example, SLEP (.55, .41, .31, .32 and so on).

(a) level of test/grade correlation varies markedly across samples,

(b) level of correlation tends to decrease with degree of restriction on placement variables (restriction on SLEP used illustratively), and

(c) differential restriction of range effects are present, thus complicating validity-related interpretation of the observed coefficients for the variables involved (Essay and SLEP data used illustratively).

Additional evidence regarding the extent to which level of test/grade correlation across the 59 samples covaried with non-validity-related sample statistics (standard deviations and size) is provided by results of analyses involving the simple correlation of sample-level coefficients for SLEP, Essay, and the placement composite (Comp), respectively, with the corresponding sample standard deviations and sample size.

Table 4.4 shows correlations reflecting the extent to which differences in level of test/grade relationship (sample test/grade coefficients) were associated with corresponding differences in sample dispersion on the test involved. Data were analyzed separately for ESSAY-restricted samples (East/Pierce) SLEP-restricted samples (all other colleges), the classification based on relative weighting of the two measures in placement composites, used in the distributional analyses reported above.

- In the overall analysis, regardless of the particular validity coefficient involved, SLEP standard deviation was the best single predictor of level of relationship.

- However, for SLEP-restricted samples, the best predictor of level of SLEP coefficient was Essay standard deviation, and for ESSAY-restricted samples, SLEP standard deviation was the best predictor of level of Essay coefficient.

- As expected from results of the tabular analyses, above, Essay correlated more closely with grades when SLEP was emphasized in placement, and the opposite was true when Essay was emphasized.

Test/Grade Coefficients in Diverse Subgroups

The foregoing analyses focus attention on distributions of coefficients, calling attention to variability in observed test/grade correlations as well as central tendencies. The findings also point up nonvalidity-related variables that influence observed outcomes, complicating validity-related interpretive inferences.

Table 4.4. Correlation of Observed Coefficients for Designated Placement Variables with Corresponding Sample Standard Deviations and Sample Size: Analyses Across 59 Samples, Weighted by Sample Size

Test/grade coefficient in sample	Sample statistic SLEP Sd r	Essay Sd r	Sd r	Comp Size N r	Mean coeff. (r)
(All Samples)					
SLEP coeff.*	.53	.18	.45	-.22	(.17)
ESSAY coeff.	.19	.14	.26	-.03	(.18)
COMPOSITE coeff.	.46	.24	.34	-.16	(.22)
Mean (statistic)	10.6	0.92	8.3	135	(N=5635)
(SLEP-restricted)**					
SLEP	.31	.46	.62	-.20	(.14)
ESSAY	.22	.37	.29	-.09	(.18)
COMPOSITE	.36	.47	.55	-.27	(.20)
Mean (statistic)	10.5	0.95	8.3	140	(N=4680)
(ESSAY-restricted)**					
SLEP	.13	-.21	-.11	.03	(.28)
ESSAY	.63	-.07	.17	.24	(.17)
COMPOSITE	.50	-.31	-.06	.31	(.26)
Mean (statistic)	11.4	0.76	8.2	112	(N= 950)

* Correlation of variable with grades in the respective samples--the criterion variable.

** SLEP-restricted samples are from City, Harbor, Mission, Trade, Valley and West; ESSAY-restricted samples are from East and Pierce. Data for samples were weighted by sample size.

Given these circumstances, emphasis in further analysis of the data was primarily descriptive, concerned primarily with providing evidence bearing on the general consistency of correlational outcomes across colleges and subgroups defined by gender, language, educational level, and age, respectively. For general perspective, findings are presented separately by grading system (A-F versus Pass/Fail).

Table 4.5 shows college-level findings: average correlations with course grade for all the placement variables--that is, shortened SLEP listening (LC), reading (RC), and combined (SLEP = LC+RC) scores, Essay score, and score on the placement composite (Comp). Colleges are listed in descending order with respect to average course-level standard deviation on the shortened SLEP.

- Average coefficients for all placement variables tend to be noticeably lower at Valley and West, with greatest within-course restriction on SLEP (and incidentally on other variables), than at the other colleges.

- The overall Essay/Grade coefficient was lower when P/F grades were involved ($r = .07$), than when A/F grades were employed ($r = .20$). This outcome was influenced strongly by the very low Essay/Grade correlation (.01) in the large sample from Mission.

- No interpretable "relative validity" pattern is discernible: in three analyses the highest coefficient was that for a shortened-SLEP section (RC), in two instances this was true for the combined SLEP score, in three instances for Essay, and in the remaining four instances, for the Composite.

Findings for the designated subgroups, presented in Table 4.6, indicate that coefficients for the respective subgroups are generally positive; also that differences in level are roughly comparable to those observed across colleges; and so on.

An Exercise in Estimation

Findings reviewed thus far indicate that differences across courses in observed test/grade correlations are associated with differences in nonvalidity-related characteristics of the samples in which the coefficients are computed, including sample standard deviations on test variables and sample size. The extent to which each of these factors, considered separately, is associated with particular test/grade coefficients was indicated in Table 4.4, above. To assess the extent to which these variables, and type of grading system, tend to account for differences in observed course-

Table 4.5. Pooled Within-Course Correlations of Tests with Grades, by College and Type of Grading System (A-F vs. Pass/Fail)

College		N	LC	RC	SLEP	ESSAY	COMP
Mission	(PF)	732	.21	.19	.23	.01	.21
East	(AF)	619	.15	.27	.26	.15	.23
Pierce	(AF)	331	.22	.32	.31	.18	.31
Trade	(AF)	83	.21	.23	.28	.20	.32
	(PF)	87	.24	.25	.31	.23	.34
Harbor	(AF)	213	.12	.28	.26	.17	.26
City	(AF)	2,011	.13	.14	.25	.27	.25
	(PF)	326	.01	.11	.09	.13	.15
Valley	(AF)	877	.11	.07	.10	.11	.14
	(PF)	95	-.05	.13	.06	.11	.08
West*	(AF)	260	.06	-.02	.02	.08	.06
Total	(AF)	4,394	.13	.15	.17	.20	.22
	(PF)	1,240	.13	.16	.18	.07	.19

Note. Colleges are listed in descending order with respect to average course-level SLEP standard deviation.

* Includes only those students with both LC and RC scores on the shortened SLEP, primarily from the Spring 1991 term.

Table 4.6. Correlations of Tests with Grades in Designated Samples: By Grading System (AF=A to F vs. PF = Pass/Fail)

Group	N	LC	RC	SLEP	Essay	Comp
Gender						
Males(AF)	1,712	.08	.16	.14	.16	.18
Females	2,682	.18	.17	.20	.21	.26
Males(PF)	555	.20	.30	.31	.04	.31
Females	685	.08	.06	.08	.08	.10
Language						
Spanish(AF)	1,471	.16	.17	.19	.17	.22
Armenian	782	.14	.15	.17	.28	.26
Korean	379	.05	.16	.13	.23	.21
Chinese	359	.05	.19	.17	.16	.18
Japanese	264	.12	.29	.26	.03	.19
All Other	1,139	.12	.17	.18	.21	.24
Spanish(PF)	555	.20	.30	.31	.04	.31
Armenian	71	-.08	.24	.13	.12	.15
Korean	86	-.03	.11	.08	.24	.19
Chinese	13	-.14	.12	.01	.24	.11
Japanese	24	.41	.39	.48	-.11	.36
All Other	273	.21	.21	.25	.10	.27
Educational Level						
NonHSgrad(AF)	870	.14	.13	.15	.15	.17
HS grad	3,147	.12	.15	.16	.20	.22
BA or higher	377	.23	.20	.25	.17	.28
NonHSgrad(PF)	517	.12	.12	.14	.10	.17
HS grad	609	.13	.17	.19	.04	.19
BA or higher	114	.16	.21	.23	.05	.23
Age						
< 22.5 yrs(AF)	1,101	.09	.19	.17	.16	.20
22.5-30.5 yrs	1,560	.10	.12	.13	.18	.18
30.5 yrs plus	1,733	.19	.16	.21	.23	.27
< 22.5 yrs(PF)	232	.12	.20	.19	.11	.22
22.5-30.5 yrs	494	.15	.14	.18	.02	.16
30.5 yrs plus	514	.12	.17	.18	.10	.21

level correlations, three stepwise multiple regression analyses were conducted, each involving the regression of a particular test/grade coefficient (treated as the dependent variable) on the following nonvalidity-related variables: standard deviations for SLEP, Essay and the Composite, respectively, sample size, and type of grading system (nominally coded 1 = AF system and 0 = PF system).

Salient findings of these stepwise analyses are shown in Table 4.7. In evaluating the findings generally, it is useful to note that negative coefficients for the composite reflect suppression effects: the composite standard deviation was positively related to course-level coefficients (see Table 4.4, above). The negative weighting shown, reflects the close relationship between restriction on the SLEP and restriction on the Composite, and the fact that SLEP standard deviation was more closely related to the criterion (course-level correlation) than was the composite. However, the negative weights for sample size are directly interpretable as indicating higher coefficients tend to be found in smaller samples.

With the foregoing in mind, it can be seen in the table that:

standard deviation on the shortened SLEP was the most highly weighted variable not only in analyses involving course-level SLEP/grade coefficients but also in analyses involving Composite/grade coefficients;

the most important predictor of level of Essay/grade correlation was type of grading system; SLEP standard deviation provided the next largest amount of information; and

the Composite standard deviation did not contribute to level of Essay/grade correlation in the stepwise regressions, after other variables were entered.

The corresponding multiple correlation coefficients are also shown in the table. Results of the regression analyses provide a direct empirical basis for estimating levels of correlation in otherwise comparable, but unrestricted samples--that is, regression equations for estimating course-level coefficients for the respective tests from corresponding sample statistics.

Table 4.7. Results of a Stepwise Regression of Course-Level Coefficients on Designated Non-Validity-Related Sample Characteristics: Data Weighted by Sample Size

Dependent coefficient		Standard partial regression weights* Standard deviation				
	(R)	SLEP	Essay	Comp	Grading A-F	Size N
SLEP(r)	(.63)	1.00	.17	-.48	.14	-.27
Essay(r)	(.47)	.57	.17	...	1.00	-.25
Composite(r)	(.59)	1.00	.08	-.76	.20	-.18

*These are conventionalized weights--that is, coefficients in each row are shown as proportions of the largest coefficient.

... Indicates variable did not enter the stepwise regression.

The raw score regression weights and contant terms for the three equations were as follows:

Estimated coefficient	Raw-score regression weights for variable: SLEP sd	Essay sd	Comp sd	AFsys	Size	Constant
SLEP(r.est)	.051	.097	-.025	.043	-.0005	-.2187
ESSAY(r.est)	.013	.064	...	.136	-.0002	-.1031
COMP(r.est)	.063	.214	-.048	.079	-.0004	-.2582

We know that standard deviations for SLEP, Essay, and Composite are, respectively, 19.3, 1.5, and 16.8, in the general population. To estimate coefficients for the shortened SLEP, the Essay, and the Composite, respectively, for course-level samples with the foregoing standard deviations, the three standard deviations were inserted in the respective raw-score regression equations, along with the average N (135), and the nominal code of "1" for course grading under that system. P/F-course estimates were similarly generated.

Results are shown below:

Grading system	Estimated coefficient: Short SLEP	Essay	Composite
A/F system	.47	.35	.50
P/F system	.43	.21	.42

These estimates reflect the observed empirical relationships between course-level correlation coefficients and sample standard deviations for the respective placement variables.

Generally speaking, these findings point up the extent to which the observed average course-level coefficients that have been reviewed in this section, tend to reflect effects associated with restriction of range on the tests, sample size, and type of grading system. They give added emphasis to the axiomatic proposition that observed levels of test/grade correlation need to be evaluated in light of the possibility that

> " . . . the magnitude of the correlation (between tests and grades) may vary as a function of the degree to which a test was used to place students in the course under investigation and/or the variation in grading standards across classrooms" (California Community Colleges, 1992: p. 20).

SECTION V. RECAPITULATION AND EVALUATION OF FINDINGS

As indicated in Section I, the Secondary Level English Proficiency (SLEP) test was initially selected for use in placement testing in the LACCD by ESL professionals (the LACCD ESL Committee) through a process that included

(a) a comprehensive, comparative assessment of the relative appropriateness (for LACCD purposes) of nine commercially available tests of ESL proficiency, and

(b) consideration of the findings of a specially commissioned study (Butler, 1989) of the performance of samples of LACCD students on all SLEP items--findings indicating that the level of difficulty of the test was generally appropriate for the students involved, and that average scores tended to covary positively with ESL placement level.

To meet purely pragmatic interest in reducing the amount of testing time needed to administer the SLEP (approximately 85 minutes) and a locally developed writing test (30 minutes), a decision was made to administer only a portion of the full-length SLEP, namely, two of four item-type susections involving listening comprehension items, and two of four item-type subsections involving reading comprehension items.

The present study was undertaken with the encouragement and support of the SLEP Testing Program at Educational Testing Service, and the collaboration of the LACCD District Office, represented by Ms. Rebecca Tillberg,

(a) to examine patterns of performance on the shortened SLEP and the writing tests in the general ESL population and in selected demographic subpopulations;

(b) to investigate levels and patterns of concurrent relationships among scores on the components of the LACCD placement battery;

(c) to obtain base-line empirical evidence regarding observed levels of correlation between scores on the shortened SLEP test, the writing test, and the placement composite, on the one hand, and student performance in ESL courses, as indexed by grade earned (a grade on the "A-F" scale, or a Pass/Fail grade), on the other, by course and by college, and in various subgroups (e.g., gender, educational level, age, language), and

(d) to provide an analytic assessment of the extent to which observed relationships in placed samples are influenced by non-validity-related factors, especially restriction of range on the tests that were used to place students.

Findings pertaining to these objectives--and the more general objective of extending evidence of the validity of the full-length SLEP test in the LACCD context--are summarized briefly below.

Summary

The ESL population being assessed with the shortened SLEP is extremely heterogeneous with respect to age, educational level, ethnicity, and language background.

- The average student is about 29 years of age; based on data of record on date of birth, some students were less than 13 years old and others were over 65 years of age when assessed in the LACCD.

- The majority of students hold secondary-school diplomas earned in schools located outside the United States, but the sample includes individuals who have not completed secondary school--some enrolled in adult education courses, some as special students--as well as individuals with bachelor's or higher degrees.

- The largest linguistic subpopulation is made up of native-speakers of Spanish (about 44 percent of the total); none of the other directly identified language groups accounts for much more than 10 percent (11 percent reported Armenian, and less than 1 percent reported Filipino).

- The shortened SLEP sections are at psychometrically efficient levels of difficulty for the general LACCD ESL population as well as for subgroups that differ markedly with respect to age, educational level, language background, and self-reported ethnic group membership.

- Results of an analysis of trends in mean scores on the shortened SLEP and the Essay, respectively, across subgroups classified by educational level indicate that the two means covaried directly. Similar patterns of covariation are discernible for the two test means across other demographic subgroups. Generally speaking, demographic subgroups with higher (lower) means on the SLEP tend to have higher (lower) means on the writing test.

- Analyses of concurrent relationships between shortened SLEP scores and scores on the writing tests were moderate to strong in each college setting: within-college correlations between scores on the locally developed writing tests and scores on the shortened SLEP centered at about .60, varied between roughly .5 and .7

across colleges. These correlations were in general (pre-placement) samples.

The relationship between the shortened SLEP reading comprehension score and Essay rating typically was somewhat stronger than that observed for the shortened SLEP listening comprehension score with Essay, consistent with expectation--that is, it is plausible that reading and writing skills should tend to be somewhat more closely related than are listening and writing skills.

At the same time, coefficients for total score on the shortened SLEP test (LC + RC) were somewhat larger than those for RC alone, indicating that each of the two sections is providing some unique information regarding aspects of the ability being measured by the writing test--a face valid measure of ability to write comprehensibly in English.

Levels of conconcurrent relationship observed for college-level samples were also observed in the results of analyses conducted by college, illustratively, for Spanish-speaking and Korean-speaking subgroups, respectively.

The foregoing findings provide strong empirical evidence of concurrent- and discriminant-validity properties for the shortened LEP and the locally developed writing tests.

In this connection, it is considered particularly noteworthy that the levels of concurrent relationship observed in the LACCD context, between scores on the shortened version of the SLEP and scores on local writing tests (with attendant differences in topic, rating procedures, and so on), are comparable to levels that have been found to obtain between scores on the Test of English as a Foreign Language and the Test of Written Expression (writing samples elicited under standard conditions, scored under controlled conditions by at least two raters, and so on)--true for,

(a) samples tested in developmental research (Carlson, Bridgeman, Camp, and Waanders, 1985) involving the TOEFL and prototypical versions of the TWE, and

(b) samples taking both the TOEFL and the TWE tests under fully operational conditions (see ETS, 1992b).

The TOEFL/TWE samples, of course, are more highly selected than is the educationally heterogeneous LACCD sample. However, the LACCD findings attest to the power of direct observation of samples of pertinent behavior in ESL proficiency assessment.

Analyses of levels and patterns of correlation of the placement tests with course grades, involving 59 course-level samples, were complicated due to the need to assess the influence of nonvalid-ity-related factors on both levels and patterns of observed test/ grade coefficients. The generalizations outlined below reflect findings reviewed in detail and evaluated in Section 4.

Scores on the shortened SLEP and the writing test, respectively, and as combined for placement purposes in different colleges, tend to be positively correlated with course grade, but there is substantial variability across courses in level of relationship. Distributions of observed course-level coefficients for the respective tests tend to center at about the .2 level, with substantial variability (ranging upward from negative values to values above .5).

Essay scores received relatively more weight than did SLEP scores in placement at East and Pierce, and in samples from those colleges, average coefficients were higher for the SLEP than for the Essay. SLEP was weighted relatively more heavily than Essay in other college samples, and in these samples, Essay coefficients typically were higher than SLEP coefficients. Such findings are interpretable as reflecting differential restriction of range effects.

Differences across 59 course-level samples, with respect to level of test/grade coefficients for the placement tests (that is, SLEP/grade coefficients, Essay/grade coefficients, and Composite/grade coefficients, respectively), were found be associated relatively strongly with nonvalidity-related sample characteristics: differences in the corresponding sample standard deviations, type of grading system employed (A/F versus Pass/Fail), and sample size.

The respective test/grade coefficients for the 59 course-level samples, treated as dependent variables, were regressed on the five nonvalidity-related variables. Results indicated that

sample standard deviation for the shortened SLEP was the most highly weighted variable not only in analyses involving course-level SLEP/grade coefficients but also in analyses involving Composite/grade coefficients;

the most important predictor of level of Essay/grade correlation was type of grading system; SLEP standard deviation provided the next largest amount of information; and

- the Composite standard deviation did not contribute to level of Essay/grade correlation in the stepwise regressions, after other variables were entered.

Regression equations developed in the foregoing analysis provided a basis for estimating test/grade coefficients (SLEP/grade, Essay/Grade, and Composite/grade coefficients, respectively), assuming no selection on the variables involved. Values for general sample standard deviations were substituted for the restricted values in the respective equations. Estimated values and observed values are shown below.

	Test/grade correlations		
	Short SLEP	Essay	Composite
A/F sys (estimated)	.47	.35	.50
A/F sys (observed)	.17	.20	.22
P/F sys (estimated)	.43	.21	.42
P/F sys (observed)	.18	.07	.19

Generally speaking, the findings involving test/grade relationships that have been reviewed should be thought of primarily as a basis for empirical assessment of the extent to which observed course-level test/grade coefficients generally, are influenced by nonvalidity-related factors, and are thereby limited as bases for inferences regarding "levels or patterns of validity" for the tests involved when applied in the general population under consideration.

In more general analyses, positive test/grade correlations varying around the observed values noted above were found to obtain in samples aggregated by gender, age, educational level, and language, respectively.

Regarding the Writing Tests

- Pooled within-school SLEP/Essay coefficients typically were larger than the "total sample" coefficients, the latter being attenuated by inconsistent ordering of school means on the respective tests--an anomalous finding, in view of the systematic, positive within-school relationships observed.

- College differences in "rating standards" are suggested by the results of analyses of differences across colleges with respect to mean discrepancies between observed Essay rating, and Essay rating estimated from SLEP performance, using a total-sample

regression equation. Mean discrepancies (residuals) were substantially more prononced in analyses by college, than in analyses for subgroups by gender, language, educational level, and age, respectively, aggregated across colleges.

- In view of apparent differences across sites in "rating standards," pending resolution of questions regarding comparability of Essay scores across colleges, future investigations of subgroup differences in the relative development of Essay-assessed "writing ability" and SLEP-assessed skills need to be conducted at the college level.

These and related issues constitute meaningful topics for further inquiry. The findings suggest the need for systematic study of the rating process on each campus, including the collection of data needed to evaluate degree of agreement among raters with respect to level as well as rank-order.

Concluding Observations

The findings of this collaborative undertaking provide direct empirical evidence that the shortened SLEP and locally developed writing tests are providing valid information regarding related aspects of ESL proficiency in the demographically diverse ESL student population being served by the LACCD; and the findings logically extend available evidence supportive of the validity of the Secondary Level English Proficiency Test for ESL assessment purposes in the LACCD and elsewhere.

Considered in light of the rigorous screening and evaluation of available psychometric options that resulted in SLEP's initial selection by the ESL Committee, and other evidence of SLEP's validity cited herein including the generally positive findings reported by Butler (1989), regarding SLEP's appropriateness for ESL assessment purposes in the LACCD, the findings that have been reviewed herein support and extend the working proposition advanced at the outset, namely, that

- the SLEP is a valid test of psychometrically distinct, albeit closely related, aspects of acquired proficiency in English as a second language: the ability to comprehend utterances in English, and the ability to read and comprehend the substance of material written in English, and

- the SLEP can be expected to provide reliable and valid information regarding these abilities, when used for ESL assessment purposes by secondary schools, colleges, and other institutions, and

- given the demographic diversity of the LACCD samples in-volved in this study, the findings extend evidence that the SLEP can be used validly with local ESL student populations that are quite heterogeneous with respect to age, educational level, language background, and national origin.

The local writing tests and the shortened SLEP appear to be providing an effective basis for placing students, within time constraints that appear to be considered necessary, from an administrative perspective.

As noted by Butler (1989), some cost in terms of diminished reliability and validity undoubtedly is entailed by the use of only a selected portion of the test. Available evidence bearing on this important question is limited, but consistent with the validity-cost assumption.

In a cooperative study currently in progress (Wilson, 1993), involving large samples of Japanese-speaking students being assessed for ESL placement at Temple University-Japan, LACCD-parallel subscores were computed post hoc from data available for the full-length SLEP. Correlations between these scores and ratings of speaking ability and writing ability, respectively, were approximately .05 correlation points lower than correlations for corresponding full-length scores, which centered around .6.

- Rudmann (1991), at Irvine Valley College, reports observed correlations centering around .40 between grades in ESL courses and scores on the full-length SLEP.

Analyses involving the full-length SLEP test would be needed in order to address questions of comparative validity: comparative validity of the current shortened, LACCD version of the SLEP, and the full-length test, and/or comparative validity of the particular SLEP item types selected versus those not selected, and so on.[16]

---

[16] In connection with the latter issue, it is noteworthy that one LACCD college (West) opted to use only the shortened SLEP reading section, with the writing test, for placement. More generally, in reporting on the basic SLEP validation study, Stansfield (1984) observed that users might elect to test either listening comprehension or reading comprehension in some circumstances (e.g., listening comprehension for level-placement decisions involving courses emphasizing the development of conversational skills). An evaluation of experience at West, where only a score on items from the SLEP "reading comprehension" ability domain is used, would be useful.

- One empirical approach to such an assessment that offers a variety of pertinent data-generating possibilities would involve administering the full-length SLEP (Form 1 or one of the other two equated forms) as an exit test, to students who were tested for placement with the shortened version. The resulting item-level data for both tests would provide a basis for assessing average (net) change in performance on the particular item-types now being used, and make it possible to evaluate the relative validity of all eight SLEP item types. Moreover, the end-of-course distributions of scores are needed to establish measured levels of proficiency associated with particular ESL instructional sequences.

Generally speaking, the use of more extensive and/or comprehensive testing procedures in ESL placement can reasonably be expected to be accompanied by benefits attendant upon reduced incidence of perceived misplacement: for example, reduction in the educational and administrative costs associated with course changes, improved satisfaction with results of the placement process on the part of both teachers and students, and improved retention of students in the ESL program.

A formal evaluation of the ESL placement process would provide empirical evidence that is pertinent to cost/benefit analysis, as well as to an overall evaluation of the ESL placement program in the LACCD.

APPENDICES

Appendix A

A.1. Holistic Scale for Rating Writing Samples in the LACCD (from Butler, 1989)

A.2. Data on Standardized English Proficiency Test Evaluated by the ESL Committee (from Butler, 1989, Appendix)

A.3. Illustrative SLEP Items

Appendix B

B.1a. Distributions of Scores on Variables Used in Placement, for All LACCD Students Assessed: Spring 1991 through Spring 1992

B.1b. Descriptive Statistics for Study Variables, by College

B.2. Descriptive Statistics for Placement Variables, by College and ESL Level (Last Course Enrolled): Classification of Courses (Other than ESL 400) According to City College Flow Chart (see text, Exhibit A)

B.3. Mean Scores on Placement Variables, by Language Group and College: All Students Assessed

B.4. Relationship Between Placement Variables and Selected Demographic Variables, By College

B.5. Results of Regression of Essay Rating on SLEP LC and RC in the All-LACCD Sample: All Students Assessed (Basis for Residual Analyses, Section 3)

Appendix C

C.1. Descriptive Statistics for Placement Variables in Unit Samples Samples (Students Classified by Course and Term {"SAMP"]: Parameters Used for Z-Scaled Transformation of Test Variables)

C.2. Descriptive Statistics for Placement Variables, in Samples Aggregated by Course, Across Terms of Enrollment: By College and Subject: 59 Course-Level Samples

C.3. Correlations of Placement Variables with Course Grade, and Standard Deviations of the Variables in Corresponding Samples, Classified by College and Subject

Appendix A.1. Holistic Scale for Rating Writing Samples in the LACCD

LOS ANGELES COMMUNITY COLLEGE DISTRICT

ESL RATING SCALE

0 Off the topic

1 **Organization & Coherence**--one or more sentences or fragments unable to supply details, examples, or elaboration in response to the prompt.
Mechanics--obvious lack of control of basic syntax, morphology, spelling.

2 **Organization & Coherence**--a loose collection of sentences or fragments (short and elliptic or long and unclear) vaguely related to the prompt. May also appear as a collection or list of "notes" on the topic.
Mechanics--may lack control over basic syntax, morphology, spelling. May exhibit an "oral" style— use of "gonna" (for going to), etc.

3 **Organization & Coherence**--generally stays on topic, but sentences show little development. Essay may ramble or skim through a number of general examples without linking them or elaborating upon them.
Mechanics--may lack control over grammatical conventions or exhibit "oral" style.

4 **Organization & Coherence**--a large paragraph or several shorter paragraphs, showing more organization and development than a 3. One or more sentences may cluster around a main point without necessarily making it clearer, but usually, a clear cause and effect relationship is expressed.
Mechanics--may contain many surface errors which are easily editable.

5 **Organization & Coherence**--more developed than a 4; essay supplies 2-3 sentences of specific detail to explain each example or explores a single example in depth; good use of transitions and linking words.
Mechanics--may contain many surface errors which are easily editable.

6 **Organization & Coherence**--several paragraphs developed in depth; several ideas or examples explored with specific details; essay uses a variety of transitional words to link thoughts together. Longer and deeper than a 5 essay, with more sophisticated sentence structures.
Mechanics--may contain some surface errors which are easily editable.

7 **Organization & Coherence**--fluent, idiomatic, capable of English 28 work; essay may still have minor mistakes, but exhibits greater depth and/or greater length than a 6.
Mechanics--fewer surface errors than a 6.

8 **Organization & Coherence**-- ready for English 101; essay is persuasive, poised, and substantial, elaborating its points with grace and style.
Mechanics--control over almost all grammatical conventions.

Commercial ESL Tests Reviewed for Use in LACCD
ESL Placement Procedure

Off-the-shelf Tests	Areas Tested
Basic English Skills Test, Literacy (BEST-L) Center for Applied Linguistics, Washington, DC (45 minutes)	reading writing
Comprehensive English Language Test (CELT) McGraw-Hill Book Company, NY (2 hours, 5 mins)	structure listening vocabulary
English-second-language Placement Test (EPT) ESL Teachers' Resource Instructor, San Francisco (30 minutes)	structure
Michigan English Placement Test (MEPT) University of Michigan, Ann Arbor, MI (1 hour, 15 mins)	structure listening vocabulary reading
Secondary Level English Proficiency Test (SLEP) Educational Testing Service, Princeton, NJ (1 hour, 30 mins)	listening reading
Structure Tests - English Language (STEL) Harper & Row Publishers, Scranton, PA (30 minutes)	structure
Test of English Proficiency Level (TEPL) Alemany Press, Hayward, CA (60 minutes)	structure reading writing
Secure Tests	
Preliminary Test of English as a Foreign Language (Pre-TOEFL) Educational Testing Service, Princeton, NJ (1 hour, 15 minutes)	structure & written expression listening vocabulary & reading
General Tests of English Language Proficiency (G-TELP), Levels 3, 2, 1 San Diego State University, San Diego, CA (Level 3, 90 minutes; Level 2, 110 minutes; Level 1, 105 minutes)	grammar listening reading

ESL Test Breakdown by Skill Areas

	# of Items	Testing Time	Comments
STRUCTURE			
CELT	75	45 mins	
MEPT	30	*	
EPT	50	30 mins	(3 options/item; 4 forms)
STEL	50	30 mins	(3 levels; 2 forms each)
TEPL	88	**	(5 options/item; items embedded in reading passages)
LISTENING			
CELT	50	45 mins	20 qs, 20 sts, 10 dialogues
MEPT	20	25 mins	3 options/item; qs, sts
SLEP	75	45 mins	25 pics, 20 sts, 12 map, 18 conv
READING			
MEPT	20	*	(sen-level only)
SLEP	70	45 mins	(27 based on pics; 43 rding comp, gr, voc based on passages)
TEPL		**	
VOCABULARY			
CELT	75	35 mins	(35 w/blank in
MEPT	30	*	(all w/blank in sen)

---

* 50 minutes to complete all three sections (str, rding, voc)
** 60 minutes to complete all sections (5 tasks); speeded test

Sample Questions

Section 1

The first section of the SLEP test measures ability to understand spoken English and is 35-40 minutes long. It is divided into four parts, with four different types of questions.

Part A

For the first type of question, the student must match one of four recorded sentences with a picture in the test book. The sentences are spoken only once and are not printed in the test book. This part contains items dealing with correct recognition of minimal pair contrasts, juncture, stress, sound clusters, tense, voice, prepositions, and vocabulary.

Sample Questions

Note: Pictures are for illustrative purposes only. Actual pictures and drawings in the test booklet are two to four times larger than sample pictures in this brochure.

1. On tape:
   Look at the picture marked 1.

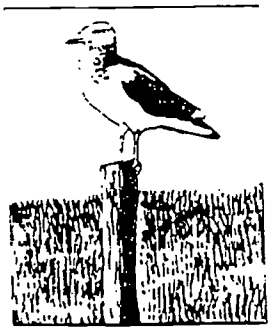
   On tape:
   (A) *There is an arrow in the sky.*
   (B) *The building has a tall tower.*
   (C) *The judge is bowing his head.*
   (D) *There is a toy in front of the building.*

1.

2. On tape:
   Look at the picture marked 2.

   On tape:
   (A) *The bird is standing on top of the pole.*
   (B) *The bird is flying over the fence.*
   (C) *The bird is digging in the sand.*
   (D) *The bird is eating the grass.*

2.



3. On tape:
Look at the picture marked 3.

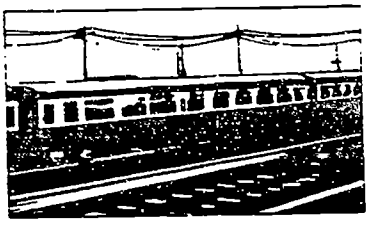
On tape:
(A) *There's a statue of a lion.*
(B) *The line is very straight.*
(C) *The wine is near the window.*
(D) *There's a lane near the building.*

3.

4. On tape:
Look at the picture marked 4.

On tape:
(A) *The brain is protected by bone.*
(B) *The train is on the track.*
(C) *The drain is stopped up.*
(D) *The rain is coming down.*

4.



Part B

These questions approximate the type of dictation exercises used frequently in English language classes: the student must match a sentence printed in the test book with a sentence heard on the tape. The questions focus on the relationship between structure and meaning.

Sample Questions

1. On tape: *The class can finish it in less than an hour.*

In test book: (A) The classes can't finish in half an hour.
(B) The class won't be finished for an hour.
(C) The classes will take at least an hour.
(D) The class can finish it in less than an hour.

2. On tape: *Why aren't they fixing the car?*

In test book: (A) Are they fixing the car?
(B) I'm fixing the car.
(C) Why aren't they fixing the car?
(D) The car has been fixed.

3. On tape: *While I was waiting for my sister, she got the news.*

In test book: (A) While I was waiting for my sister, she got the news.
(B) While my sister was waiting for me, she got the news.
(C) I was waiting for my sister to get the news.
(D) I was waiting for my sister when I got the news.

Part C

For the second type of question, the student refers to a map in the test book (see page 11). Streets and buildings on the map are labeled, and there are four cars, marked A, B, C, and D. The student must choose the one car that is the source of a brief conversation on the recording. The questions in this part assess a variety of linguistic, cultural, and pragmatic concepts. These include directions, recognition of building names and associated vocabulary, distance, and time.

Sample Questions

1. On tape:
   (man) *The museum has a special exhibit this week. Why don't we go?*
   (woman) *I'd like to very much. If we continue on Mackerel to the circle and go around to Salmon, we can park on Cod Lane.*
   (third voice) *Which car are the people in?*

2. On tape:
   (man) *I would like to find the way to the circle. From there, I know how to get home.*
   (woman) *It's not too hard. If we bear right into Bass and then go south on Salmon, we will end up at the circle.*
   (third voice) *Which car are the people in?*

3. On tape:
   (woman) *The judges are going to hear a very interesting case today. Let's stop at the courts.*
   (man) *That's a good idea. I'll go north at the next intersection and cross Pike Avenue. We can park in the lot across the street from the courts.*
   (third voice) *Which car are the people in?*

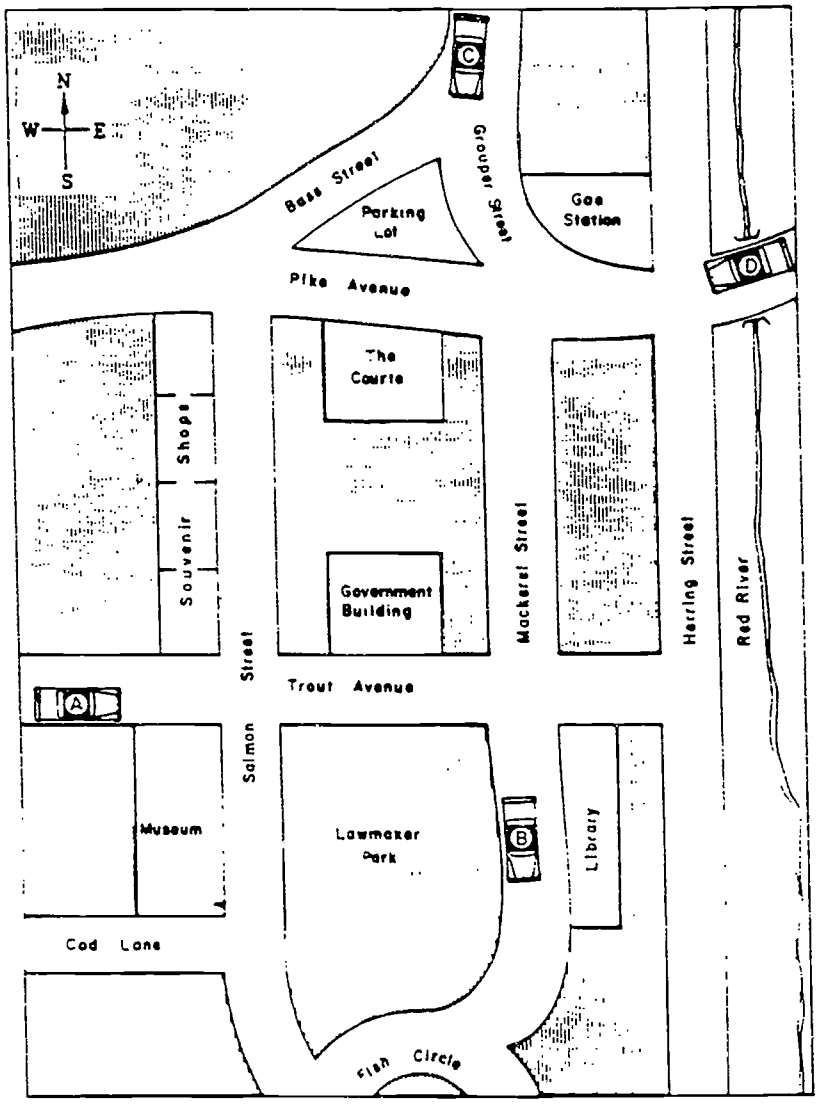
4. On tape: *He didn't know how to get to the gym.*

   In test book: (A) He didn't go to the gym.
   (B) He explained how to use the gym.
   (C) He told us to get to the gym.
   (D) He didn't know how to get to the gym.

5. On tape: *Bill has one brother and one sister, and so does Jane.*

   In test book: (A) Bill has one brother and one sister, and so does Jane.
   (B) Bill has one brother and a sister named Jane.
   (C) Bill and Jane are brother and sister.
   (D) Bill's brother and sister like to be with Jane.

Appendix A.3., con't: Listening Comprehension



Part D

The questions in this part are based on conversations, recorded by American high school students, that represent typical secondary school situations. The conversations take place in various parts of a school and deal with events that typically occur in each location. Conversations also deal with extracurricular activities, academic subjects, school closings, and holidays. For each recorded question, the student must choose one of four answers printed in the test book.

Sample Questions

1. On tape:

(Bob) *I heard that it is supposed to be a very good band. Since the game starts at 7:30, Nancy, I'll pick you up at 7.*

(Nancy) *That's fine. I'll be ready. It takes 15 minutes to get to the gym, so we'll have time.*

(third voice) *At what time will they arrive at the gym?*

In test book: (A) 6:45.
(B) 7:00.
(C) 7:15.
(D) 7:30.

For questions 2 and 3.

2. On tape:

(Nancy) *Jane, what are you going to wear to the game?*

(Jane) *I'm not sure yet. I don't want to have a heavy sweater on at the dance. It'll be pretty warm in the gym. I'll probably wear a light dress, even though the weather outside might not be so warm.*

(third voice) *What is the girl going to wear?*

In test book: (A) A heavy sweater.
(B) A heavy coat.
(C) Some light slacks.
(D) A light dress.

3. (On tape) *What is the girl's reason for this decision?*

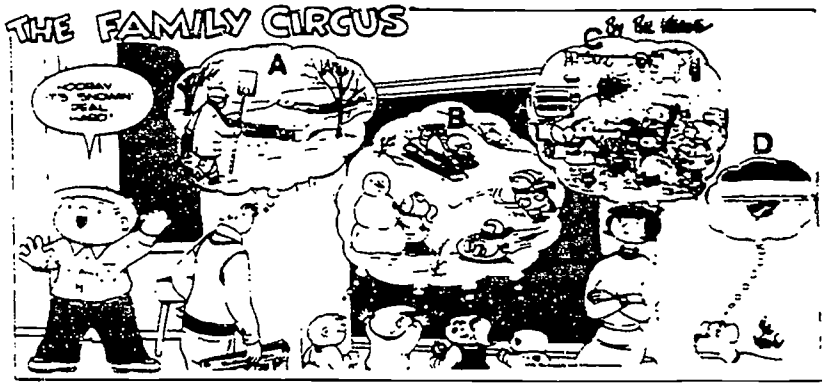
In test book: (A) She expects it to be cold outside.
(B) She expects it to be warm inside.
(C) It is going to snow.
(D) It will be very windy.

Section 2

The second section of the test is 40 minutes long and measures ability to understand Written English. The questions cover grammar, vocabulary, and reading comprehension. There are three parts to Section 2.

Part A

For each question in this part, the student must match the reaction of one of four characters in a cartoon with a printed sentence.



Sample Questions

1. All those wet clothes. The children will want to stay outside and I'll spend my time trying to keep them dry.

2. I can hardly wait to make the first snowball. I've been waiting all year to get back at her.

3. Oh, my aching back. The car will be covered and I'll have to shovel it out.

4. Isn't it great that school might be closed? I'd much rather have fun outside than stay in school. What better way to spend a snowy day.

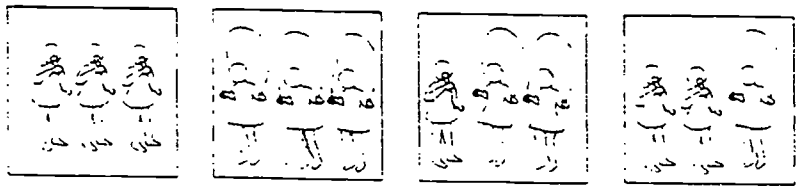
5. I'm going to be awfully hungry. I shouldn't have hidden that bone. It would have been better to leave it in the house.

Part B

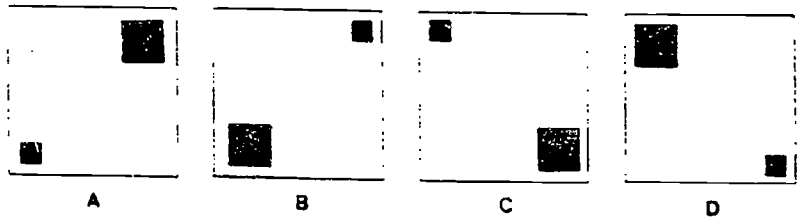
For the questions in this part, the student must match a printed sentence with one of four drawings. The particular focus of this item type is the use of prepositions, pronouns, adverbs, and numbers.

Sample Questions

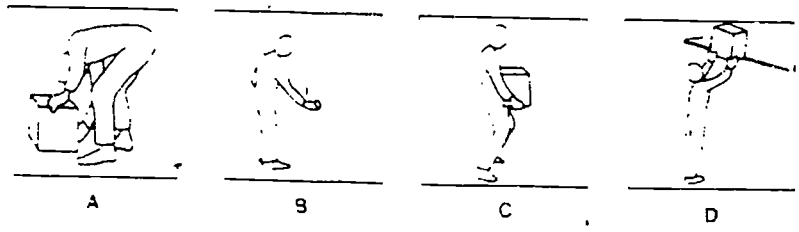
1. One girl is eating ice cream but two aren't.



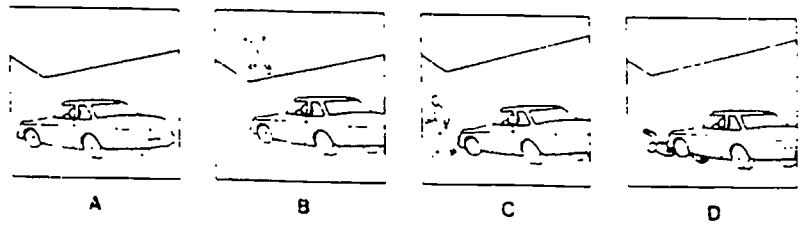
2. The small square is in the upper left corner.



3. He is bending over to pick up the box.



4. The car almost hit him while he was crossing the street.



Part C

This part of Section 2 contains questions of two types. In one, the student must complete passages by selecting the appropriate words or phrases from among four choices printed at intervals in the passages.

Sample Passage and Questions

1. Sound is something we [(A) hears. (B) hearing. (C) heard. (D) hear.] It comes to your

2. [(A) eyes (B) nose (C) ears (D) mouth] in different ways. It might be pleasant,

3. like the voice of a friend, [(A) when (B) as (C) or (D) since] unpleasant, like the screech

4. of a train's wheels on a railroad [(A) station. (B) track. (C) light. (D) conductor.] Some sounds are loud,

5. and some are soft; some are high, and some are [(A) full. (B) low. (C) quiet. (D) big.] Sound is

6. very [(A) importance (B) importantly (C) important (D) import] to us because it is the basic means of communication.

In the second type of question, the student must answer questions about the passage for which he or she supplied the missing words or phrases.

Sample Questions

7. What does screech in line 3 mean?

(A) noise (B) moon (C) place (D) piece

8. Which of the phrases below is another example of a pleasant sound, similar to the phrase in the sentence that begins in line 2, "like the voice of a friend"?

   (A) Like the ring of an alarm (B) Like the wail of a siren
   (C) Like the honk of a horn (D) Like the song of a bird

9. Which sentence below has almost the same meaning as the sentence that begins in line 5?

   (A) It is meaningful to communicate with sound.
   (B) The main way we communicate is with sound.
   (C) The meaning of sound is basic to communication.
   (D) In order to communicate, we need basic sounds.

Part D

In this part of Section 2, the student must read a short literary passage and answer questions about it.

Sample Passage and Questions

> The footsteps began about a quarter past one o'clock in the morning, a rhythmic, quick-cadenced walking around the dining room table. My mother was asleep in one room upstairs, my brother Herman in another; grandfather was in the attic, in the old walnut bed. I had just stepped out of the bathtub and was busily rubbing myself with a towel when I heard the steps. They were the steps of a man walking rapidly around the dining room table downstairs.

1. What did the writer hear?

   (A) A soldier marching (B) His brother snoring
   (C) His mother talking (D) A person walking

2. Where did the sounds come from?

   (A) The attic (B) The dining room
   (C) The bathroom (D) The stairs

3. What was most of the family doing?

   (A) Listening (B) Working (C) Bathing (D) Sleeping

Appendix B.1a. Distribution of Scores on Variables Used in Placement, for All LACCD-Students Assessed: Spring 1991 through Spring 1992

SLEP LC (Short version)

Score	Freq	Pct	Cum Pct	Score	Freq	Pct	Cum Pct	Score	Freq	Pct	Cum Pct
1	17	0	0	16	190	2	12	31	301	3	46
2	36	0	1	17	179	2	14	32	362	3	50
3	27	0	1	18	210	2	16	33	420	4	54
4	40	0	1	19	161	2	17	34	464	4	58
5	37	0	1	20	192	2	19	35	484	5	63
6	46	0	2	21	197	2	21	36	505	5	67
7	51	0	2	22	221	2	23	37	512	5	72
8	62	1	3	23	226	2	25	38	521	5	77
9	70	1	4	24	202	2	27	39	557	5	82
10	77	1	4	25	231	2	29	40	470	4	87
11	102	1	5	26	260	2	31	41	415	4	91
12	110	1	6	27	265	3	34	42	355	3	94
13	120	1	8	28	292	3	37	43	299	3	97
14	112	1	9	29	349	3	40	44	247	2	99
15	158	1	10	30	351	3	43	45	84	1	100

Valid cases 10587

SLEP RC (Short version)

Score	Freq	Pct	Cum Pct	Score	Freq	Pct	Cum Pct	Score	Freq	Pct	Cum Pct
1	24	0	0	20	302	3	30	39	233	2	85
2	25	0	0	21	375	4	33	40	196	2	87
3	22	0	1	22	377	4	37	41	223	2	89
4	53	1	1	23	327	3	40	42	172	2	91
5	31	0	1	24	380	4	43	43	176	2	93
6	69	1	2	25	370	3	47	44	177	2	94
7	103	1	3	26	337	3	50	45	140	1	96
8	78	1	4	27	340	3	53	46	95	1	96
9	137	1	5	28	328	3	56	47	83	1	97
10	123	1	6	29	317	3	59	48	79	1	98
11	139	1	8	30	344	3	63	49	66	1	99
12	184	2	9	31	312	3	66	50	53	1	99
13	186	2	11	32	295	3	68	51	31	0	99
14	229	2	13	33	261	2	71	52	35	0	100
15	278	3	16	34	260	2	73	53	19	0	100
16	271	3	18	35	232	2	75	54	8	0	100
17	289	3	21	36	309	3	78	55	5	0	100
18	267	3	24	37	271	3	81				
19	323	3	27	38	228	2	83				

Valid cases 10587

Total SLEP Score (Short version)

Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct	Value	Freq	Pct	Cum Pct
3	10	0	0	36	113	1	16	69	175	2	71
4	9	0	0	37	146	1	17	70	216	2	73
5	12	0	0	38	137	1	19	71	173	2	74
6	8	0	0	39	127	1	20	72	203	2	76
7	6	0	0	40	103	1	21	73	172	2	78
8	5	0	0	41	141	1	22	74	187	2	80
9	20	0	1	42	143	1	24	75	181	2	81
10	26	0	1	43	138	1	25	76	181	2	83
11	30	0	1	44	168	2	27	77	158	1	84
12	22	0	1	45	149	1	28	78	166	2	86
13	27	0	2	46	129	1	29	79	131	1	87
14	22	0	2	47	114	1	30	80	144	1	89
15	23	0	2	48	159	2	32	81	143	1	90
16	37	0	2	49	177	2	33	82	112	1	91
17	15	0	3	50	185	2	35	83	131	1	92
18	36	0	3	51	194	2	37	84	126	1	93
19	44	0	3	52	168	2	39	85	124	1	95
20	55	1	4	53	157	1	40	86	77	1	95
21	40	0	4	54	188	2	42	87	75	1	96
22	34	0	5	55	176	2	43	88	73	1	97
23	75	1	5	56	189	2	45	89	73	1	97
24	63	1	6	57	186	2	47	90	68	1	98
25	76	1	7	58	213	2	49	91	41	0	98
26	72	1	7	59	208	2	51	92	48	0	99
27	60	1	8	60	205	2	53	93	20	0	99
28	78	1	9	61	197	2	55	94	28	0	99
29	88	1	9	62	205	2	57	95	12	0	100
30	89	1	10	63	200	2	59	96	23	0	100
31	93	1	11	64	223	2	61	97	21	0	100
32	108	1	12	65	217	2	63	98	2	0	100
33	113	1	13	66	240	2	65	99	3	0	100
34	108	1	14	67	200	2	67	100	3	0	100
35	88	1	15	68	205	2	69				

Valid cases 10587

Essa. rating

Rating	Frequency	Percent	Valid Percent	Cum Percent
0	175	1.7	1.7	1.7
1	1553	14.7	14.7	16.3
2	2173	20.5	20.5	36.8
3	2384	22.5	22.5	59.4
4	2510	23.8	23.8	83.2
5	1207	11.4	11.4	94.6
6	466	4.4	4.4	99.0
7	95	.9	.9	99.9
8	14	.1	.1	100.0
Total	10587	100.0	100.0	

Valid cases 10587

CITY

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	27.00	10.2[illegible]	1	45	3847
RC	23.68	10.16	1	51	3851
SLEP	50.82	18.95	3	95	3833
ESSAY	3.22	1.50	0	7	3872
TOTAL	47.50	17.87	1	88	3872
GRADEAF	2.50	1.21	0	4	2067
PASSFAIL	.71	.46	0	1	374
LCPCT	59.99	22.[illegible]3	2.2	100	3847
RCPCT	43.05	18.47	1.8	92.7	3851
ESSPCT	40.23	18.70	.0	87.5	3872
GENDER	1.60	.49	1	2	3872
EDLEVEL	1.86	.67	0	4	3872
AGE	30.86	9.23	15.8	79.8	3872
PERSIST	.63	.48	0	1	3872

EAST

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	34.50	6.96	9	45	942
RC	30.82	9.60	5	54	941
SLEP	65.35	15.10	19	97	941
ESSAY	2.49	.94	0	7	942
TOTAL	45.24	11.48	12	84	942
GRADEAF	2.67	1.13	0	4	628
PASSFAIL	Variable is missing for every case.				
LCPCT	76.68	15.46	20.0	100	942
RCPCT	56.04	17.46	9.1	98.2	941
ESSPCT	31.12	11.79	.0	87.5	942
GENDER	1.56	.50	1	2	942
EDLEVEL	1.82	.67	0	4	942
AGE	28.39	8.42	10.3	71.3	942
PERSIST	.67	.47	0	1	942

PIERCE

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	39.18	5.39	8	45	493
RC	37.62	9.18	8	55	493
SLEP	76.80	13.44	17	100.00	493
ESSAY	4.66	1.28	1	8	490
TOTAL	66.04	13.43	19	98	493
GRADEAF	2.71	.98	0	4	347
PASSFAIL	Variable is missing for every case.				
LCPCT	87.07	11.99	17.8	100	493
RCPCT	68.40	16.70	14.5	100	493
ESSPCT	58.24	15.95	12.5	100	490
GENDER	1.60	.49	1	2	493
EDLEVEL	2.23	.76	1	4	493
AGE	27.20	8.04	16.8	59.7	493
PERSIST	.70	.46	0	1	493

TRADE

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	29.82	7.31	1	43	477
RC	22.59	7.93	1	47	474
SLEP	52.51	13.10	11	88	474
ESSAY	4.51	1.44	0	8	477
TOTAL	54.64	12.20	19	85	477
GRADEAF	1.89	1.32	0	4	83
PASSFAIL	.60	.49	0	1	115
LCPCT	66.27	16.25	2.2	95.6	477
RCPCT	41.08	14.42	1.8	85.5	474
ESSPCT	56.37	17.98	.0	100	477
GENDER	1.48	.50	1	2	477
EDLEVEL	1.57	.57	0	4	477
AGE	30.18	8.09	16.1	59.2	477
PERSIST	.42	.49	0	1	477

Note. "Persist" is variable coded "1" for individuals who received a grade of A-F or Pass/Fail versus "0" for those who did not do so. "Edlevel" reflects recoding of the "educational status" variable to a sacle (0,4) in which "high school graduate" is coded "2."

HARBOR

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	35.01	7.11	7	45	431
RC	32.53	9.14	4	55	431
SLEP	67.54	15.00	11	100.00	431
ESSAY	2.87	1.24	0	7	431
TOTAL	57.38	13.00	12	91	431
GRADEAF	2.99	1.09	0	4	216
PASSFAIL	1.00	.00	1	1	45
LCPCT	77.79	15.80	15.6	100	431
RCPCT	59.15	16.62	7.3	100	431
ESSPCT	35.85	15.53	.0	87.5	431
GENDER	1.50	.50	1	2	431
EDLEVEL	1.52	1.20	0	4	431
AGE	27.40	7.49	12.3	56.4	431
PERSIST	.61	.49	0	1	431

MISSION

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	27.47	10.62	1	45	1698
RC	27.87	11.18	1	52	1651
SLEP	56.13	19.59	3	96	1642
ESSAY	2.72	1.36	0	7	1717
TOTAL	47.76	18.45	4	91	1717
GRADEAF	Variable is missing for every case.				
PASSFAIL	.66	.47	0	1	855
LCPCT	61.05	23.59	2.2	100	1698
RCPCT	50.67	20.34	1.8	94.5	1651
ESSPCT	33.98	17.00	.0	87.5	1717
GENDER	1.54	.50	1	2	1717
EDLEVEL	1.32	1.16	0	4	1717
AGE	31.00	8.68	13.8	68.4	1717
PERSIST	.50	.50	0	1	1717

VALLEY

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	35.20	6.79	8	45	1616
RC	30.17	10.04	3	54	1616
SLEP	65.36	14.89	11	97	1616
ESSAY	2.85	1.27	0	7	1616
TOTAL	56.04	13.20	12	91	1616
GRADEAF	2.87	1.12	0	4	911
PASSFAIL	.74	.44	0	1	137
LCPCT	78.21	15.09	17.8	100	1616
RCPCT	54.85	18.26	5.5	98.2	1616
ESSPCT	35.65	15.88	.0	87.5	1616
GENDER	1.58	.49	1	2	1616
EDLEVEL	2.10	.83	0	4	1616
AGE	29.16	8.89	15.6	67.5	1616
PERSIST	.65	.48	0	1	1616

WEST

Variable	Mean	Std Dev	Minimum	Maximum	Valid N
LC	32.46	8.85	5	45	431
RC	26.78	10.06	2	55	1238
SLEP	59.44	17.89	7	100.00	428
ESSAY	3.72	1.22	0	8	1241
TOTAL	50.22	14.71	9	96	1241
GRADEAF	2.62	1.29	0	4	803
PASSFAIL	Variable is missing for every case.				
LCPCT	72.13	19.68	11.1	100	431
RCPCT	48.68	18.29	3.6	100	1238
ESSPCT	46.48	15.20	.0	100	1241
GENDER	1.51	.50	1	2	1241
EDLEVEL	1.71	1.06	0	4	1241
AGE	29.03	9.82	12.5	83.5	1241
PERSIST	.65	.48	0	1	1241

Appendix B.2. Descriptive Statistics for Placement Variables, by College and ESL Level (Last Course Enrolled)

College	ESL Level	N	Shortened SLEP LC M	LCSD Sd	RC M	RCSD Sd	SLEP M	SLSD Sd	Essay ESSAY M	ESSD Sd	Composite COMP M	TOSD Sd
City	1	944	17	9	14	7	32	15	1.9	1.2	29	13
1	2	613	21	7	18	6	39	10	2.4	1.1	36	9
1	3	554	27	7	22	6	49	9	3.1	1.0	46	8
1	4	736	32	5	27	6	59	10	3.7	.9	55	8
1	5	619	36	5	32	6	68	10	4.4	.9	64	8
1	6	331	38	6	37	9	75	13	4.9	1.1	71	12
1	7	75	36	5	34	9	70	12	4.8	1.1	67	11
East	2	71	28	8	21	7	49	12	1.6	.7	33	8
2	3	362	34	6	28	8	62	13	2.2	.7	42	8
2	5	163	38	4	35	7	73	10	2.9	.8	51	8
2	6	134	38	6	38	9	76	13	3.3	1.1	56	12
2	7	212	33	8	31	11	64	17	2.4	.9	44	12
Harbor	2	100	33	7	30	9	62	15	2.4	1.2	52	13
3	4	96	32	8	27	8	58	14	2.3	1.1	49	11
3	5	156	37	6	36	8	73	13	3.1	1.0	62	10
3	6	49	38	6	37	8	75	13	3.8	1.3	67	12
3	7	30	38	5	37	8	75	11	3.3	1.3	64	11
Mission	1	361	16	9	16	8	33	14	1.1	.4	24	12
4	2	217	23	9	22	9	45	16	1.9	.7	38	12
4	3	361	28	8	26	8	53	13	2.7	.9	48	8
4	4	374	32	7	31	8	63	12	3.4	.8	56	9
4	5	281	36	5	37	7	73	10	3.9	1.1	66	7
4	6	123	38	6	41	8	79	13	4.2	1.4	71	12
Pierce	4	138	39	5	36	8	75	11	4.1	.5	61	6
5	5	127	40	4	39	8	79	10	4.7	.6	67	7
5	6	113	42	3	44	7	85	9	6.0	.9	80	9
5	7	115	35	7	32	10	68	16	3.9	1.6	57	17
Trade	1	28	26	6	19	7	45	11	2.8	1.6	42	12
6	2	82	27	8	19	7	47	12	4.0	1.4	48	11
6	3	170	27	7	20	7	48	12	4.3	1.3	51	11
6	4	143	33	6	25	7	57	11	4.9	1.1	60	8
6	5	54	36	3	32	7	68	7	5.6	1.4	70	7
Valley	4	935	32	7	25	8	57	12	2.2	.9	48	9
7	5	267	38	5	35	7	73	11	3.2	1.0	62	9
7	6	206	40	5	40	9	80	11	4.1	1.3	71	10
7	7	208	39	4	38	8	77	11	3.9	1.2	68	10
West	3	510	2-	8	20	7	42	14	2.9	1.0	38	9
8	4	358	35	6	28	8	62	10	3.9	.9	52	10
8	5	257	38	4	33	8	71	9	4.4	1.0	62	9
8	6	116	40	3	41	7	82	6	5.2	1.0	71	10

Note. All levels not represented at all colleges.

Appendix B.3. Mean Score on Placement Variables, by Language Group and College: All Students Assessed

LANG (code)	College	N	Shortened SLEP LC	RC	SLEP	ESSAY	COMP	
					English			
1	City	1	41	25.7	22.7	50.1	3.4	45.4 City
1	East	2	30	33.4	29.0	62.5	2.2	41.6 East
1	Harb	3	8	39.6	35.8	75.4	2.6	61.8 Harbor
1	Miss	4	30	32.9	33.9	66.9	3.4	59.1 Mission
1	Pier	5	21	40.2	40.5	80.7	4.7	67.6 Pierce
1	Trad	6	8	24.5	22.8	47.3	3.3	45.1 Trade
1	Vall	7	38	34.5	34.2	68.7	3.0	58.6 Valley
1	West	8	23	39.4	34.1	77.1	3.8	57.1 West
					Armenian			
2		1	1019	23.1	18.1	41.3	2.5	38.3
2		2	16	33.3	25.8	59.1	2.4	42.7
2		4	12	28.0	25.0	53.0	2.3	45.4
2		5	18	40.1	35.9	76.0	4.5	64.7
2		7	154	32.7	23.1	55.8	2.6	48.7
2		8	4	12.0	20.3	18.0	2.5	35.0
					Chinese			
3		1	97	29.8	27.9	58.5	3.8	54.4
3		2	369	34.8	31.1	65.9	2.o	46.1
3		3	34	35.3	35.8	71.1	3.3	61.4
3		4	5	39.2	28.4	67.6	4.6	65.4
3		5	37	36.5	34.3	70.8	4.2	59.9
3		6	11	33.8	23.5	61.0	3.9	50.9
3		7	33	38.5	33.9	72.4	3.3	62.8
3		8	33	31.7	26.9	58.0	3.7	49.7
					Farsi			
4		1	21	29.4	24.4	53.8	3.1	49.5
4		2	2	38.0	18.5	56.5	2.5	42.5
4		3	3	35.0	27.7	62.7	3.7	57.7
4		4	3	41.7	37.7	79.3	3.7	68.7
4		5	62	39.5	34.4	73.9	4.2	61.4
4		6	1	28.0	10.0	38.0	4.0	44.0
4		7	103	35.0	26.2	61.2	2.5	52.1
4		8	73	33.7	25.4	54.8	3.6	48.8
					Filipino			
5		1	34	36.6	35.1	72.7	4.5	67.4
5		2	2	39.0	24.5	63.5	2.5	45.0
5		3	20	38.3	38.4	76.7	2.8	62.8
5		4	1	44.0	52.0	96.0	5.0	85.0
5		5	4	42.3	47.5	89.8	5.5	77.3
5		7	22	39.4	38.7	78.1	3.5	67.2
5		8	3	43.0	27.7	60.0	4.7	61.0

LANG (code)	College	N	Shortened SLEP LC	RC	SLEP	ESSAY	COMP
							Japanese
6	1	98	34.7	30.8	65.5	4.3	62.3
6	2	31	36.4	32.0	68.4	2.6	47.3
6	3	81	38.1	35.9	74.0	3.3	63.4
6	4	18	33.7	33.5	67.2	3.7	60.4
6	5	37	40.2	38.4	78.6	4.9	68.9
6	6	1	41.0	29.0	70.0	6.0	73.0
6	7	121	36.8	32.0	68.9	3.2	60.0
6	8	225	36.0	30.0	65.6	4.1	55.7
							Korean
7	1	574	32.8	29.3	62.1	3.6	56.8
7	2	15	36.8	32.8	69.6	2.5	46.9
7	3	17	37.9	33.9	71.9	3.0	60.9
7	4	21	34.7	31.0	65.7	3.4	58.6
7	5	39	38.8	37.9	76.7	4.7	66.3
7	6	25	34.6	24.6	59.2	4.1	57.1
7	7	98	35.7	32.4	68.1	2.7	57.1
7	8	77	35.3	27.5	66.1	4.0	53.1
							Russian
8	1	166	27.0	22.4	49.4	3.4	47.5
8	2	23	31.7	24.2	55.9	1.9	37.3
8	4	7	23.9	24.7	48.6	4.0	49.6
8	5	11	40.6	40.5	81.1	4.8	68.9
8	6	1	40.0	39.0	79.0	4.0	68.0
8	7	124	37.1	29.4	66.5	3.0	57.6
8	8	28	38.0	26.6	59.0	4.4	54.1
							Spanish
9	1	1499	24.8	23.1	48.0	3.1	45.3
9	2	336	34.4	31.5	66.0	2.4	45.1
9	3	153	31.1	29.4	60.8	2.5	50.8
9	4	1629	26.3	26.9	54.0	2.6	45.5
9	5	83	39.4	38.8	78.2	4.7	67.1
9	6	495	28.5	21.2	49.7	4.0	50.5
9	7	564	34.4	30.7	65.1	2.8	55.4
9	8	500	29.6	25.0	54.5	3.5	46.7
							Vietnamese
10	1	55	28.5	24.3	52.8	4.0	52.4
10	2	70	30.4	27.7	58.2	2.4	41.9
10	3	5	28.6	25.4	54.0	2.2	45.8
10	4	7	36.4	30.6	78.8	4.7	57.6
10	5	69	34.5	32.4	66.9	4.2	58.4
10	6	20	22.1	18.0	40.1	4.5	46.7
10	7	75	30.5	25.6	56.1	2.7	49.5
10	8	9	.	25.8	.	4.8	54.7

LANG (code)	College	N	Shortened SLEP LC	RC	SLEP	ESSAY	COMP
						Other Language	
11	1	263	31.3	26.8	58.4	3.8	54.4
11	2	38	37.9	32.0	69.9	2.6	48.3
11	3	27	38.2	35.4	73.6	2.6	60.4
11	4	47	30.9	35.3	66.8	3.6	58.7
11	5	111	41.0	40.1	81.0	4.9	69.9
11	6	12	31.3	21.8	53.0	4.4	54.6
11	7	273	36.2	31.0	67.1	2.9	57.3
11	8	259	31.7	27.0	59.3	3.9	50.9
						Lang Missing	
0	1	101	29.9	28.7	58.6	3.7	54.9
0	2	20	39.0	38.0	77.0	3.0	53.3
0	3	168	30.2	27.1	57.3	2.5	49.2
0	4	272	27.8	27.3	55.4	2.6	48.3
0	5	22	42.0	41.7	83.6	5.0	71.2
0	6	79	30.0	20.7	51.0	4.3	52.8
0	7	36	33.5	33.0	66.5	2.8	56.3
0	8	40	34.4	27.1	63.6	2.9	50.4

Number of cases read = 92

Appendix B.4. Relationship Between Placement Variables and Selected Demographic Variables, By College

College/ demographic	Correlation with designated demographic variable: LC	Shortened SLEP: RC	SLEP	ESSAY	TOTAL
City					
GENDER	-.20**	-.20**	-.22**	-.13**	-.19**
AGE	-.28**	-.23**	-.26**	-.16**	-.27**
EDLEVEL	.13**	.14**	.14**	.16**	.16**
East					
GENDER	-.13**	-.13**	-.14**	.04	-.05
AGE	-.08*	.04	-.01	.02	-.00
EDLEVEL	.14**	.18**	.18**	.11**	.16**
Harbor					
GENDER	-.12**	-.09*	-.10*	-.08	-.12**
AGE	-.26**	-.20**	-.23**	-.10*	-.24**
EDLEVEL	.29**	.33**	.33**	.16**	.31**
Mission					
GENDER	-.09**	-.08**	-.09**	-.04	-.09**
AGE	-.19**	-.03	-.12**	-.02	-.11**
EDLEVEL	.21**	.26**	.25**	.21**	.25**
Pierce					
GENDER	.00	-.02	-.01	.01	.01
AGE	-.10*	-.05	-.08	-.06	-.07
EDLEVEL	.08	.09*	.09*	.04	.07
Trade					
GENDER	-.05	.01	-.03	-.05	-.04
AGE	-.19**	-.11**	-.15**	-.11**	-.19**
EDLEVEL	.15**	.22**	.21**	.16**	.22**
Valley					
GENDER	-.16**	-.16**	-.18**	-.02	-.14**
AGE	-.21**	-.09**	-.16**	-.13**	-.18**
EDLEVEL	.09**	.11**	.11**	.15**	.14**
West					
GENDER	-.17**	-.11**	-.22**	.00	-.07**
AGE	-.23**	-.09**	-.20**	.00	-.06*
EDLEVEL	.18**	.12**	.16**	.26**	.18**
Total Sample					
GENDER	-.13**	-.13**	-.14**	-.06**	-.12**
AGE	-.25**	-.15**	-.21**	-.10**	-.19**
EDLEVEL	.20**	.19**	.20**	.18**	.22**

* - Signif. LE .05 ** - Signif. LE .01 (2-tailed)

Note. Correlation coefficients reflecting relationships between placement variables, on the one hand, and selected demographic variables, indicate generally that (a) females tend to earn lower scores on the placement variables than do males, (b) older students tend to learn lower scores than to younger students, and (c) students reporting more formal education tend to perform better than do students with less formal education. The general negative relationship between age and performance on the placement tests is stronger for listening comprehension than for either reading comprehension or writing ability (essay rating).

Appendix B.5. Results of Regression of Essay Rating on SLEP LC and RC in the All-LACCD Sample: All Students Assessed, All Colleges (Basis for Residual Analysis, Section III)

	Mean	Std Dev	Label
ESSAY	3.092	1.483	
LC	30.195	9.912	
RC	26.986	10.910	
SLEP	57.181	19.335	
TOTAL	50.410	16.770	

N of Cases = 10,587

Correlation:

	ESSAY	LC	RC	SLEP	TOTAL
ESSAY	1.000	.486	.491	.526	.788
LC	.486	1.000	.723	.921	.851
RC	.491	.723	1.000	.935	.844
SLEP	.526	.921	.935	1.000	.913
TOTAL	.788	.851	.844	.913	1.000

Equation Number 1 Dependent Variable.. ESSAY

Variable(s) Entered on Step Number
1.. RC
2.. LC

Multiple R	.52625
R Square	.27694
Adjusted R Square	.27680
Standard Error	1.26116

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	6447.54537	3223.77268
Residual	10584	16834.10949	1.59052

F = 2026.86160 Signif F = .0000

---------- Variables in the Equation ----------

Variable	B	SE B	Beta	T	Sig T
RC	.0399	.001627	.293	24.515	.0000
LC	.0409	.001791	.273	22.833	.0000
(Constant)	.7812	.039589		19.733	.0000

Residuals Statistics:

	Min	Max	Mean	Std Dev	N
*Predicted Essay	.9019	4.8152	3.0924	.7804	10587
*Residual	-4.2537	4.9499	.0000	1.2610	10587
*Z-scaled pred	-2.8068	2.2075	.0000	1.0000	10587
*Z-scaled res	-3.3729	3.9249	.0000	.9999	10587

Total Cases = 10,587

Table C.1. Descriptive Statistics for Placement Variables in Unit Samples (Students Classified by Course and Term ["SAMP"]): Parameters Used for Z-Scaled Transformation of Basic Test Variables

COL	LEV	SUBJ	COUR	SAMP		N	LC M	RC M	SHORT SLEP M	ESS M	COMP M	LCSD SD	RCSD SD	SLSD SD	ESSD SD	COMP SD
City	1	263	61		1	27	17	14	31	1.7	29	7	7	11	1.1	11
1	1	263	71	S91	1	94	20	16	37	2.0	33	8	8	15	1.1	12
1	1	263	71	F91	2	49	12	11	23	1.8	24	6	5	9	1.0	9
1	1	263	71	S92	3	87	18	14	31	1.7	28	9	7	13	1.1	12
1	1	263	74		1	51	21	17	38	2.1	34	9	7	15	1.2	12
1	1	263	74		2	50	14	13	27	1.8	26	8	7	14	1.2	12
1	1	263	74		3	81	18	14	32	1.7	29	9	8	15	1.1	13
1	1	263	75		1	69	20	18	38	2.1	34	8	7	14	1.1	11
1	1	263	75		2	40	14	14	29	1.7	26	8	7	15	.9	12
1	1	263	75		3	74	19	15	34	1.9	31	10	7	15	1.2	13
1	2	263	72		1	103	22	19	41	2.9	40	8	6	12	1.1	10
1	2	263	72		2	44	20	17	37	2.4	35	6	6	10	1.1	6
1	2	263	72		3	34	22	17	39	2.2	35	8	7	12	1.1	9
1	2	263	76		1	56	21	19	40	2.6	38	8	5	10	1.1	9
1	2	263	76		2	31	20	16	36	2.3	34	6	5	8	1.0	6
1	2	263	76		3	30	21	17	38	2.4	35	8	5	9	1.0	8
1	2	263	78		1	45	22	19	40	2.8	39	7	5	9	1.0	9
1	2	263	78		2	11	19	18	37	1.7	32	9	6	12	1.3	9
1	2	263	78		3	27	22	18	40	2.4	37	8	6	11	1.0	9
1	3	263	63		3	14	27	22	49	2.8	45	5	2	6	.7	4
1	3	263	73		1	22	27	22	49	3.7	49	7	5	8	.9	8
1	3	263	73		2	48	27	21	48	3.1	46	5	5	8	1.0	7
1	3	263	73		3	35	26	22	47	3.1	45	7	5	9	.8	7
1	3	263	77		1	37	28	23	51	3.2	48	7	5	8	1.0	6
1	3	263	77		2	30	25	22	47	3.0	44	6	4	7	.8	4
1	3	263	77		3	35	27	22	49	2.9	45	6	6	9	.9	7
1	3	263	79		1	48	29	22	52	3.4	49	6	5	7	.8	6
1	3	263	79		2	43	26	20	47	3.3	46	6	5	8	.9	7
1	3	263	79		3	31	28	22	50	3.1	47	6	6	9	.9	7
1	4	399	80		1	15	32	27	59	4.1	57	4	5	6	.7	3
1	4	399	83		1	37	31	27	58	3.9	55	6	6	10	.9	9
1	4	399	84		1	142	31	28	59	3.8	56	5	6	9	.8	6
1	4	399	84		2	70	33	28	61	4.0	58	4	6	8	.7	5
1	4	901	71		1	52	28	24	52	3.3	49	7	7	11	1.3	12
1	4	901	71		2	12	33	25	58	3.8	56	4	7	8	.6	5
1	4	901	71		3	24	33	26	59	3.8	56	6	5	9	.9	9
1	5	399	81		1	21	33	32	65	4.2	62	7	6	12	.9	10
1	5	399	81		2	25	36	32	68	4.5	65	4	4	7	.9	5
1	5	399	85		1	104	35	34	69	4.6	66	4	6	7	.8	5
1	5	399	85		2	87	36	33	69	4.6	66	4	5	6	.9	5
1	5	399	85		3	78	37	33	71	4.6	67	4	5	5	.7	4
1	5	901	72		1	38	32	29	61	4.1	59	8	9	15	1.3	14
1	5	901	72		2	16	38	31	69	4.8	66	3	5	6	.9	4
1	5	901	72		3	24	36	30	66	4.3	63	7	10	16	1.2	14
1	6	399	86		1	34	39	41	80	5.3	75	3	5	6	.7	5
1	6	399	86		2	49	39	39	78	5.0	73	3	6	8	1.0	8
1	6	399	86		3	35	40	41	81	5.4	77	3	5	6	.6	4
1	6	399	87		1	16	38	40	78	5.5	75	3	6	8	.6	4
1	6	901	73		1	44	32	30	62	4.1	59	7	10	16	1.2	15
1	6	901	73		2	12	41	42	83	5.7	79	3	4	4	.5	3
1	6	901	73		3	11	38	36	74	5.1	71	4	6	8	1.2	8
1	7	901	113		1	22	35	31	67	4.7	65	5	8	12	1.1	12
1	7	901	113		2	12	39	38	77	5.4	74	5	8	12	1.2	11
1	7	901	113		3	12	39	33	71	4.3	67	4	10	11	.7	9

COL	LEV	SUBJ	COUR	SAMP	N	LC M	RC M	SLEP M	ESS M	COMP SD	LCSD SD	RCSD SD	SLSD SD	ESSD SD	COMP SD
East	2	263	76	1	14	30	20	49	1.4	31	8	6	12	.5	5
2	2	263	76	2	19	29	22	51	1.7	34	7	8	12	.7	9
2	2	263	76	3	19	29	21	50	1.8	34	7	6	11	.8	8
2	4	399	83	1	20	33	27	61	2.4	43	7	8	13	.8	9
2	4	399	83	2	15	35	28	63	1.9	40	5	7	10	.5	5
2	4	399	83	3	19	31	28	59	2.3	41	8	10	16	.7	9
2	4	399	84	1	72	35	31	65	2.4	45	6	9	13	.7	9
2	4	399	84	2	52	35	30	65	2.3	43	5	7	11	.5	5
2	4	399	84	3	78	34	29	63	2.3	43	5	7	10	.7	6
2	5	399	85	1	27	38	36	74	3.1	53	4	6	8	.6	5
2	5	399	85	2	37	39	35	73	2.9	52	3	7	8	.7	6
2	5	399	85	3	40	38	35	73	3.0	52	3	6	8	.7	6
2	6	399	86	1	20	40	41	81	3.6	59	4	7	9	.8	8
2	6	399	86	2	21	40	41	81	3.7	61	3	6	8	.8	7
2	6	399	86	3	33	38	37	76	3.0	53	5	8	10	1.0	10
2	7	901	113	1	65	33	31	64	2.4	44	8	10	17	.9	12
2	7	901	113	2	41	35	33	69	2.7	48	5	10	13	.8	9
2	7	901	113	3	27	33	33	66	2.5	46	8	10	17	.8	12
Harb	2	263	76	1	19	33	32	65	2.7	55	9	10	18	1.7	16
3	2	263	76	2	22	32	28	60	2.2	50	7	10	16	.9	12
3	2	263	76	3	13	37	35	72	3.2	62	4	6	9	.9	8
3	4	399	84	1	18	34	27	61	2.4	51	6	7	11	1.3	10
3	4	399	84	2	18	34	29	63	2.1	52	4	7	9	.8	7
3	4	399	84	3	23	34	29	62	2.3	52	7	6	10	1.1	7
3	5	399	85	1	18	40	39	79	3.6	68	3	5	7	1.0	5
3	5	399	85	2	36	38	39	77	3.1	64	4	7	9	.9	7
3	5	399	85	3	17	39	37	76	3.4	65	2	4	5	.9	4
3	6	399	86	1	12	40	38	79	4.4	71	3	6	8	1.2	8
3	6	399	86	2	17	39	38	77	3.6	66	5	7	10	1.4	11
Miss	1	400	1	1	35	13	11	24	1.1	21	9	7	14	.3	10
4	1	400	1	2	81	18	17	35	1.1	28	7	7	12	.3	8
4	2	400	2	1	21	24	20	44	2.0	39	10	8	16	.8	11
4	2	400	2	2	71	25	25	50	1.9	41	9	9	16	.5	11
4	3	400	3	1	56	28	22	50	2.8	46	10	6	11	.8	8
4	3	400	3	2	67	27	27	54	2.9	48	8	7	13	.9	8
4	3	400	3	3	16	28	31	59	2.8	51	6	7	11	.9	8
4	4	400	4	1	79	33	29	62	3.8	58	8	7	13	.5	9
4	4	400	4	2	90	31	34	64	3.3	57	6	7	10	.8	7
4	5	400	5	1	45	38	34	72	4.4	69	5	8	12	1.5	8
4	5	400	5	2	94	35	38	72	4.1	65	6	7	10	.8	7
4	6	400	6	1	43	39	40	79	4.1	72	8	10	17	1.7	15
4	6	400	6	2	34	38	43	81	4.2	72	4	6	8	1.1	7
Pier	4	399	84	1	38	41	37	78	4.0	62	4	6	9	.3	3
5	4	399	84	2	38	38	35	73	4.0	60	7	7	12	.5	6
5	4	399	84	3	19	38	34	72	4.1	60	4	9	12	.2	5
5	5	399	85	1	26	41	40	82	4.7	68	3	5	6	.5	3
5	5	399	85	2	34	41	38	79	4.6	67	3	10	12	.7	8
5	5	399	85	3	33	40	39	79	4.7	67	3	7	10	.6	6
5	6	399	86	2	18	40	42	82	5.6	75	3	6	7	.9	8
5	6	399	86	3	13	41	41	82	5.9	78	2	5	6	.3	3
5	6	399	87	2	11	43	48	91	6.5	85	2	4	5	.5	4
5	6	399	87	3	23	42	45	87	6.2	81	5	8	12	1.0	11
5	7	901	113	1	26	39	35	74	4.0	60	4	9	12	1.5	15
5	7	901	113	2	33	34	31	65	4.0	57	7	9	15	1.7	18
5	7	901	113	3	19	33	32	65	4.1	57	8	10	18	1.7	19

COL	LEV	SUBJ	COUR	SAMP	N	LC M	RC M	SLEP M	ESS M	COMP M	LCSD SD	RCSD SD	SLSD SD	ESSD SD	COMP SD
Trade	2	400	2	2	15	29	21	50	5.0	56	7	7	11	.8	8
6	3	400	3	2	72	27	20	47	4.5	52	7	7	11	1.0	10
6	4	400	4	2	38	33	23	56	4.8	59	6	6	10	1.2	10
6	4	400	4	3	24	32	24	56	4.4	56	6	9	14	.9	8
6	5	400	5	2	11	37	30	67	5.7	70	2	5	7	1.3	5
6	5	400	5	3	10	35	36	71	5.0	68	2	6	6	1.7	7
Vall	4	399	83	2	95	29	21	50	1.9	42	7	6	10	.8	7
7	4	399	84	1	53	37	27	63	2.7	54	5	6	8	.9	6
7	4	399	84	2	72	35	29	65	2.6	54	5	7	8	.8	6
7	4	399	84	3	46	36	28	64	2.5	54	5	5	7	.7	5
7	4	901	64	1	74	32	22	54	2.5	47	6	6	10	1.0	8
7	4	901	64	2	110	32	26	58	2.3	49	7	9	13	.9	10
7	4	901	64	3	60	32	27	59	2.3	50	6	6	9	.8	7
7	5	399	85	1	44	39	35	74	3.5	64	5	6	8	1.0	6
7	5	399	85	2	73	39	36	74	3.2	64	4	7	10	1.0	7
7	5	399	85	3	48	37	32	69	3.0	59	5	6	10	1.0	8
7	6	399	86	1	38	40	38	78	4.5	71	5	8	10	1.1	8
7	6	399	86	2	55	41	40	81	3.7	70	2	8	8	1.0	7
7	6	399	86	3	28	39	42	81	3.9	70	5	8	12	1.2	10
7	6	399	87	2	15	40	39	79	4.3	71	4	10	13	1.2	12
7	6	399	87	3	11	39	43	81	5.4	77	5	9	14	.8	11
7	7	901	111	1	36	39	38	77	4.4	70	5	10	12	1.2	12
7	7	901	111	2	29	40	42	82	4.5	74	4	8	12	1.2	11
7	7	901	113	1	27	40	34	74	3.7	66	3	7	8	1.0	5
7	7	901	113	2	46	40	36	76	3.3	65	3	7	9	.8	6
7	7	901	113	3	12	39	39	78	3.6	67	3	5	7	1.1	4
West*	3	400	3	1	86	24	17	40	2.4	38	7	6	12	1.1	10
8	3	400	3	2	23	28	27	56	2.7	40	7	8	12	.9	8
8	4	400	4	1	51	34	25	60	3.7	56	5	6	7	.7	5
8	5	400	5	1	71	38	32	69	3.9	63	4	6	8	1.0	7
8	6	400	6	1	29	40	42	82	5.2	77	3	5	6	1.0	6

Note. SAMP codes are 1 (Spring 1991), 2 (Fall 1991), and 3 (Spring 1992).

* Students with all SLEP scores only.

Number of cases read = 140 Number of cases listed = 140

Table C.2. Descriptive Statistics for Placement Variables, In Samples Aggregated By Course, Across Terms of Enrollment: By College and Subject

COLL	SUBJ	COUR	N	LCMN	RCMN	SLMN	ESMN	COMP	LCSD	RCSD	SLSD	ESSD	COSD
City	263	61	27	17	14	31	1.7	29	7	7	11	1.1	11
1	263	63	14	27	22	49	2.8	45	5	2	6	.7	4
1	263	71	230	18	14	32	1.8	29	9	7	14	1.1	12
1	263	72	181	21	18	39	2.6	38	8	6	11	1.1	10
1	263	73	105	27	21	48	3.3	46	6	5	8	1.0	7
1	263	74	182	18	15	32	1.9	30	9	7	15	1.2	13
1	263	75	183	18	16	34	1.9	31	9	7	15	1.1	13
1	263	76	117	21	18	38	2.5	36	7	5	9	1.0	8
1	263	77	102	27	22	49	3.0	46	6	5	8	.9	6
1	263	78	83	22	18	40	2.5	38	8	6	10	1.1	9
1	263	79	122	28	22	49	3.3	47	6	5	8	.9	7
1	399	80	15	32	27	59	4.1	57	4	5	6	.7	3
1	399	81	46	34	32	67	4.4	63	6	5	9	.9	8
1	399	83	37	31	27	58	3.9	55	6	6	10	.9	9
1	399	84	212	32	28	60	3.8	56	5	6	8	.8	6
1	399	85	269	36	34	70	4.6	66	4	5	6	.8	5
1	399	86	118	39	40	79	5.2	75	3	5	7	.8	6
1	399	87	16	38	40	78	5.5	75	3	6	8	.6	4
1	901	71	88	30	25	55	3.5	52	7	7	11	1.2	11
1	901	72	78	34	30	64	4.3	61	7	8	14	1.2	13
1	901	73	67	35	33	68	4.5	64	7	10	16	1.3	15
1	901	113	46	37	33	71	4.8	68	5	9	12	1.1	11
East	263	76	52	29	21	50	1.6	33	7	7	11	.7	8
2	399	83	54	33	28	61	2.2	41	7	8	13	.7	8
2	399	84	202	35	30	64	2.3	44	6	8	11	.7	7
2	399	85	104	38	35	74	3.0	52	3	6	8	.7	6
2	399	86	74	39	39	79	3.4	57	4	7	10	.9	9
2	901	113	133	34	32	66	2.5	46	7	10	16	.9	12
Harb	263	76	54	34	31	65	2.6	54	7	9	16	1.3	13
3	399	84	59	34	28	62	2.3	52	6	7	10	1.1	8
3	399	85	71	39	38	77	3.3	65	3	6	8	.9	6
3	399	86	29	39	38	77	3.9	68	4	6	9	1.3	10
Miss	400	1	116	17	15	32	1.1	26	8	7	13	.3	9
4	400	2	92	25	24	49	1.9	41	9	9	16	.6	11
4	400	3	139	27	25	53	2.8	48	9	7	13	.9	8
4	400	4	169	32	31	63	3.5	57	7	7	11	.7	8
4	400	5	139	36	37	72	4.2	66	6	7	10	1.1	7
4	400	6	77	38	41	80	4.2	72	7	8	14	1.5	12
Pier	399	84	95	39	36	75	4.0	61	5	7	11	.4	5
5	399	85	93	41	39	80	4.7	67	3	8	10	.6	6
5	399	86	31	40	42	82	5.7	76	3	5	7	.7	6
5	399	87	34	42	46	88	6.3	83	4	7	11	.9	9
5	901	113	78	36	32	68	4.0	58	7	10	15	1.6	17
Trad	400	2	15	29	21	50	5.0	56	7	7	11	.8	8
6	400	3	72	27	20	47	4.5	52	7	7	11	1.0	10
6	400	4	62	33	23	56	4.7	58	6	7	11	1.1	9
6	400	5	21	36	33	69	5.4	69	2	6	7	1.5	6
Vall	399	83	95	29	21	50	1.9	42	7	6	10	.8	7
7	399	84	171	36	28	64	2.6	54	5	6	8	.8	6
7	399	85	165	38	35	73	3.2	62	5	7	10	1.0	7
7	399	86	121	40	40	80	4.0	70	4	8	10	1.1	8
7	399	87	26	39	41	80	4.7	73	5	10	13	1.2	12
7	901	64	244	32	25	57	2.3	49	6	8	12	.9	9
7	901	111	65	40	40	79	4.5	72	5	9	12	1.2	11
7	901	113	85	40	36	76	3.5	65	3	7	8	.9	6
West	400	3	109	25	19	44	2.4	38	8	8	13	1.0	9
8	400	4	51	34	25	60	3.7	56	5	6	7	.7	5
8	400	5	71	38	32	69	3.9	63	4	6	8	1.0	7
8	400	6	29	40	42	82	5.2	77	3	5	6	1.0	6

Table C.3. Correlations of Placement Variables with Course Grade, and Standard Deviations of Variables in Corresponding Courses, Classified by College and Subject (DC [263], ENG [399], ESL [400], SPEECH [901])

				Shortened SLEP					Standard deviations				
COLL	SUBJ	COUR	N	List r	Read r	SLEP r	Essay r	Comp r	LC	RC	SLEP	Comp	Essay
City	263	61	27	.45	.22	.41	.43	.48	6.9	7.1	11.4	10.8	1.1
1	263	63	14	.02	.54	.23	.32	.44	5.3	2.2	5.6	3.8	.7
1	263	71	230	.31	.23	.31	.31	.38	8.1	6.6	13.2	11.1	1.1
1	263	72	181	.05	.15	.12	.13	.17	7.8	6.1	11.3	9.2	1.1
1	263	73	105	-.02	.16	.09	.32	.24	6.3	5.1	8.2	7.3	.9
1	263	74	182	.28	.19	.26	.33	.35	8.9	7.3	14.7	12.5	1.2
1	263	75	183	.12	.12	.13	.25	.20	8.7	7.3	14.6	12.3	1.1
1	263	76	117	.03	.08	.05	.33	.23	7.4	5.2	9.3	7.8	1.0
1	263	77	102	-.08	.08	-.02	.39	.24	6.3	4.9	8.0	5.7	.9
1	263	78	83	.17	.03	.14	.17	.21	7.9	5.6	10.4	8.9	1.1
1	263	79	122	-.04	-.04	-.05	.34	.15	6.0	5.3	8.0	6.7	.9
1	399	80	15	.10	.35	.35	-.16	.25	3.7	5.4	6.3	3.3	.7
1	399	81	47	.04	.16	.13	.26	.24	5.4	5.4	9.0	7.4	.9
1	399	83	37	.06	-.08	-.02	.30	.14	6.0	6.4	10.3	8.9	.9
1	399	84	212	.00	.10	.07	.10	.12	4.6	5.9	8.2	5.9	.8
1	399	85	269	-.01	.01	.01	.12	.09	3.9	5.2	6.3	4.7	.8
1	399	86	118	.09	.00	.04	.27	.18	3.0	5.4	6.9	5.6	.8
1	399	87	16	-.03	.34	.26	.00	.25	3.0	6.2	7.7	4.3	.6
1	901	71	88	.23	.24	.28	.44	.40	6.1	6.7	10.4	9.9	1.1
1	901	72	78	.47	.50	.55	.35	.52	6.8	8.1	13.6	11.9	1.2
1	901	73	67	.33	.21	.32	.25	.38	6.0	8.5	12.9	11.5	1.1
1	901	113	46	.34	.31	.35	.37	.41	4.5	8.6	11.7	10.8	1.0
East	263	76	52	.17	.21	.23	.33	.35	7.1	6.7	11.6	7.7	.7
2	399	83	54	.20	17	.20	.40	.39	6.5	8.5	13.4	8.0	.7
2	399	84	202	.15	.28	.26	.21	.33	5.6	7.6	11.4	7.0	.7
2	399	85	104	-.08	.41	.30	-.19	-.02	3.2	6.3	7.8	5.7	.7
2	399	86	74	-.04	.03	.01	-.22	-.20	3.8	7.0	9.5	8.4	.9
2	901	113	133	.40	.34	.39	.34	.41	6.7	10.1	15.6	11.4	.9
Harb	263	76	54	.28	.35	.34	.29	.39	6.7	9.0	14.8	12.3	1.2
3	399	84	59	.35	.35	.46	.32	.54	6.0	6.6	9.9	8.3	1.1
3	399	85	71	-.09	.12	.06	-.02	-.02	3.3	5.9	7.7	5.7	.9
3	399	86	29	-.14	.37	.20	.15	.18	3.9	6.6	9.3	9.6	1.3
Miss	400	1	116	.21	.19	.24	.01	.21	7.3	6.9	12.5	8.8	.3
4	400	2	92	.15	.14	.16	.07	.17	9.2	8.5	15.6	10.9	.6
4	400	3	139	.22	.16	.26	-.04	.23	8.5	6.3	12.2	8.1	.9
4	400	4	169	.02	.04	.03	-.08	.00	6.7	6.9	11.3	7.8	.7
4	400	5	139	.28	.11	.22	-.08	.21	5.4	7.2	10.4	6.9	1.0
4	400	6	79	.29	.45	.45	.17	.36	6.4	8.2	13.1	11.8	1.5
Pier	399	84	95	.29	.34	.35	.17	.40	4.8	7.1	10.7	4.8	.4
5	399	85	93	.36	.28	.34	.25	.42	3.2	7.4	9.6	5.8	.6
5	399	86	31	-.20	.19	.11	.13	.06	2.6	5.5	6.6	5.9	.7
5	399	87	34	.08	.54	.42	.16	.25	4.1	6.9	9.8	8.7	.9
5	901	113	78	.17	.31	.27	.14	.19	6.4	9.5	14.8	17.1	1.6
Trad	400	2	15	-.07	-.06	-.08	.54	.18	6.7	7.2	11.4	7.7	.8
6	400	3	72	.13	.20	.20	.26	.28	6.7	6.6	11.3	9.6	1.0
6	400	4	62	.26	.21	.30	.21	.29	6.0	7.2	11.2	9.2	1.1
6	400	5	21	.06	.29	.24	.17	.39	2.3	5.8	6.6	6.2	1.5

N = 59

COLL	SUB	COUR	N	Shortened SLEP List r	Read r	SLEP r	Essay Comp r	r	Standard deviations LC	RC	SLEP	Comp	Essay
Vall	399	83	95	-.01	.02	.00	.19	.11	6.3	5.6	9.2	6.9	.8
7	399	84	171	.05	-.07	-.03	.20	.09	5.0	6.2	8.0	5.7	.8
7	399	85	165	.16	.13	.16	.12	.20	4.7	6.5	9.3	7.1	1.0
7	399	86	121	.10	.10	.13	.02	.11	3.8	7.7	9.8	8.1	1.1
7	399	87	26	.22	.03	.10	.16	.15	4.6	9.5	13.3	11.6	1.0
7	901	64	244	.12	.04	.09	.10	.12	6.4	7.5	11.2	8.9	.9
7	901	111	65	.22	.33	.33	.06	.26	4.6	9.2	12.2	11.1	1.2
7	901	113	85	.22	.03	.10	.16	.15	3.3	6.5	8.2	5.5	.9
West	400	3	109	.16	.01	.10	.20	.17	7.4	6.5	12.0	9.2	1.0
8	400	4	51	.07	.09	.12	-.06	.07	4.9	5.5	6.9	5.1	.7
8	400	5	71	-.04	-.15	-.13	-.08	-.15	3.9	6.3	8.3	7.0	1.0
8	400	6	31	-.11	-.02	-.06	.30	.19	2.8	4.6	6.4	5.5	1.0

N = 59

References

Alderson, C. J., Krahnke, K. J., & Stansfield. C. W. (Eds.). (1987). Reviews of English language proficiency tests (437). Washington, D.C.: Teachers of English to Speakers of Other Languages.

Butler, F. A. (1989). Placement testing for a district-wide ESL matriculation model: Los Angeles Community College District. Los Angeles, CA: LACCD (Project Report).

California Community Colleges (1992). Standards, policies and procedures for the evaluation of assessment instruments used in the California community colleges. Los Angeles: Author.

Carlson, S., Bridgeman, B., Camp, R., and Waanders, J. (1985). Relationship of admission test scores to writing performance of native and nonnative speakers of English (TOEFL Research Report No. 19). Princeton, NJ: Educational Testing Service.

Carroll, J. B. (1967). The foreign language attainments of language majors in the senior year: A survey conducted in U.S. colleges and universities (Final Report, Contract OE-4-14-048). Cambridge, MA: Harvard University Graduate School of Education).

Educational Testing Service (1988). SLEP test manual. Princeton, NJ: Author.

Educational Testing Service (1991). SLEP test manual. Princeton, NJ: Author.

Educational Testing Service (1992a). TOEFL test and score manual. Princeton, NJ: Author.

Educational Testing Service (1992b). TOEFL Test of Written English Guide. Princeton, NJ: Author.

Guilford, J. P. (1965). Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company.

Hale, G. A. (1986). An overview of research related to the TOEFL. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the Second TOEFL Invitational Conference (TOEFL Research Report No. 21: pp. 10-16): Princeton, NJ: Educational Testing Service.

Holloway, D. M. (1984). Validation study of SLEP test. Gainesville, FL: University of Florida (M.A. thesis).

Oller, J. W. (1983). (Ed.). Issues in language testing research. Rowley, MA: Newbury House.

Maculaitis, J. D. (1982). Technical manual for Maculaitis Assessment Program: A multi-purpose test for non-native speakers of English in kindergarten through twelfth grade (MAC:K-12). San Francisco: The Alemany Press.

Pike, L. W. (1979). An evaluation of alternative item formats for testing English as a foreign language (TOEFL Research Report No. 2 and ETS RR-79-6). Princeton, NJ: Educational Testing Service.

Rudmann, J. (1991). A study of the Secondary Level English Proficiency Test (SLEP) at Irvine Valley College. Irvine, CA: Author (mimeo).

SPSS Inc. (1990). SPSS reference guide. Chicago, IL: Author.

Stansfield, C. S. (1984). Reliability and validity of the Secondary Level English Proficiency Test. System, 12(1).

Thorndike, R. L. (1949). Personnel Selection. NY: John Wiley.

von Schilling, J. M. (1988). A study to assess the level of English language proficiency of the middle and high school Cambodian population in the Richmond, Virginia, public schools, with an analysis of related student characteristics. Richmond, VA: Virginia Commonwealth University, unpublished doctoral disseration).

Wilson, K. M. (1989). Enhancing the interpretation of a norm-referenced second-language test through criterion-referencing: A research assessment of experience in the TOEIC testing context (TOEIC Research Report No. 1 & ETS RR-89-39. Princeton, NJ: Educational Testing Service.

Wilson, K. M. (1993). Uses of the Secondary Level English Proficiency Test (SLEP): A Survey of Current Practice (Draft project report). Princeton, NJ: Educational Testing Service.

Wilson, K. M. (in progress). A cooperative study of factors in performance on the Secondary Level English Proficiency Test in the Temple University-Japan context.