

DOCUMENT RESUME

ED 397 727

HE 029 327

AUTHOR Weimer, Don
 TITLE Applying Linear and Logistic Regression to a Required English Proficiency Test. AIR 1996 Annual Forum Paper.
 PUB DATE May 96
 NOTE 20p.; Paper presented at the Annual Forum of the Association for Institutional Research (36th, Albuquerque, NM, May 5-8, 1996).
 PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS Age Differences; *English; Failure; Grades (Scholastic); Grammar; Higher Education; *Identification; Institutional Research; *Language Proficiency; Liberal Arts; *Predictor Variables; Private Colleges; *Screening Tests; Sex Differences; Statistical Analysis; *Student Placement; Writing Ability
 IDENTIFIERS *AIR Forum; Cardinal Stritch College WI

ABSTRACT

Prompted by repeated failure on a required English language proficiency test by a large proportion of undergraduates, educators at Cardinal Stritch College (Wisconsin) conducted a study to determine predictors of failure or success on the test, and whether revision of course placement standards should be considered to improve test performance. Academic and demographic data on 1,587 students from 1987-92 were analyzed for predictors of failure on both grammar and essay sections of the test. Certain subgroups (males, non-whites, nursing majors, community college transfers, and students entering as transfers) consistently performed less well on both test parts. Scores on standardized tests, letter grades in English department courses, and age were consistently high predictors of success on the test. (Contains 16 references.) (Author/MSE)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED 397 727

Applying Linear And Logistic
Regression To A Required
English Proficiency Test

Don Weimer
Director of Institutional Research
Cardinal Stritch College
6801 N. Yates Rd.
Milwaukee, WI 53217
(414) 242-2815
dweimer@acs.stitch.edu

BEST COPY AVAILABLE

Presented to:
Association of Institutional Research
36th Annual Forum

May 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it
- Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

AIR

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

AE 029 327



for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Sixth Annual Forum of the Association for Institutional Research held in Albuquerque, New Mexico, May 5-8, 1996. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Jean Endo
Editor
AIR Forum Publications

Abstract

The problem discussed in this paper is faced by an independent Midwestern liberal arts college. Many students repeatedly fail a test of English proficiency which is required in order to proceed to upperclassman status. The College would like to determine predictors of success or failure on the test, and whether revision of course placement standards should be considered in order to improve success on that test. This paper will describe the test and the problem of repeated student failures, and will report the results of linear and logistic regression analyses to determine a) whether predictors of success or failure on this test could be established or b) whether distinct student groups likely to fail could be identified. If distinct student groups with a high probabilities of failing the EPRT could be identified, then placement standards for remedial and introductory courses may be modified.

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
LITERATURE REVIEW	3
HYPOTHESIS	6
METHODOLOGY	7
RESULTS.....	10
CONCLUSION.....	12
REFERENCES	13
BIBLIOGRAPHY.....	14

Introduction

Cardinal Stritch College, Milwaukee, WI, with total student enrollment of over 5000, is the second largest independent college in the state. Students are enrolled into three program or administrative divisions: namely the Divisions of Arts and Sciences, Education, and Business and Management. Undergraduate students in the Arts and Sciences and Teacher Education Divisions are required to pass an English Proficiency Requirement Test (EPRT) before registering for classes for the semester following completion of 45 credits (Bachelor degree) or 30 credits (Associate degree). This test, developed and scored internally by the English Department, consists of Grammar and Essay sections. The Grammar section tests for errors in the areas of fused sentences, fragments, comma splices, subject-verb agreement, pronoun usage, verb forms, modifiers, parallelism, all uses of commas, and apostrophes. This grammar section is graded on a scale covering a range of scores. The Essay section requires the student to address a given topic and formulate a thesis, body, and conclusion, while paying attention to the transition between ideas and overall clarity. A student may receive a grade of fail, low pass, pass, or high pass on the Essay section of the EPRT. For purposes of this paper, the grading method is considered pass/fail.

The English Department faculty expressed concern that many students have had difficulty passing the test, and asked the Institutional Research Department to conduct a study. Anecdotal evidence from the English Department suggested that many students take and fail the test repeatedly. Some never pass. This had led to frustration and demoralization for the student as well as the faculty.

Based on anecdotal information, the English Department faculty suspected that students failing the test were disproportionately concentrated among certain student populations and hoped that predictors could be determined. Areas of interest included whether transfer students were failing at a higher rate than students entering as freshmen, and/or whether students placed into remedial English classes were failing at a higher rate than students enrolled directly into freshman-level English courses. Another question was whether students taking English 101 and English 102 did significantly better than students not taking those regular beginning courses.

It was suspected, for example, that transfer students failed the test at a rate disproportionate to their numbers. If this were the case, there might be several reasons: this college admits a relatively high number of transfer students from the local community college and the local branch of the state university, many of whom transfer English courses to Stritch satisfying lower-level requirements. It was deemed conceivable that a passing grade in a comparable 100-level English course at one of these schools represented a lower standard than a passing grade at Cardinal Stritch College. Another reason for the suspicion that transfer students failed the EPRT at a disproportionate rate was that some students have transferred into and out of several schools in their careers as students. Perhaps this constant changing of schools was a proxy for the inability to succeed in any academic environment.

Another inquiry was whether students in some majors failed at a disproportionately higher rate. English Department faculty suspected that students enrolled in Nursing and Education majors were having relatively more difficulty on the test than students in other majors.

Finally, scores on standardized admission tests, specifically the American College

Testing (ACT) and the Scholastic Aptitude Test, (SAT) were of interest as predictors, as were sex and racial/ethnic group.

To summarize, if measurable and significant differences existed among student groups, then revision of course placement standards would be appropriate in order to help these students succeed. If students from the community college were failing at a disproportionate rate, it might be appropriate to revise course placement criteria for these students. Likewise, if Education or Nursing majors were more likely to fail than students enrolled in other majors, revising placement criteria for these students would be a viable option to increase their chances of passing the test.

Literature review

A review of literature in the Educational Resources Information Center (ERIC) demonstrated that certain measures are used in higher education in conjunction with specific methodologies to predict academic success. Johnson (1992) used a variety of demographic variables such as age, gender, and race combined with reading and language scores on standardized placement tests to predict grades in a first-quarter English course. In this study, age and language scores on one standardized test were significant predictors of students' grades, while gender, race, and scores on a different test were not significant.

Boese, Sheppard, Sandusky, and Ghoston (1990) examined the relationship between academic success, basic skills assessment test scores and recommended placement levels, and demographic and educational background variables. Academic success was measured in terms

of cumulative grade point average (GPA) and the number of semesters the students remained enrolled. Among other findings, age was a significant predictor of cumulative GPA but not of some assessment scores; being female was a significant predictor of higher English test scores; English and reading scores varied significantly between ethnic groups, as did the measures of academic success; and English course level significantly predicted the measures of success.

The validity of American College Testing Program (ACT) scores and high school grades for predicting grades in specific freshman courses was studied by Noble and Sawyer (1988). Equations were developed for specific college courses in English, mathematics, social studies, and natural sciences using ACT research data. The validity of the prediction equations was tested using cross-validation techniques. Further analysis tested to determine if the predictive equations varied among the types of institutions.

In studying the utility of using both the ACT composite score and the high school GPA, Myers and Pyles (1992) reported a significant relationship between these two independent variables and the grade-point average of freshman year. The study had been conducted to address objections about the use of the ACT score alone, because the ACT score alone may be a poor predictor of success for minority students. In a different study, ACT and other standardized tests scores, final GPA, and overall GPA were correlated in a study by Sibert (1989). The study concluded that the ACT was the strongest and most consistent of any of the measures. This study by Sibert takes place in the contemporary climate of using testing as a criterion for admission to Teacher Education programs as well as the teaching profession. The author argues that the multiplicity of tests administered to future teachers has been a drain on resources, and

results suggest that many are, at best, redundant and that over-testing of students planning to become teachers is occurring

One study by Chou and Huberty (1990) analyzed the effectiveness of a regression equation which had been used to predict freshman performance at the University of Georgia. Predictors included verbal and mathematics scores on the Scholastic Aptitude Test (SAT), high school GPA, sex, race, and high school grouping. They found that the equation which had been in operation was not effective and developed a replacement model. It was also found that the predictors' effectiveness varied with characteristics of the institution. A similar methodology was used by Riggs and Riggs (1990) to test the validity of several predictors against several indicators of student success in a teacher education program. The predictors included undergraduate GPA scores from the California Basic Education Skills Test, grades in Education prerequisite courses and scores from the National Teacher Examination. A reading methods course grade and GPA functioned as best predictors of performance, while standardized tests scores did not.

In summary, regression and correlation analysis of predictors are standard tools for the type of analysis proposed for the EPRT. Commonly used variables included predictors such as course grades, ACT or SAT scores, grade-point average, and demographic information such as age, sex, and ethnicity. A similar methodology using comparable independent variables is used in this research, first, to analyze the relationship between independent variables such as ACT and SAT scores, demographic information, grades on English courses, as well as other relevant data and, second, to develop a model of predictors of success or failure on the EPRT.

Hypothesis

It should be noted that this study makes certain assumptions, including that the grading process of the EPRT is valid and the standards are consistently applied. It is beyond the scope of this study to analyze and review this grading process, but certain points need to be made. The first is that the English Department does not use a group grading process as a means to standardize grading: essay papers are graded privately by individual instructors, and no attempt has been made to look at inter-faculty differences. This is not an issue with regard to the grammar portion, as the grammar portion covers basic sentence structures using grammar rules found in mainstream grammar texts. A second point is that the Essay section offers the student a choice of questions which often address contemporary political and social issues and may involve controversial topics. It is conceivable that responses to these controversial topics elicit different evaluations based on the political or social view of the grader, even though this may occur at the subconscious level. Because this topic deserves separate treatment, and absent evidence to the contrary, the stated assumptions of this paper are that the EPRT grading process takes place in an atmosphere of trust and that the process is valid and uniformly applied.

The premise of this study is that the drawbacks of a false positive are greater than the drawbacks of a false negative. Put another way, to place an individual inappropriately in a remedial class on the basis of this study would have more negative consequences than to place an individual in a class that was too difficult. This premise is induced for several reasons. The first relates to marketing: as a practical matter, it is not the intent of this study to place barriers in front of potential customers of the College. The second reason is that this College has in place a

system to provide services for students encountering difficulty in classes. A student incorrectly placed into a class that was too difficult would find appropriate assistance readily available.

In this study, then, the premise is that risks due to falsely assessing that students need remediation when they do not are greater than the risks due to falsely assessing that students do not need remediation when they do. Consequently, a standard used to evaluate and possibly change college placement policies, for instance, must have a very high level of predictability, or the current system should remain in place. In terms of the scientific method, the null hypothesis is that the current system should remain in place absent compelling evidence to revise the system.

In summary, two questions were considered: First, an objective basis was necessary to determine whether a deficiency in English skills existed as measured by the EPRT, and, as part of that question, could populations be identified or predictors be determined? The second question concerned the alternative courses of action available to deal with the deficiencies and balancing these alternatives against market realities. To paraphrase this latter question, if placement standards were changed, would that change adversely affect student admissions and, consequently, revenue? In considering these questions, there is a predisposition to maintain the current system, absent very strong evidence to the contrary.

Methodology

This section will report the methodology performed to determine whether measurable differences existed among these students and results of the study. A data file was generated

consisting of traditional undergraduate students during approximately 1987-1992. The file consisted of academic background data such as: American College Testing (ACT) and Scholastic Aptitude Test (SAT) scores, high school grade point average (GPA), letter grades for courses English courses numbered 010, 101, and 102 (English 010, English 101, and English 102, respectively) along with the semester and year the course was taken; freshman or transfer status; and the previous college(s) attended by transfer students, total number of transfer credits, previous college GPA transferred in, and major field of study. Demographic data including age, race, and sex were also included. EPRT scores were matched to the students' names and entered. The raw score of the Grammar section of the test was entered as shown on record. Essay section "High Pass," "Pass," and "Low Pass" scores were entered as "Pass," and "Fail" results were entered as such. After test scores were entered, records of student for whom no test scores were available were eliminated. The result was a data file of 1537 records.

Letter grades in English 010, English 101, and English 102 were assigned numerical values from 0 through 11: A = 11, A- = 10, B+ = 9, ..., F/W=0. Dichotomous academic and demographic variables were assigned the following values:

Previous colleges:	State University = 1, other = 0
	Community College = 1, other = 0
Major:	Education = 1, other = 0
	Nursing = 1, other = 0
Enrollment status:	Freshman = 1, Transfer = 0.
Race:	White = 1, Non-white = 0
Sex:	Female = 1, Male = 0
Essay test:	Pass = 1, Fail = 0

Of these 1587 students, 481 (30.3%) were male and 1106 (69.7%) were female. White students numbered 1366 (86.1%), non-white students numbered 179 (11.3%), with 42 (2.6%) unreported. Two hundred forty three (15.3%) of these students were Nursing majors, and 37

(2.3%) were Education majors. One hundred twelve (7.1%) students had previously attended the community college, and 91 (5.7%) had previously attended the state university.

These records indicated that 1496 (93.2%) passed Grammar portion of the test on their first attempt, 91 (5.7%) failed, with 19 (1.2%) missing. The minimum passing score was 36 points, average score was 41.8 points. Regarding the Essay portion of the test 196 (12.6%) of these students failed on their first attempt, 1365 (86.0%) passed, with 26 (1.6%) records missing.

Some student populations under study had significantly lower scores and passing rates on the Grammar and Essay portions of the EPRT. Specifically, non-white, male, nursing majors, transfers from the community college, and other students entering as transfers also received lower mean scores on the Grammar portion of the EPRT than did their counterparts. Non-white, male, nursing majors, transfers from the community college, and students entering as transfers failed the Essay portion of the EPRT more than did their counterparts. This leaves open the question of whether group means and percentages should appropriately be used as criteria for placement of student groups into specific courses. Consequently, the significance of each variable was examined in the framework of the entire student population. In other words, assigning group means and rates to individuals may not be appropriate if in the overall context of the student population the significance of these variables diminished.

A correlation matrix established that the Verbal section of the SAT and the English section of the ACT had, predictably, the highest correlation to success on the EPRT. Accordingly, of the set of standardized test scores, these were the only sections used in the analysis in order to avoid multicollinearity. A series of least squares linear regression equations was run to determine which combination of academic and demographic variables produced the

best prediction of the Grammar Score. A series of logistic regression equations was run to predict the probability of passing the Essay portion.

Results

Grammar

- 1) Standardized tests: both the ACT English score and the Verbal SAT score were consistently and significantly positive; the higher the score on either of these tests, the higher the result on the grammar portion of the EPRT.
- 2) English 010, English 101 and English 102 grades: English department courses were consistently positive, and usually significant. In equations including both courses, English 010 overshadowed English 101 as a predictor, as English 101 overshadowed English 102 when both appeared in the same equation. When English 010 entered the equation, the “match” of the equation to the data fell noticeably, possibly due to the smaller number of students enrolled in the course, thereby decreasing sample size.
- 3) High School GPA, College Transfer GPA: these indicators of previous academic success were inconsistent and not significant.
- 4) Nursing or Education Major: these majors were negative in all equations. However, neither approached any level of significance.
- 5) Previous college: in all equations, having transferred from the community college was a negative indicator, but never significant. Transferring from the state university was consistently positive, significant in some equations, not significant in most.

6) Freshman/transfer status: in all equations, having entered the Stritch as a freshman was positive, but not significant.

7) Age: in most cases, this variable was positive and often was significant. It appears that the older a student is, the more likely he or she is to do better on the Grammar portion of the EPRT.

8) Sex: this predictor was inconsistent. In some cases this variable was positive, in some cases negative, and never significant.

9) Race: status as a white student was consistently negative, although never significant.

10) Repeating the test: No equation could produce a reasonable "match" which could explain the number of times a student took the Grammar portion of the EPRT.

Essay

The combination of variables which produced the best match of the above set of equations was used to determine which, if any, combination would serve to predict the probability of passing the Essay portion of the EPRT. The most successful models were those that included English 010 and, consequently, contained a much smaller sample. This suggests that the logistic regression model serves as a better predictor for students who enrolled in English 010.

Of all the variables listed above, High School GPA, English 010 and English 101 grades were positive and significant factors which increase the probability of passing the Essay portion. Most of the remaining variables were inconsistent and do not approach any level of significance.

Only the state university as a previous college was consistently negative, but not at any level of significance.

As in the Grammar portion, no equation could produce a reasonable "match" which could explain the number of times a student took the Essay portion of the EPRT. The logistic regression equations examining each variable individually accurately predicted most variables between 80 and 90 percent of occurrences. Only the grade for English 010, the remedial English course, was noteworthy in the percent of incorrectly predicted failures. The equation using all variables accurately predicted 83.8 percent, and incorrectly predicted failures in 7% of cases.

Conclusion

In many respects, the findings in this study were consistent with the findings reported in the literature. Specifically, scores on standardized tests and letter grades in English Department courses were consistent and significant predictors of success on the EPRT, as was age. Other predictors, such as race and sex, provided mixed results.

Future research on this topic would include a study of the validity of the grading process of the Essay section of the EPRT and a study of the recently-implemented Grammar test. A study of the validity of the Essay section grading process would lead in a completely different direction and must wait for another day. Because a new Grammar test was recently implemented, another avenue would be to see if the results in this study would be replicated in a study of the new Grammar test. This would indicate consistency of the two tests, at the very least, and if results were different, then future research would be called for.

To summarize, this study found no reason to conclude that any identifiable student groups are more likely than any other to fail this test given the overall characteristics of students. Therefore, it is concluded that the College should take no action to change any placement or admission standards. In other words, the College should do nothing.

References

- Boese, L., Sheppard, M., Sandusky, S. and Ghoston, A. (1990). The relationship of academic skills to basic skills, educational background, and demographic characteristics: a retrospective study of Sacramento City College students. Sacramento, CA: Sacramento City College. (ERIC Document Reproduction Service No. ED 346 941)
- Chou, T. and Huberty, C.J. (1990). A freshman admissions prediction equation: an evaluation and recommendation. Athens, GA; University of Georgia. (ERIC Document Reproduction Service No. ED 333 018)
- Johnson, R.G. (1992). A study in predicting English Grades for first-time-in-college students using the Career Planning Program (CPP) and Multiple Assessment and Program Services (MAPS) reading and language scores, age, gender, and race. Lawrenceville, GA: Gwinnett Technical Institute. (ERIC Document Reproduction Service No. ED 367 392)
- Mason, R.D. and Lind, D.A. (1993). Statistical Techniques in Business and Economics. Homewood, IL: Richard D. Irwin, Inc.

Myers, R.S. and Pyles, M.R. (1992). Relationships among high school grades, ACT test scores, and college grades. Cleveland, MS; Delta State University. (ERIC Document Reproduction Service No. ED 353 317)

Noble, J.P. and Sawyer, R. (1988). Predicting grades in specific college freshman courses from ACT test scores and self-reported high school grades. Iowa City; IA: American College Testing Research Series. (ERIC Document Reproduction Service No. ED 322 162)

Riggs, I. M. and Riggs, M.L. (1990). A test of the validity of selected predictors of student success in a teacher education program. San Bernadino, CA; California State University. (ERIC Document Reproduction Service No. ED 324 324)

Sibert, P. C. (1989). Relationships among the ACT, PPST, NTE, ACT COMP, and the GPA. Cookeville TN; Tennessee Technological University. (ERIC Document Reproduction Service No. ED 305 374)

Bibliography

Costrell, Robert M. (1993). An economic analysis of college admission standards. Education Economics, 1 (3), 227-241.

Cotter, M., Erdman, P., Koch, B., Nagy, J., Ramer, S., Roberts, N., and Wiley, J. (1993). How TQM helped change an admissions process. Change, May/June, 36-40.

Hawkins, D. and Phillip, M. (1993) Calls to raise admissions standards open Pandora's box of issues for campuses, students. Black Issues in Higher Education. January 28, 32-34.

- McDonough, Patricia M. (1994). Buying and selling higher education. Journal of Higher Education. 65(4), 427-446.
- Norusis, M.J. (1993). Advanced Statistics. Chicago, IL: SPSS, Inc.
- Norusis, M.J. (1993). SPSS for Windows: Base system user's guide. Chicago, IL: SPSS, Inc.
- Peacock, Ross (1993) Designing an effective admissions rating system, The Journal of College Admissions. Spring, 21-25.
- Townsley, Michael K. (1993) A strategic model for enrollment-driven private colleges, Journal for Higher Education Management. 8 (2), 57-66.