DOCUMENT RESUME

ED 102 199 TM 004 256

AUTHOR Calkins, Dick S.; Whitworth, Randolph
TITLE Differential Prediction of Freshmen Grade Point Average for Sex and Two Ethnic Classifications at a Southwestern University.
PUB DATE [74]
NOTE 13p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *College Admission; Freshmen; *Grade Point Average; Grade Prediction; *Grades (Scholastic); *Individual Differences; Mexican Americans; *Predictive Ability (Testing); Sex Differences; Statistical Analysis; Test Bias
IDENTIFIERS Scholastic Aptitude Test

ABSTRACT

This study stems from attempts to develop prediction equations for freshman grade point average for two demographic classifications of students at a southwestern university. Although the particular results are probably of interest only to that university, certain aspects of the study should be of general interest, because, of the 3,237 students comprising the sample, 1,241 were classifiable as Spanish-surname individuals. Comparisons of the regression equations resulting from using SAT scores and high school quartile rank to predict freshmen grade point average indicate that different equations are necessary for different classifications of students. Since these data were obtained from an educational environment created by a semi-open admissions policy and thus represent no restriction in the range of SAT scores, the results of this study may be of special interest to individuals and educational institutions who deal with similar populations. (Author)

DIFFERENTIAL PREDICTION OF FRESHMAN GRADE POINT AVERAGE FOR SEX AND TWO ETHNIC CLASSIFICATIONS AT A SOUTHWESTERN UNIVERSITY

Dick S. Calkins and Randolph H. Whitworth

The University of Texas at El Paso

Whether for the purpose of conservation of resources, or for humanitarian or other reasons, most colleges and universities like to assess the prospects of academic success for prospective students before the student is admitted. Usually only those students who are predicted to succeed are then admitted. Such endeavors, however, often prove difficult, given the interaction between the state of the technology of testing and measurement and the political pressures for social justice. Kendrick (1964) expressed this problem in the following manner:

> If values were divorced from science, and admissions practice from both, it would be both unnecessary and impossible to say anything about the relationship of cultural deprivation to College Board scores. But sentiment suggests it is wrong that students from minority cultures so seldom attend college. And experience with tests shows fairly clearly that some tests misrepresent the intellectual competence of students whose early environment has been different from that of most of us. Beyond that we cannot go with assurance (p. 7).

However, if such misrepresentations of intellectual competence by tests can be discovered, perhaps compensations can be made in the interpretation of tests, more accurate predictions made possible, and social justice obtained.

Davis and Temp (1971) indicated two general ways that test scores which could possibly misrepresent the intellectual competence of minority students could be misused by college admission officials:

1. Bias in the use of an admissions test could be legitimately claimed if an educational institution excluded members of a minority group who obtained scores on the test below a certain level without attention to the fact that on the average minority subgroup members obtain lower

scores on the test than does the subgroup for which the arbitrary cutoff level was established.

and 2.* Bias would be evidenced if scores on a test used as an admissions screening device (a) were related to college success for the dominant subgroup but not for the minority subgroup, (b) were differentially related to college success such that error of prediction was greater for the minority subgroup, or (c) were similarly related to college success but members of the minority subgroup with a given score on the test are more successful in college than members of the dominant subgroup with the same score level.

Several investigations of possible bias in the use of the Scholastic Aptitude Test as a college screening device for members of the black culture have been reported. Reviews by Thomas and Stanley (1969), Kendrick and Thomas (1970), and Stanley (1971) provide extensive coverage of this literature. Later studies by Davis and Temp (1971), Kallingal (1971), Pfeifer and Sedlacek (1971), and Temp (1971) have also been reported. Investigations of differential prediction equations for males and females were reported by Stroup (1970) and Siegelman (1971) and reviews of previous literature can be found in Thomas and Stanley (1969) and Kendrick and Thomas (1970). These studies indicate that black students on the average score considerably lower than white students on both sections of the SAT, but in many cases when these socres are used along with high school grades to predict college grades, predictions for black students are at least as accurate as for white students. Prediction of grades for female students is generally better than for males.

---

*Thorndike (1971) presents a selection procedure which is designed to compensate for this type of bias.

There appear to be no published studies investigating the possibility of bias in the use of the Scholastic Aptitude Test as a college screening device for members of the Mexican-American culture. Borup (1971) has reported an investigation of the ACT as a screening device for this group.

This paper presents an investigation of the SAT as an admissions tool at a Southwestern university with a Mexican-American enrollment of 33-50%. The main emphasis of the study involves the investigation of the differential predictive validity of the Scholastic Aptitude Test and quartile rank in high school as indicators of grade point average in the freshman college year for classifications of students based on sex and ethnicity.

PROCEDURE

The data for the 3, 237 students utilized in this study were obtained from the records of The University of Texas at El Paso. This group constitutes all entering freshmen for the period Fall semester, 1969 to Fall semester, 1971 who had data for all variables. Since this university was pursuing essentially an open admissions policy during that period, the data represent a set which is unusually free of any restriction in the range of the variables. (Although students with low SAT scores were denied admission during each Fall semester, the admission policy was essentially open in that students were provisionally admitted in the Spring and Summer semesters without regard to SAT scores or high school grades.)

The data collected consisted of the scores on the Scholastic Aptitude Test (SAT), which has two subscales labeled Verbal and Quantitative, each of which is designed to have a national mean of 500 and a standard deviation of 100; quartile rank in high school class; grade point average attained for the first thirty or less college hours; and a determination of sex and of ethnicity (as indicated by the surname for each individual). Quartile rank

in high school was coded such that a one indicated that the individual was ranked in the highest quartile and a four indicated that the individual was ranked in the lowest quartile according to high school grades. These variables have several disadvantages, some of which are as follows:

1. For a correlational study, quartile rank is not the most expedient coding for a variable indicating achievement in high school, but it was the only form of data available for this variable.
2. Freshman grade point average is probably a poor indicator of college success for many reasons, the most important being that GPA is not necessarily comparable across subjects because of the variance in the difficulty of the courses chosen by various students.
3. The determination of ethnicity or culture by surname is not a particularly good classification scheme for this type of investigation, since persons with Spanish surnames may be products of an Anglo environment and persons with non-Spanish surnames may be products of the Mexican-American culture with accompanying language deficits.

RESULTS

The data were analyzed utilizing the CDC 3150 computer system at The University of Texas at El Paso and a program package for multiple linear regression documented in An Introduction to Linear Models by Ward and Jennings (1973). The results of these analyses are presented in Tables 1-4.

Table 1 shows the means, standard deviations, and F statistics for comparisons of the means for the groups based on sex and ethnicity for the variables SAT V, SAT Q, quartile rank in high school, freshman grade point average, and SAT composite score (SAT V + SAT Q). It should be noted that there are significant differences among the group means for all variables.

Consistent with previous findings for blacks, the means for SAT V and SAT Q appear to be lower for Spanish surname than for non-Spanish surname individuals, with apparent sex differences only for SAT Q, where males attain higher means than females. The means for quartile rank and grade point average appear to differ primarily between the sexes, with females attaining ranks closer to the top of their class and higher grade point averages.

Table 2 contains the correlations between the variables for each of the groups. It should be noted that all these correlations are of sufficient magnitude that it is unlikely that any represent random values from populations in which the correlations are zero. (At the .01 level with $\underline{df} = 500$, correlations above .115 are significant.) Also the correlational matrices for the groups appear to be fairly similar.

Table 3 shows the equation for predicting grade point average from knowledge of SAT V, SAT Q and quartile rank in high school for each group. In addition, this table shows the squared multiple correlation coefficient for each equation and information concerning the tests for evaluating statistically the relevance of each predictor. It should be noted that the squared multiple correlation coefficients are generally low and appear to be lower for the Spanish surname groups than for the non-Spanish surname groups. In addition, results of the F tests for relevance of the predictor variables indicate that SAT V and quartile rank are "good" predictors of grade point average for all the groups. SAT Q appears to be relevant for the prediction of grade point average for the non-Spanish surname individuals.

Table 4 presents the term by term comparison of the equations for the groups which were obtained to discover whether the various groups differ. These statistical tests involve comparisons of the constants or coefficients for a particular predictor variable across the groups to determine the

TABLE 1

Means and Standard Deviations for All Groups on All Variables

	Spanish Surname				Non-Spanish Surname						
	Female N = 530		Male N = 711		Female N = 804		Male N = 1192				
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*F*	*df*	*p*
SAT V	377.3	85.11	376.1	87.05	459.1	97.31	443.0	93.82	165.3	1, 3233	<.0001
SAT Q	384.4	82.52	416.6	94.36	439.5	89.48	476.5	99.47	138.3	1, 3233	<.0001
Rank	1.953	1.042	2.332	1.100	1.858	.9619	2.499	1.071	74.39	1, 3233	<.0001
HSPA	2.089	.8338	1.682	.9170	2.223	.9749	1.741	.9687	63.12	1, 3233	<.0001
SAT Composite	761.7	148.3	792.6	161.5	898.6	165.8	919.5	169.3	171.7	1, 3233	<.0001

TABLE 2

Correlations between All Variables for All Groups Separately

Variable	SAT V	SAT Q	Rank	GPA	SAT Composite
Spanish Surname Females					
SAT V	1.000	.565	-.348	.335	.888
SAT Q		1.000	-.428	.284	.881
Rank			1.000	-.338	-.438
GPA				1.000	.3507
Spanish Surname Males					
SAT V	1.000	.585	-.311	.246	.881
SAT Q		1.000	-.365	.226	.900
Rank			1.000	-.358	-.381
GPA				1.000	.265
Non-Spanish Surname Females					
SAT V	1.000	.576	-.343	.368	.897
SAT Q		1.000	-.411	.340	.877
Rank			1.000	-.380	-.423
GPA				1.000	.400
Non-Spanish Surname Males					
SAT V	1.000	.534	-.325	.324	.868
SAT Q		1.000	-.301	.335	.883
Rank			1.000	-.419	-.419
GPA				1.000	.3767

TABLE 3

Prediction Equations for GPA with SAT V, SAT Q and Rank as Predictors for Groups by Sex and Ethnicity

Group	Equation				R^2
Spanish Surname Female	GPA = 1.4145 + .0022 SAT V + .0006 SAT Q - .1887 Rank				.1702
		F	*df*	*p*	
	SAT V	20.42	1, 526	<.0001	
	SAT Q	1.38	1, 526	>.20	
	Rank	28.22	1, 526	<.0001	
Spanish Surname Male	GPA = 1.6005 + .0014 SAT V + .0004 STA Q - .2530 Rank				.1493
		F	*df*	*p*	
	SAT V	8.90	1, 707	<.004	
	SAT Q	.82	1, 707	>.30	
	Rank	65.22	1, 707	<.0001	
Non-Spanish Surname Female	GPA = 1.2040 + .0022 SAT V + .0012 SAT Q - .2653 Rank				.2159
		F	*df*	*p*	
	SAT V	30.99	1, 800	<.0001	
	SAT Q	7.47	1, 800	<.007	
	Rank	56.93	1, 800	<.0001	
Non-Spanish Surname Male	GPA = 1.1689 + .0016 SAT V + .0013 SAT Q - .2888 Rank				.2263
		F	*df*	*p*	
	SAT V	24.63	1, 1188	<.0001	
	SAT Q	17.01	1, 1188	<.0001	
	Rank	129.62	1, 1188	<.0001	

TABLE 4

Comparison of the Coefficients in the Prediction Equations for the Various Groups

Coefficient	Spanish Surname		Non-Spanish Surname		*F*	*df*	*p*
	Female	Male	Female	Male			
SAT V	.0022	.0014	.0022	.0016	.95	3, 3221	>.40
SAT Q	.0006	.0004	.0012	.0013	1.16	3, 3221	>.30
Rank	-.1887	-.2530	-.2653	-.2888	1.57	3, 3221	>.15
Intercept*	1.5551	1.2188	1.4673	1.1442	35.17	3, 3230	<.0001

*The intercept values are different than those shown in Table 3 because the values for Table 4 result from a different model. Since the results in Table 4 show that there are no significant differences between the coefficients for the various groups for predictor variables SAT V, SAT Q and rank, this indicates that the groups do not differ in terms of prediction using these variables and so values of the coefficients for these variables can be estimated from the data for all the groups. The intercepts shown in Table 4 are the values of the intercepts from a model in which all the groups had the same coefficients for each predictor -- i.e., all groups had the same coefficient for SAT V, the same coefficient for SAT Q, and the same coefficient for rank.

likelihood that the coefficients are random samples from populations with equal values, i.e., the groups do not differ in terms of the coefficients of a particular predictor. (This type of test is equivalent to the comparison of slopes and intercepts in the analysis of covariance.)

The results of the F tests shown in Table 4 indicate that the groups probably only differ in terms of the regression constants (intercepts). This implies that although the most accurate prediction of freshman grade point average would stem from the use of a regression equation for each group with different weights for each predictor, insufficient difference among these equations is evidenced to statistically disallow the reduction of these four equations to four equations which use the same predictor weights for each group but with each group having a different intercept. The intercepts for these equations are those shown in Table 4, and the common set of weights are .0018 for SAT V, .0010 for SAT Q, and -.2586 for quartile rank. Since the prediction equations differ statistically only in terms of the intercepts, and since differences among the intercepts appear to be attributable to differences in the sexes rather than to differences in ethnicity, differences in the prediction equations are probably attributable only to differences due to sex.

DISCUSSION

It has often been noted that in this country members of some minority groups on the average obtain lower scores on intelligence and aptitude tests than members of the white majority. Unless those who use the scores from such tests are willing to assume that such tests indicate true differences in aptitude between such subgroups (and hence that such minority groups are in fact inferior in terms of the aptitudes assessed), they should face the problems involved with the use of tests which produce different distributions of scores for different demographically definable groups. Failure to take such group differences

into consideration when these test scores are being utilized can lead to the often cited charges that such tests are biased and that the use of such tests is discriminatory.

The problems involving the use of such tests can be ameliorated to some degree through (a) the establishment of different expectations for scores on the tests for different groups of individuals, and (b) the investigation of possible differential predictability of the test scores for different groups of individuals. This investigation was undertaken to investigate possible sources of bias which could result from the use of the SAT and rank in high school as predictors of freshman grade point average at a Southwestern university with a sizeable Mexican-American enrollment.

Results of this investigation indicate that:

1. There is considerable difference between the SAT V and SAT Q scores for Spanish surname and non-Spanish surname individuals, with non-Spanish surname individuals generally attaining higher scores. In addition, females make higher grades in college and are nearer the top of their high school classes in terms of grades than males.

2. The accuracy of the prediction equations for all the groups as indicated by the squared value of the multiple correlation coefficient is low, so prediction of freshman grade point average is not good for any group, but predictions are slightly poorer for Spanish surname than for non-Spanish surname groups.

3. Quartile rank in high school is the best predictor of freshman grade point average, with SAT V next best and SAT Q the poorest.

4. Statistically there appears to be no justification for utilizing different predictor weights for SAT V, SAT Q or rank for the various

groups. The groups probably do differ in terms of the regression constants however. Overall the differences in the prediction equations appear to be attributable to differences in the sexes and not to differences in ethnicity.

This study indicates that in terms of the magnitude of SAT scores Spanish surname individuals tend to score about a half a standard deviation below non-Spanish surname individuals. However, in terms of prediction of freshman grade point average, there does not appear to be any bias against Spanish surname individuals, since the regression constants and slopes for SAT V, SAT Q and quartile rank did not differ significantly for the groups based on ethnicity. On the other hand, there does appear to be some bias against women since the regression equations over-predict female grade point average as compared to males. This indicates that females probably can be successful at this university with lower SAT scores and quartile rank than can males.

REFERENCES

Borup, Jerry E. The validity of American College Test for discerning potential academic achievement levels: Ethnic and sex groups. Journal of Educational Research, 1971, 65, 3-6.

Davis, Junius A. and Temp, George. Is the SAT biased against black students? College Board Review, No. 81, Fall 1971, 4-9.

Kallingal, Anthony. The prediction of grades for black and white students at Michigan State University. Journal of Educational Measurement, 1971, 8, 263-266.

Kendrick, S. A. College Board scores and cultural bias. College Board Review, Winter 1964-65, 7-9.

Kendrick, S. A. and Thomas, Charles L. Transition from school to college. Review of Educational Research, 1970, 40, 151-173.

Pfeifer, C. Michael Jr., and Sedlacek, William E. The validity of academic predictors for black and white students at a predominantly white university. Journal of Educational Measurement, 1971, 8, 253-262.

Siegelman, Marvin. SAT and high school average predictions of four-year college achievement. Educational and Psychological Measurement, 1971, 31, 947-950.

Stanley, Julian C. Predicting college success of the educationally disadvantaged. Science, 1971, 171, 640-647.

Stroup, Atlee L. The prediction of academic performance from personality and aptitude variables. Journal of Experimental Education, 1970, 38, 83-86.

Temp, George. Validity of the SAT for blacks and whites in thirteen integrated institutions. Journal of Educational Measurement, 1971, 8, 245-252.

Thomas, Charles L. and Stanley, Julian C. Effectiveness of high school grades for predicting college grades of black students: A review and discussion. Journal of Educational Measurement, 1969, 6, 203-216.

Thorndike, Robert L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Ward, Joe H. and Jennings, Earl. Introduction to Linear Models. Inglewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.