# THE PREDICTION OF GRADES OF FIRST-TIME FLORIDA FRESHMEN AT THE FLORIDA STATE UNIVERSITY

W. L. Bashaw
Florida State University

This study deals with the prediction of grade point averages earned by first-time Florida freshmen at the Florida State University. Independent variables used in this study were the total percentile score on the Florida State-Wide Twelfth-Grade Testing Program battery, high school academic grade point average, American Council on Education Psychological Examination raw score, the English percentile score from the Florida State-Wide Twelfth-Grade Testing Program, and sex.

This study was done to investigate the utility of variables used as admission criteria by the Florida State University and of other pertinent variables used as predictors of the quality of work produced by Florida high school graduates during their first semester at the University. Admission requirements at FSU are graduation from an accredited Florida secondary school and attainment of 200 percentile points on the Florida State-Wide Twelfth-Grade Testing Program, or approval after special consideration by the FSU Admissions Committee.

The utility of these variables as predictors has been investigated three times since 1956 by the FSU Office of Institutional Research and Service.

The first study (1) was based on 947 of the 1,201 first-time Florida freshmen who entered the University in the fall of 1955. They were not screened on the basis of their Florida State-Wide Twelfth-Grade Testing Program scores (TGTP). Correlations between their FSU grade point averages (UG) and TGTP, American Council on Education Psychological Examination total scores (ACE), and high school academic grade point averages (HSG) were reported. These correlations appear in Table 1.

Table 1

Correlations of Independent Variables
with University Grades:
1956, 1957, and 1958

| Variable | 1956 | 1957 | 1958 |
|----------|------|------|------|
| TGTP | .62 | .57 | .51 |
| HSG | .63 | -- | .62 |
| ACE | .44 | .48 | .50 |

The study was repeated in 1957 using 1,169 of the 1,300 new first-time Florida freshmen (2). The group was selected according to the admission criteria previously described. The correlation between (TGTP) and (UG) was lower than that reported the previous year due to the curtailment of TGTP distribution occasioned by application of selection criteria. The obtained correlations appear in Table 1.

The study was again repeated in 1958 using 95% of 1,137 first-time Florida freshmen. The obtained correlations appear in Table 1.

These three studies reported consistently high positive correlations between the various predictors and UG. A further re-study of these relationships was considered necessary due to the continued curtailment of the TGTP distribution by admissions selection.

More stringent selection requirements, specifically raising the minimum acceptable score on TGTP, have the effect of curtailing the TGTP distribution represented by first-time Florida freshmen and consequently diminish the relationship between it and the criterion variable. Therefore, the restudy dealt not only with the predictors used in past studies, but also with sex and English.

## Method

### Sample

The population for this study consisted of the 2,014 students who enrolled for the first time at Florida State University in the fall semester, 1960-61, and who were graduates of Florida high schools. None of these students was a transfer from another college or university.

Two samples were drawn from this population. Every fourth name on an alphabetical list of these students was chosen for the first sample. From these, all students with incomplete data were discarded. This sample contained 435 (21.6%) students. A second sample was drawn consisting of every tenth remaining student on the list. All students with incomplete data were again discarded leaving a sample of 111 (5.5%). The latter sample was used for cross-validation.

### Variables

The dependent variable was University grade point average (UG). Grade point averages were based on the following weights: A=4, B=3, C=2, D=1, and F=0.

The independent variables were the Florida State-Wide Twelfth-Grade Testing Program (TGTP), English (ENG), American Council on Education Psychological Examination (ACE), high school academic grade point average (HSG), and sex.

The English score used was the percentile score from the Cooperative English test which is part of the TGTP battery.

The ACE score used is the total raw score on the American Council on Education Psychological Examination. This examination was administered by the University during the fall Freshman Orientation preceding the beginning of classes.

HSG is based on high school transcripts. This variable includes only grades in academic courses. Academic courses include English, social studies, history, mathematics, foreign languages, and natural science. Credit must have been earned during the first seven semesters of high school (the last semester was omitted). Averages were evaluated by the following weights: A=4, B=3, C=2, D=1, and F=0. Whenever number grades were given, the following conversions were made: A=95-100, B=88-94, C=80-87, D=70-79, F=69 or less.

The fifth variable used as a predictor was sex. A score of "1" was assigned to all males in the sample and a score of "0" was assigned to all females.


## Procedure

Data were analyzed by computing means, standard deviations, and correlation coefficients for both samples.

The smaller sample was used for cross-validating the regression results obtained from the larger sample. A thorough cross-validation would have entailed predicting criterion scores for the cross-validation sample and then correlating the predicted scores and scores actually obtained on the criterion. This correlation coefficient would be an index of the accuracy with which the regression equation predicts scores on an independent sample. The square of the correlation between predicted and actual scores was estimated by $r^2_{pa}$ as shown below (5).

$$r^2_{pa} = \frac{[W'V]^2}{W'RW}$$

W  is the vector of beta weights from the original sample,
W' is the transpose of W,
V  is the vector of validities for the new sample,
R  is the intercorrelation matrix of the independent variables for the new sample.

29

Regression equations were calculated from the results of the larger sample whenever they were considered meaningful by the investigator. These equations were computed by a method that selects the independent variables in the order of their contributions to the predicted variance (4).

## Results

The means and standard deviations of all variables are listed in Table 2.

Table 2

Means and Standard Deviations of All Variables

| | Samples | | | | | |
|---|---|---|---|---|---|---|
| | I N=435 | | II N=111 | | Total N=546 | |
| | m | sd | m | sd | m | sd |
| 1. UG | 2.16 | .74 | 2.30 | .76 | 2.19 | .75 |
| 2. ENG | 71.60 | 21.00 | 72.00 | 22.70 | 71.70 | 21.40 |
| 3. TGTP | 346.00 | 81.90 | 344.00 | 85.50 | 346.00 | 83.20 |
| 4. ACE | 108.00 | 20.00 | 107.00 | 21.30 | 108.00 | 20.60 |
| 5. HSG | 2.73 | .65 | 2.83 | .69 | 2.75 | .66 |

The data in Table 2 indicate a great deal of similarity between means and standard deviations for the two samples.

The intercorrelation matrix for the large sample is presented in Table 3; that of the smaller sample is presented in Table 4.

The results based on the data of the larger sample were cross-validated with the smaller sample by computing $r^2_{pa}$. The square root of $r^2_{pa}$ was .73. A t-test of the significance of this coefficient yielded a t of 19.3 with 432 degrees of freedom. This is highly significant and indicates that the regression equations based on the larger sample hold up well under cross-validation.

The regression equation for all five independent variables on UG is

$$(UG) = .0005 \ (ENG) + .0018 \ (TGTP) + .0010 \ (ACE) + .5937 \ (HSG) - .23$$

(Add .0099 to total for all female students.)

Table 3

Intercorrelation Matrix for Larger Sample (N=435)

|         | 1 | 2 | 3 | 4 | 5 | 6 |
|---------|---|-----|-----|-----|-----|------|
| 1. UG   |   | .34 | .46 | .40 | .61 | -.06 |
| 2. ENG  |   |     | .71 | .57 | .36 | -.20 |
| 3. TGTP |   |     |     | .80 | .45 | .08  |
| 4. ACE  |   |     |     |     | .40 | .00  |
| 5. HSG  |   |     |     |     |     | -.28 |
| 6. SEX  |   |     |     |     |     |      |

Table 4

Intercorrelation Matrix for Smaller Sample (N=111)

|         | 1 | 2 | 3 | 4 | 5 | 6 |
|---------|---|-----|-----|-----|-----|------|
| 1. UG   |   | .34 | .59 | .57 | .66 | -.10 |
| 2. ENG  |   |     | .63 | .44 | .41 | -.12 |
| 3. TGTP |   |     |     | .80 | .41 | .18  |
| 4. ACE  |   |     |     |     | .45 | .09  |
| 5. HGS  |   |     |     |     |     | -.26 |
| 6. SEX  |   |     |     |     |     |      |

The multiple correlation of these five variables with UG is .65 which is of the order of the correlation between HSG and UG alone (.61). It appears that the additional variables provide little additional information. The relative contribution of each variable to the predicted variance is

| | |
|-------|-------|
| ENG   | .4%   |
| FSTGTP | 9.0%  |
| ACE   | 1.1%  |
| HSGPA | 31.9% |
| SEX   | .4%   |

It is apparent that, for this combination of independent variables, HSG contributes a great deal of information and TGTP adds a little information.

The multiple correlation of HSG and TGTP on UG was .64. The inclusion of the other three variables does not add much additional information. The F ratio for the difference between the multiple correlation using all variables and the multiple correlation of HSG and TGTP with UG is 3.19 which is significant at 3 and 429 degrees of freedom. Thus the difference between the two multiple correlations is large enough to claim that one or more of the three variables ENG, ACE, or sex does add accuracy to the prediction.

The variables HSG and TGTP provide a fairly efficient regression equation when taken alone. The regression of these variables on UG is given by

$$UG = .5713 \ (HSG) + .0021 \ (TGTP) - .14$$

The multiple correlation of these variables with UG is .64 and the standard error of prediction is .57.

The difference between this multiple correlation of .64 and the correlation of HSG with UG (.61) yields an F ratio of 27.50 which at 1 and 433 degrees of freedom is highly significant. However, predictions based on HSG or TGTP alone should provide a fairly reliable prediction when both scores are unavailable. The regression of HSG on UG is

$$UG = .691 \ (HSG) + .274.$$

The standard error of this equation is .54.

The regression of TGTP on UG is given by

$$UG = .0042 \ (TGTP) + .714.$$

The correlation between UG and TGTP is .46 and the standard error of prediction is .66.

These last two regression equations are plotted in Figures 1 and 2. The center line in both figures is the actual regression line. The upper and lower lines are plus and minus one standard error of prediction from the regression line, respectively.

The probability is about .67 that an actual UG will fall between the two outer lines for any given predictor score.
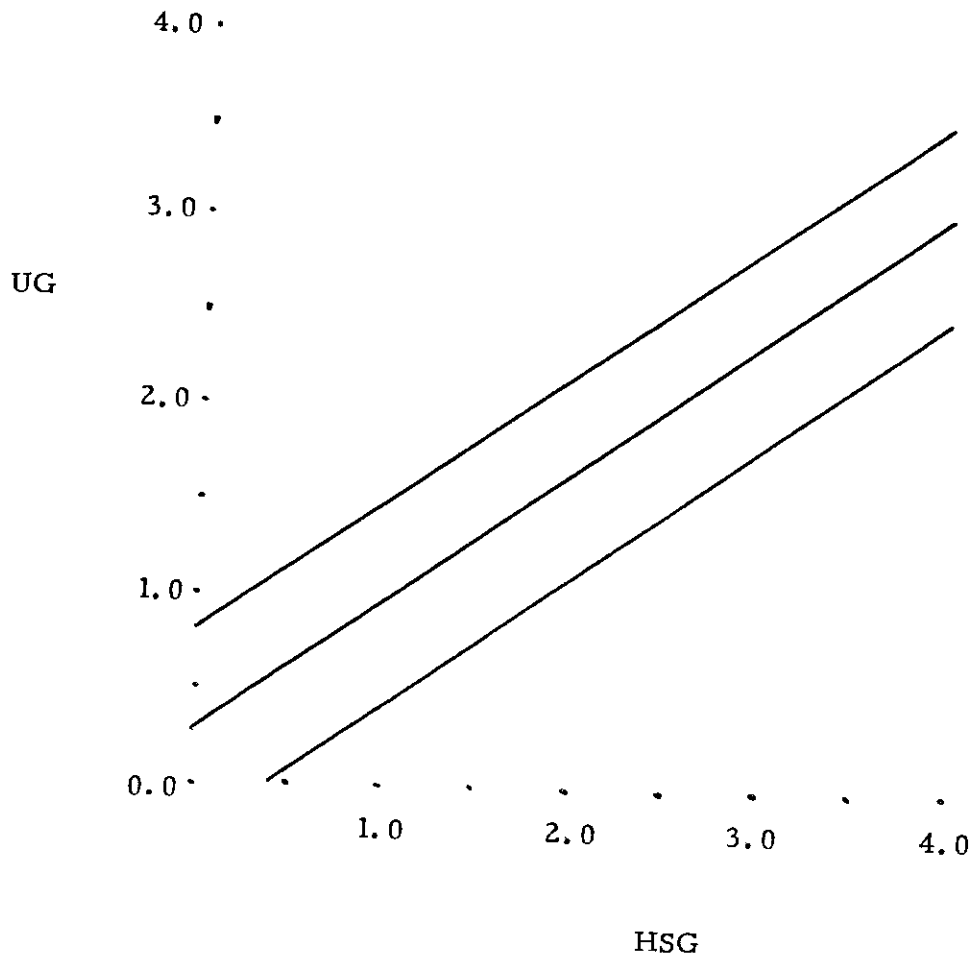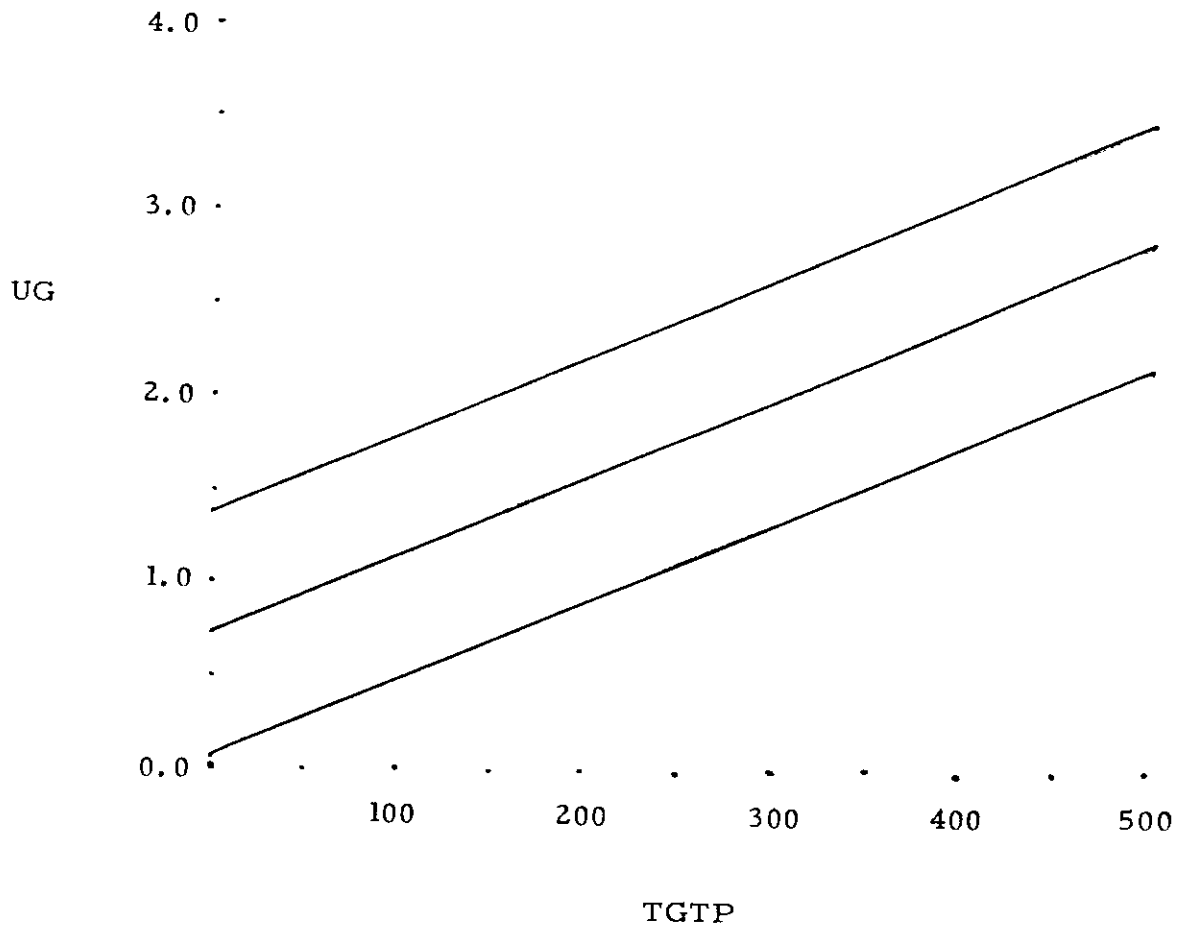
Fig. 1. The regression of HSG on UG.

Fig. 2. The regression of TGTP on UG.

# Discussion

It is apparent that curtailment of the TGTP distribution over the years in which it has been used as a selection criterion by the University has decreased its value as a predictor of achievement at FSU. This does not imply that the TGTP is a poor selection criterion, but rather indicates that its use in predicting actual averages has been weakened. Averages computed on academic high school units have consistently been the best predictors.

The English test score distribution has likewise been curtailed and it is a percentile distribution. Both of these conditions affect the correlation of this variable adversely, hence it appears to be a weak predictor. It was hoped that this variable might correlate with dependent variables on the same order as the total TGTP battery does. However, inspection of Tables 3 and 4 show that this is not the case.

The ACE scores were neither curtailed nor reported in percentiles. The distribution of raw scores was approximately normal. The correlations of the ACE with UG has been consistently between .40 and .50 over the past several years. However, since this variable correlates so highly with other independent variables, it adds little information to the prediction of UG when HSG and TGTP are used as independent variables.

The artificial binary variable sex did not add information to the prediction. Tables 2 and 3 show that this is due to low correlations with UG. It is interesting to note that the correlations of sex with HSG are considerably higher than those with UG. The high school sex difference is large enough to be of significant importance in predicting high school achievement. The difference favors the female student.

# Summary

This study reports an attempt to predict academic performance of Florida State University first-time Florida freshmen. Regression equations were computed for the regression of high school academic grade point average, Florida State-Wide Twelfth-Grade Testing Program, ACE, English, and sex on FSU grade point average. It was shown that HSG and TGTP together provide an efficient prediction of UG. Other variables appear to add little information. The best single predictor of UG is performance on academic grades in high school (r=.61). The multiple correlation of HSG and TGTP with UG was .64.

Cross-validation with an independent sample showed a high relationship between predicted and actual grades.

# References

1. Office of Educational Research and Service. Studies of educational progress among first-time Florida freshmen who entered the Florida State University in the fall semester, 1955-56. Florida State University, May, 1956.

2. Office of Educational Research and Service. A study of relationships between Florida State-Wide Twelfth-Grade Testing Program Scores and quantity and quality of academic work in the Florida State University, fall semester, 1956-57. Florida State University, May, 1957.

3. Office of Institutional Research and Service. A study of grades earned by first-time Florida freshmen in relation to Florida State-Wide Twelfth-Grade Testing Program scores, American Council on Education Psychological Examination scores, and high school records: fall semester, 1957-1958. Florida State University, June, 1958.

4. Rigney, Bob. Multiple Regression Program IBM 650. Lackland Air Force Base, Texas.

5. Ward, J. H. The application of a set of beta weights to new data. Personal communication, February 24, 1961.