DOCUMENT RESUME

ED 397 107 TM 025 237

AUTHOR Lyerla, Rob L.; Elmore, Patricia B.
TITLE Predicting Academic Success: An Application of Young's Universal Scale for Grades.
PUB DATE Apr 96
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Black Students; College Entrance Examinations; *College Freshmen; Ethnicity; *Grade Point Average; Grades (Scholastic); Higher Education; Intellectual Disciplines; Item Response Theory; Liberal Arts; Natural Sciences; Physical Sciences; *Prediction; Regression (Statistics); Sex Differences; Social Sciences; Test Results; *Undergraduate Students; White Students
IDENTIFIERS *Partial Credit Model; *Universal Scale for Grades

ABSTRACT

The prediction of academic success of undergraduate students using an item response theory partial credit model adjusted grade point average (IRTGPA) is presented and its results are compared to those from an unweighted grade point average (GPA). The sample included 2,444 freshmen admitted in the fall of 1987 to 681 courses at a large midwestern state university. For each student the individual courses, course grades, terms of enrollment, American college testing scores, ethnicity, and gender information were obtained. Analysis focused on three domains: social sciences, liberal arts, and natural and physical sciences. The calculation of an IRTGPA for freshmen and cumulative analyses enhanced the proportion of variance accounted for in the regression models for students in general, men and women as separate groups, and African American and White students as separate groups for the three academic domains. The differences found for students in general, men and women, and ethnic groups indicates that separation of courses into domain-specific groups and the calculation of an IRTGPA enhances prediction of academic success for underrepresented groups in particular. (Contains six tables and nine references.) (SLD)

# Predicting Academic Success:

# An Application of Young's Universal Scale for Grades

Rob L. Lyerla
Centers for Disease Control
Epidemiology Program Office
1600 Clifton Road, NE
Mailstop C08
Atlanta, GA 30333

Patricia B. Elmore
Department of Educational Psychology
and Special Education
Mailcode 4618
Southern Illinois University
at Carbondale
Carbondale, IL 62901-4618

Paper Presented at the Annual Meeting of the
American Educational Research Association
New York

April, 1996

# Predicting Academic Success:
# An Application of Young's Universal Scale for Grades

Prediction of academic success of college students has been a topic of interest for college administrators and faculty for many years (Elliot & Strenta, 1988; Fishman & Pasanella, 1969). Problematic areas have included not only the stability of the grade point average (GPA) across time, but also the comparability of the grade point average as a measure of academic success across disciplines. Even with these limitations many scholarships and awards as well as graduate school admission are determined using college GPA as the criterion.

This research furthers work begun by Young (1990, 1991a, 1991b) that utilized an item response theory adjusted GPA (IRTGPA) known as Young's Universal Scale for Grades. Young's scale is an application of item response theory, in particular the Partial Credit Model (Masters, 1982), in which course grade is considered a measure of ability within a specific unidimensional academic domain or discipline. Each course is analogous to an item and the subsequent course grade is the response to that item. Within this theoretical framework, Young (1990, 1991a, 1991b) predicted academic success measured by IRTGPA more accurately than GPA using Scholastic Achievement Test (SAT) scores as preadmission

measures for a cohort of students from a highly selective private university.

This research study extends the application of Young's (1990, 1991a, 1991b) Universal Scale for Grades using the IRTGPA with a more heterogeneous cohort of undergraduate students enrolled at a large midwestern state university that uses the American College Testing Program Examination (ACT) as the preadmission measure.

The purpose of this study was to compare the prediction of academic success of undergraduate students using two different measures of the dependent variable: (1) an item response theory partial credit model adjusted GPA (IRTGPA) and (2) an unweighted GPA (GPA).

## Method

### Participants

The sample included 2444 freshmen admitted in the fall of 1987 and 681 courses completed at a large midwestern state university.

### Variables

For each student the individual courses, course grades, term of enrollment for each course, ACT subtest and composite scores, ethnicity, and gender were obtained from university records. Grades such as pass/fail, incomplete, withdrew, and dropped were eliminated from all analyses.

Unidimensional academic domains used in this study were the three domains (social sciences, liberal arts, and natural and physical sciences) already established by Young (1990,

1991a, 1991b) and two additional domains (business and technical careers) appropriate for the institution.

Procedure

In order to establish a data matrix that consisted of students representative of the domain, only students who had completed five or more courses within a particular domain were included in the analyses. Additionally, the PARSCALE computer program used to compute IRTGPA required that all courses included in the analyses be represented by at least five percent of the subject pool.

In parameter estimation using the PARSCALE program, each subset of courses was scaled independently. The data matrix consisted of rows representing students and columns representing course grades. The cells of the matrix contained course grades coded A=5, B=4, C=3, D=2, F=1. Each unidimensional subset of courses was represented by a different data matrix. One ability estimate was determined for each student for each unidimensional set of courses. The estimate of each student's ability (theta), called the IRTGPA, was then used as the criterion for evidence of predictive validity.

A second criterion used for evidence of predictive validity was the unweighted GPA calculated as the mean grade received in the same courses used to calculate the IRTGPA. Traditional GPA is determined by incorporating a weight based on the number of credit hours of each course. To make the results of this study comparable to the results reported by

Young (1990), the unweighted GPA was calculated in this study following Young's procedure.

Statistical Analyses

Twenty multiple regression analyses were run, one for each of the five domains with freshman GPA, freshman IRTGPA, four-year GPA, and four-year IRTGPA as the criterion variables. The predictor variables in each equation were the four ACT subtests: English, Mathematics, Social Science, and Natural Science. The same regression analyses were also run separately for males and females and for four ethnic groups (African American, Asian, Hispanic, and White). Differences in the $R^2$ for the models using the GPA and IRTGPA were determined.

## Results

For most domains there was an increase in predictability measured by $R^2$ when the criterion variable was the IRT adjusted GPA (IRTGPA) rather than the unweighted GPA (GPA). The Pearson product-moment correlation coefficient between the IRT adjusted GPA and the unweighted GPA; the proportion of variance ($R^2$) explained by the four predictors (ACT English, Mathematics, Natural Science, and Social Science) using GPA and IRTGPA as criterion variables; the change in $R^2$ predicting IRTGPA as compared to GPA; and the number of courses and students by domain are shown in Table 1 for freshman analyses and in Table 2 for four-year cumulative analyses. The proportion of variance accounted for by each of the twenty full regression models was significantly

different than the proportion of variance accounted for by the null model at the .05 level of significance (see Tables 1 and 2 for p values). For freshman and cumulative IRTGPA compared to GPA the social science domain showed an increase in $R^2$ of .0581 and .0591, respectively. Similarly, freshman and cumulative increases for liberal arts were .0214 and .0221 and for business were .0379 and .0385, respectively. The largest freshman and cumulative increases of .1069 and .1049, respectively, were for the natural and physical sciences. For technical careers the freshman IRTGPA had an increase of .0629 but the cumulative IRTGPA had a decrease of .0923 compared to GPA.

Identical regression analyses were also run separately for men and women. The proportion of variance ($R^2$) explained by the four predictors (ACT English, Mathematics, Natural Science, and Social Science) using GPA and IRTGPA as criterion variables for men and women; the change in $R^2$ predicting IRTGPA as compared to GPA for men and women; and the number of courses, men, and women by domain are shown in Table 3 for freshman analyses and in Table 4 for four-year cumulative analyses. The proportion of variance accounted for by each of the twenty full regression models for both men and women was significantly different than the proportion of variance accounted for by the null model at the .05 level of significance except for the GPA and IRTGPA cumulative models for women in technical careers (see Tables 3 and 4 for p values). The IRTGPA was a better criterion variable for the

models for all domains for both freshman and cumulative analyses except cumulative technical careers for women in which there was a decrease of .0973 compared to GPA. It should be noted that only 28 women were included in the freshman analyses and 30 in the cumulative analyses for technical careers; therefore, the findings for women in technical careers may not be stable and generalizable. With the exception of the technical careers domain, the largest freshman increase for men (.1142) and women (.0676) was for the natural and physical sciences. Similarly, the largest cumulative increase was also found for the natural and physical sciences with .1101 for men and .0680 for women.

Regression analyses were also run separately for four ethnic groups (African American, Asian, Hispanic, and White). For the social science, liberal arts, and natural and physical science domains there were not sufficient numbers of Asian and Hispanic students to run the analyses; therefore, only results for African American and White students are presented. For the business and technical careers domains there were not sufficient numbers of African American, Asian, or Hispanic students to run the analyses; therefore, no results are reported for those two domains. The proportion of variance ($\underline{R}^2$) explained by the four predictors (ACT English, Mathematics, Natural Science, and Social Science) using GPA and IRTGPA as criterion variables for African American and White students; the change in $\underline{R}^2$ predicting IRTGPA as compared to GPA for African American and White

students; and the number of courses, White, and African American students by domain are shown in Table 5 for freshman analyses and in Table 6 for cumulative analyses. The proportion of variance accounted for by each of the twelve full regression models for both White and African American students was significantly different than the proportion of variance accounted for by the null model at the .05 level of significance (see Tables 5 and 6 for p values). The IRTGPA was a better criterion variable for the models for all domains for both freshman and cumulative analyses for both African American and White students. The largest freshman increases for White (.1054) and African American (.1894) students were in the natural and physical sciences. Similarly, the largest increases for cumulative analyses were also found in the natural and physical sciences with .1032 for White and .1797 for African American students.

## Discussion

The purpose of this study was to compare the prediction of academic success of undergraduate students using an item response theory partial credit model adjusted GPA (IRTGPA) compared to an unweighted GPA (GPA). It appears that the calculation of an IRT adjusted GPA for freshman and cumulative analyses enhances proportion of variance accounted for in the regression models for students in general, men and women as separate groups, and African American and White students as separate groups for the social science, liberal arts, and natural and physical science domains.

Researchers predicting academic success in college have reported $\underline{R}^2$ values ranging from .1178 (Wilson, 1981) to .3040 (Price & Kim, 1976) for all students combined. Farver, Sedlacek and Brooks (1975) studied ethnic and gender groups separately and found an $\underline{R}^2$ of .2514 for African American males but an $\underline{R}^2$ of .3721 for African American females, White males, and White females.

Young's (1990) original study reported larger $\underline{R}^2$ values than found in this study for models predicting cumulative GPA (.2774 versus .1190 for natural science, .2794 versus .1268 for social science, and .2672 versus .1740 for humanities/liberal arts) and for cumulative IRTGPA (.3729 versus .2239 for natural science, .2966 versus .1859 for social science, and .2617 versus .1961 for humanities/liberal arts). The larger $\underline{R}^2$ values reported by Young (1990) may be due to an additional predictor variable, high school grade point average, used in his regression analyses. The increases in $\underline{R}^2$ when using the IRTGPA were similar in Young's (1990) study and this study (.0955 versus .1049 for natural science, .0172 versus .0591 for social science, and -.0055 versus .0221 for humanities/liberal arts).

When comparing the same models for men and women separately, the technical careers domain should probably be excluded due to the small sample size for women. The $\underline{R}^2$ values for women are higher than for men especially in business (.3454 versus .1337 for freshman GPA, .3714 versus .1546 for freshman IRTGPA, .3501 versus .1368 for cumulative

GPA, and .3826 versus .1587 for cumulative IRTGPA). The largest freshman and cumulative increases in predictability using IRTGPA rather than GPA for men and women were in the natural and physical sciences.

When comparing African American and White students separately, the largest differences in the $\underline{R}^2$ values for the same models are in the natural and physical sciences with the $\underline{R}^2$ values larger for African American students than for White students (.1712 versus .1031 for freshman GPA, .3606 versus .2085 for freshman IRTGPA, .1706 versus .0959 for cumulative GPA, and .3503 versus .1991 for cumulative IRTGPA). The largest freshman and cumulative increases in predictability using IRTGPA rather than GPA for African American and White students were again in the natural and physical sciences.

The implications of this research are important for the use of GPA in selection situations such as for scholarships, awards, and graduate school admission. This study found differences for all students in general, men and women, and African American and White students. It appears that the separation of courses into domain specific groups and the calculation of an IRT adjusted GPA enhanced the prediction of academic success for all students and especially for under-represented groups.

References

Elliot, R., & Strenta, A. C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. Journal of Educational Measurement, 25, 333-347.

Farver, A. S., Sedlacek, W. E., & Brooks, G. C. (1975). Longitudinal predictions of university grades for blacks and whites. Measurement and Evaluation in Guidance, 7, 243-250.

Fishman, J. A., & Pasanella, A. K. (1969). College admissions-selection studies. Review of Educational Research, 30, 298-310.

Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47(2), 149-174.

Price, F. W., & Kim, S. (1976). The association of college performance with high school grades and college entrance test scores. Educational and Psychological Measurement, 36, 965-970.

Wilson, K. M. (1981). Analyzing the long-term performance of minority and nonminority students: A tale of two studies. Research in Higher Education, 15, 351-375.

Young, J. W. (1990). Adjusting the cumulative GPA using item response theory. Journal of Educational Measurement, 27, 175-186.

Young, J. W. (1991a). Gender bias in predicting college academic performance: A new approach using item response theory. Journal of Educational Measurement, 28, 37-47.

Young, J. W. (1991b). Improving the prediction of college performance of ethnic minorities using the IRT-based GPA. Applied Measurement in Education, 4, 229-239.

Table 1

$R^2$ Values for Models Predicting Freshman Unadjusted GPA and IRT Adjusted GPA for Five Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Number of Students | Criterion | r between IRTGPA and GPA | $R^2$ | p | $R^2$ Change |
|---|---|---|---|---|---|---|---|
| Social Science | 45 | 884 | GPA | .933 | .1365[a] | .0001 | .0581 |
| | | | IRTGPA | | .1946[a] | .0001 | |
| Liberal Arts | 27 | 1702 | GPA | .929 | .1769[a] | .0001 | .0214 |
| | | | IRTGPA | | .1983[a] | .0001 | |
| Natural/ Physical Sciences | 54 | 844 | GPA | .897 | .1235[a] | .0001 | .1069 |
| | | | IRTGPA | | .2304[a] | .0001 | |
| Technical Careers | 9 | 181 | GPA | .931 | .1912[a] | .0001 | .0629 |
| | | | IRTGPA | | .2541[a] | .0001 | |
| Business | 29 | 283 | GPA | .955 | .1818[a] | .0001 | .0379 |
| | | | IRTGPA | | .2197[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.

Table 2

$R^2$ Values for Models Predicting Cumulative Unadjusted GPA and IRT Adjusted GPA for Five Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Number of Students | Criterion | r between IRTGPA and GPA | $R^2$ | p | $R^2$ Change |
|---|---|---|---|---|---|---|---|
| Social Science | 44 | 938 | GPA | .927 | .1268[a] | .0001 | .0591 |
| | | | IRTGPA | | .1859[a] | .0001 | |
| Liberal Arts | 27 | 1793 | GPA | .929 | .1740[a] | .0001 | .0221 |
| | | | IRTGPA | | .1961[a] | .0001 | |
| Natural/ Physical Sciences | 55 | 873 | GPA | .897 | .1190[a] | .0001 | .1049 |
| | | | IRTGPA | | .2239[a] | .0001 | |
| Technical Careers | 8 | 189 | GPA | .700 | .1900[a] | .0001 | -.0923 |
| | | | IRTGPA | | .0977[a] | .0008 | |
| Business | 29 | 291 | GPA | .955 | .1873[a] | .0001 | .0385 |
| | | | IRTGPA | | .2258[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.

Table 3

$R^2$ Values for Models Predicting Freshman Unadjusted GPA and IRT Adjusted GPA Separately for Men and Women for Five Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Criterion | Number of Men | $R^2$ Men | p Men | $R^2$ Change Men | Number of Women | $R^2$ Women | p Women | $R^2$ Change Women |
|---|---|---|---|---|---|---|---|---|---|---|
| Social | 45 | GPA | 403 | .1398[a] | .0001 | .0648 | 481 | .1681[a] | .0001 | .0422 |
| Science | | IRTGPA | | .2046[a] | .0001 | | | .2103[a] | .0001 | |
| Liberal | 27 | GPA | 941 | .1817[a] | .0001 | .0229 | 761 | .2193[a] | .0001 | .0195 |
| Arts | | IRTGPA | | .2046[a] | .0001 | | | .2388[a] | .0001 | |
| Natural/ | 54 | GPA | 529 | .1066[a] | .0001 | .1142 | 315 | .2046[a] | .0001 | .0676 |
| Physical | | IRTPGA | | .2208[a] | .0001 | | | .2722[a] | .0001 | |
| Sciences | | | | | | | | | | |
| Technical | 9 | GPA | 153 | .2498[a] | .0001 | .0486 | 28 | .1688 | .3508 | .0703 |
| Careers | | IRTGPA | | .2984[a] | .0001 | | | .2391 | .1619 | |
| Business | 29 | GPA | 166 | .1337[a] | .0001 | .0209 | 117 | .3454[a] | .0001 | .0260 |
| | | IRTGPA | | .1546[a] | .0001 | | | .3714[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.

Table 4

$R^2$ Values for Models Predicting Cumulative Unadjusted GPA and IRT Adjusted GPA Separately for Men and Women for Five Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Criterion | Number of Men | $R^2$ Men | p Men | $R^2$ Change Men | Number of Women | $R^2$ Women | p Women | $R^2$ Change Women |
|---|---|---|---|---|---|---|---|---|---|---|
| Social | 44 | GPA | 434 | .1212[a] | .0001 | .0609 | 504 | .1636[a] | .0001 | .0493 |
| Science | | IRTGPA | | .1821[a] | .0001 | | | .2129[a] | .0001 | |
| Liberal | 27 | GPA | 994 | .1727[a] | .0001 | .0215 | 799 | .2150[a] | .0001 | .0218 |
| Arts | | IRTGPA | | .1942[a] | .0001 | | | .2368[a] | .0001 | |
| Natural/ | 55 | GPA | 553 | .0996[a] | .0001 | .1101 | 320 | .2073[a] | .0001 | .0680 |
| Physical | | IRTGPA | | .2097[a] | .0001 | | | .2753[a] | .0001 | |
| Sciences | | | | | | | | | | |
| Technical | 8 | GPA | 159 | .2508[a] | .0001 | .0099 | 30 | .1117 | .5452 | -.0973 |
| Careers | | IRTGPA | | .2607[a] | .0001 | | | .0144 | .9844 | |
| Business | 29 | GPA | 171 | .1368[a] | .0001 | .0219 | 120 | .3501[a] | .0001 | .0325 |
| | | IRTGPA | | .1587[a] | .0001 | | | .3826[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.

Table 5

$R^2$ Values for Models Predicting Freshman Unadjusted GPA and IRT Adjusted GPA Separately for White and African American Students for Three Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Criterion | Number of White | $R^2$ White | p White | $R^2$ Change White | Number of African American | $R^2$ African American | p African American | $R^2$ Change African American |
|---|---|---|---|---|---|---|---|---|---|---|
| Social Science | 45 | GPA | 711 | .1074[a] | .0001 | .0501 | 146 | .0879[a] | .0110 | .0447 |
| | | IRTGPA | | .1575[a] | .0001 | | | .1326[a] | .0005 | |
| Liberal Arts | 27 | GPA | 1435 | .1516[a] | .0001 | .0259 | 212 | .1416[a] | .0001 | .0066 |
| | | IRTGPA | | .1775[a] | .0001 | | | .1482[a] | .0001 | |
| Natural/ Physical Sciences | 54 | GPA | 719 | .1031[a] | .0001 | .1054 | 89 | .1712[a] | .0031 | .1894 |
| | | IRTPGA | | .2085[a] | .0001 | | | .3606[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.

Table 6

$R^2$ Values for Models Predicting Cumulative Unadjusted GPA and IRT Adjusted GPA Separately for White and African American Students for Three Domains Using ACT Scores as Predictors

| Domain | Number of Courses | Criterion | Number of White | $R^2$ White | p White | $R^2$ Change White | Number of African American | $R^2$ African American | p African American | $R^2$ Change African American |
|---|---|---|---|---|---|---|---|---|---|---|
| Social Science | 44 | GPA | 750 | .0981[a] | .0001 | .0503 | 160 | .0851[a] | .0077 | .0460 |
| | | IRTGPA | | .1484[a] | .0001 | | | .1311[a] | .0002 | |
| Liberal Arts | 27 | GPA | 1503 | .1501[a] | .0001 | .0256 | 229 | .1256[a] | .0001 | .0062 |
| | | IRTGPA | | .1757[a] | .0001 | | | .1318[a] | .0001 | |
| Natural/ Physical Sciences | 55 | GPA | 742 | .0959[a] | .0001 | .1032 | 91 | .1706[a] | .0027 | .1797 |
| | | IRTGPA | | .1991[a] | .0001 | | | .3503[a] | .0001 | |

[a] $R^2$ for the model is statistically significant at .05.