

ETHNIC AND SEX DIFFERENCES IN THE PREDICTIVE
VALIDITY OF THE SCHOLASTIC ACHIEVEMENT
TEST FOR COLLEGE GRADES¹

RICHARD LYNN

AND

WEI-CHENG MAU

Ulster Institute for Social Research

Wichita State University

Summary.—The predictive validity of SAT-Verbal and SAT-Mathematics scores for college grades was examined for males and females and for four ethnic groups in a nationally representative American sample. SAT scores “overpredicted” the grades of males and of Asians, Hispanics, and blacks, i.e., these groups did not obtain as good grades as would be predicted from their SAT scores. Conversely, SAT scores “underpredicted” the grades obtained by females and whites, i.e., these groups obtained better grades than would be predicted from their SAT scores. The possible explanations for these differential predictive validities are discussed.

A number of studies have been carried out on how well cognitive tests used to select applicants for admission to colleges and universities predict the subsequent academic achievements of the students who have been admitted. In the United States these studies have generally examined how well students' performance on the SAT (Scholastic Achievement Test), taken as part of the admissions procedure, is related to the grades measured as GPAs (Grade Point Averages) during their years at college. These studies have typically shown that SAT scores have some positive correlations with grades but that these correlations are quite low and have generally been in the range of .2 to .4 (e.g., Linn, 1990). In more recent studies, there has been interest in differential prediction for males and females and for different ethnic groups. Differential prediction means that SAT scores predict the subsequent academic attainment of some groups better than others. In this research, a distinction is made between “overprediction” and “underprediction.” Overprediction means that the group performs less well on grades than would be predicted from its SAT scores. Overprediction has been found among men as compared with women and among blacks as compared with whites in samples analysed by Linn (1990) and Young (1994). Underprediction means that the group performs better than would be predicted from SAT scores and has been found in these studies among women as compared with men and among whites as compared with blacks. The inference drawn from these studies is that women and whites have some advantages in obtaining college grades which is not measured by their SAT scores and, conversely, that men

¹Address correspondence to R. Lynn, Ph.D., Ulster Institute for Social Research, University of Ulster, Coleraine, Northern Ireland BT52 1SA.

and blacks have some corresponding disadvantages. In this paper we report a further study of these phenomenon.

METHOD

Participants

Data for this study were obtained from respondents in the Baccalaureate and Beyond 1993-1994 survey (B&B: 93/94; National Center for Educational Statistics, 1996) sponsored by the U.S. Department of Education. The participants ($N=10,080$) were a subsample of the students selected for the 1993 National Postsecondary Student Aid Study (NPSAS: 93) sample, a nationally representative sample of all postsecondary students. The B&B: 93/94 sample represents all postsecondary students in the United States who received a baccalaureate degree between July 1, 1992 and June 20, 1993.

The sampling procedure for the NPSAS: 93 survey was carried out by utilizing a stratified, multistage sample design with postsecondary institutions as the first stage unit and students within the institutions as the second stage unit. The institution sample was stratified by type of control (private vs public), highest degree offered, size of enrollment in professional programs, graduate student enrollment, and the number of bachelor degrees awarded in education. Of the 12,731 B&B: 93/94 eligible students from NPSAS: 93, 92% or 10,080 were successfully contacted and participated in the first B&B: 93/94 follow-up study. The final sample consisted of 43.4% males ($n=4,377$) and 56.6% females ($n=5,703$). Eighty-two percent of the sample categorized themselves as white, 6.0% as African American, 5.0% as Hispanic, 0.7% as American Indian, 4.0% as Asian/Pacific Islander, and 2.3% as 'Other' or 'Missing'.

Data Analyses

Of the 10,080 B&B: 93/94 participants, 3,512 for whom there is complete information on both SAT and GPA were used for the analysis. Differences in the SAT-Verbal and SAT-Mathematics and in the GPA among sex and ethnic groups were analyzed using multivariate analysis of variance. To examine the differential prediction of college grades by sex and by ethnicity, separate multiple regressions for each sex and ethnic group were conducted with SAT verbal score and SAT mathematics score as predictors and college GPA as the dependent variable. Mean residuals for each sex and ethnic group were also calculated based on a single prediction equation derived from the total sample. The analyses were based on weighted samples created to adjust for the oversampling bias, which redistributed the observations to represent the distribution in the population. Design effects were also used to adjust standard error for hypothesis testing (National Center for Educational Statistics, 1997).

RESULTS

Descriptive statistics for the sample are presented in Table 1. This shows for males and females and for the four ethnic groups the means and standard deviations obtained on the GPA, the SAT-Verbal, and the SAT-Mathematics. With regard to sex, it will be seen that males obtained higher scores on the SAT-Verbal and the SAT-Mathematics while females obtained higher GPAs. A multivariate analysis of variance indicated that there are statistically significant differences on all three measures as a function of sex ($F=106.34$, $p<.001$). The analyses also indicated that there are statistically significant differences on all three measures as a function of race ($F=36.27$, $p<.03$). Subsequent univariate analyses indicated the presence of significant differences in the SAT-Verbal ($F=52.84$, $p<.001$), the SAT-Mathematics ($F=73.34$, $p<.001$), and the GPA ($F=27.35$, $p<.001$).

TABLE 1
MEANS AND STANDARD DEVIATIONS BY SEX AND RACE

| | Male ($n=1,605$) | Female ($n=1,907$) | Asian ($n=173$) | Black ($n=183$) | Hispanic ($n=164$) | White ($n=2,919$) | Significant Differences* |
|-----------------|-----------------------|-------------------------|----------------------|----------------------|-------------------------|------------------------|-----------------------------|
| GPA | | | | | | | |
| M | 2.96 | 3.08 | 2.97 | 2.76 | 2.93 | 3.05 | B < A, H, W |
| SD | .50 | .50 | .54 | .43 | .47 | .50 | F < M |
| SAT-Verbal | | | | | | | |
| M | 491.95 | 477.40 | 485.05 | 404.07 | 462.87 | 489.26 | B < A, H, W |
| SD | 99.33 | 104.11 | 134.36 | 109.81 | 107.79 | 96.53 | H < A, W |
| SAT-Mathematics | | | | | | | M > F |
| M | 560.05 | 515.54 | 590.50 | 447.38 | 505.29 | 539.44 | B < H < W < A |
| SD | 107.27 | 104.15 | 110.70 | 110.71 | 115.95 | 103.55 | M < F |

* $p<.05$.

The statistical significance of differences among the four ethnic groups was tested by Scheffé *post hoc* analyses. These showed that the black group scored significantly lower than other groups on GPAs. On the SAT-Verbal, black students scored significantly lower than Asian, Hispanic, and white students, whereas Hispanic students scored significantly lower than Asian and white students. On the SAT-Mathematics, Asian students scored significantly higher than white students, whereas white students scored significantly higher than Hispanic students, and Hispanic students scored significantly higher than black students.

The correlations of the SAT-Verbal and the SAT-Mathematics with college GPA are shown in Table 2. All correlations are statistically significant ($p<.001$), but they are nevertheless quite low, suggesting the SAT scores are relatively weak predictors of grade point average. It should be noted that the students attended a number of different colleges that will have had different

TABLE 2
ZERO-ORDER CORRELATIONS WITH GPA

| | All | Male | Female | Asian | Black | Hispanic | White |
|-----------------|-----|------|--------|-------|-------|----------|-------|
| SAT-Verbal | .30 | .28 | .34 | .24 | .32 | .20 | .29 |
| SAT-Mathematics | .27 | .29 | .31 | .27 | .42 | .19 | .25 |

Note.—All coefficients are significant at $p < .001$.

grading standards. This will have reduced the reliability of the GPAs and hence the magnitude of the correlations.

The magnitude of the effect of the SAT scores on grade point average was tested further by calculating prediction equations for GPAs for the two sexes and for the four ethnic groups. The results are shown in Table 3. These results confirm that the predictive power of SAT scores for GPA is quite low. Nevertheless, the R^2 for each group are statistically significant at $p < .001$ for white, black, and Asian groups and at $p < .01$ for the Hispanic group. The prediction equation obtained for the entire sample is

$$\text{GPA} = 2.18 + .0011 \text{ SAT-Verbal} + .0006 \text{ SAT-Mathematics}$$

TABLE 3
PREDICTION EQUATIONS BY SEX AND RACE

| | Sex | | Race | | | | |
|-----------------|--------|--------|--------|--------|----------|--------|--------|
| | Male | Female | Asian | Black | Hispanic | White | All |
| Intercept | 2.28† | 2.15† | 2.15† | 2.01† | 2.47† | 2.20† | 2.18† |
| SAT-Verbal | .0008† | .0011† | .0005 | .0001 | .0006 | .0011† | .0011† |
| SAT-Mathematics | .0009† | .0008† | .0010* | .0016† | .0004 | .0005† | .0006† |
| R^2 | .10† | .13† | .08† | .18† | .04* | .09† | .10† |

* $p < .05$. † $p < .001$.

Table 4 shows the residual values calculated from the prediction equation for all students. Residual values greater than zero suggest an underprediction of GPA, whereas residual values smaller than zero suggest an over-

TABLE 4
MEANS AND STANDARD DEVIATIONS OF THE RESIDUAL VALUES BASED ON TOTAL SAMPLE

| | <i>M</i> | <i>SD</i> |
|-----------|----------|-----------|
| Sex | | |
| Male | -.0901 | .49 |
| Female | .0826 | .47 |
| Ethnicity | | |
| Asian | -.0709 | .56 |
| Black | -.1469 | .40 |
| Hispanic | -.0347 | .46 |
| White | .0234 | .48 |

prediction of GPA. The results show that college GPA was overpredicted for male students and underpredicted for female students. With regard to the ethnic differences, the results show that there was an underprediction for white students and an overprediction for the three minority groups.

DISCUSSION

The results contain four principal points of interest. First, with regard to the descriptive statistics set out in Table 1, the results show that males obtained higher means on both the SAT-Verbal and the SAT-Mathematics. The male advantage is greater on the latter, where it amounts to approximately .42 of a standard deviation (d), than on the SAT-Verbal, where it amounts to approximately .14 of a standard deviation (d). Hence, measured in ds (standard deviation units), the male advantage on the SAT-Mathematics is about three times greater than on the SAT-Verbal. These results confirm previous findings. Thus, in the early 1980s males obtained a higher mean on the SAT-Verbal by .11 d (Ramist & Arbeiter, 1986) and on the SAT-Mathematics by .40 d (Arbeiter, 1985). These differences are very close to those obtained in the present study. However, despite the male advantage on both the SAT-Verbal and the SAT-Mathematics, females obtained significantly better GPAs than males. It appears that females have some advantages for GPAs that are not measured by the SAT-Verbal or the SAT-Mathematics. These advantages are probably both cognitive and motivational. Females have cognitive advantages in essay writing skills and spelling (Hyde & Linn, 1988) which are not measured in the SAT-Verbal but which are likely to contribute to GPAs. In addition, females appear to have stronger work motivation than males (Farmer, 1983; Nevill & Perotta, 1985; Nevill & Super, 1988; Watson & Stead, 1990; Luzzo, 1994), and this also is likely to contribute to their obtaining higher GPAs.

Second, with regard to the descriptive statistics on ethnic differences, the results show that on the SAT-Verbal blacks obtained significantly lower scores than the other three groups. This confirms the results obtained by Young (1994). On the SAT-Mathematics, the highest scores are obtained by Asians, followed in descending order by the white, Hispanic, and black groups. There is a discrepancy here with the results obtained by Young, who found that the white group obtained the highest score on the SAT-Mathematics, followed by the Asian, Hispanic, and black groups. This discrepancy is probably due to the different samples. Young's sample was drawn from one university, whereas the present sample was drawn nationwide. Nevertheless, apart from the reversal of the scores for white or Asian groups, the results obtained in the two studies are similar.

Third, the results set out in Tables 2 and 3 show that, although the SAT scores have significant positive correlations with GPAs, these correla-

tions and the predictive value of SAT scores for GPAs are quite modest. Evidently there are important determinants of GPAs other than the cognitive abilities measured by SAT scores.

Fourth, the analyses presented in Table 4 were carried out to examine the differential predictive validity of the SAT scores for males and females and for the four ethnic groups. With regard to the sex differences, the results confirm that the SAT scores significantly underpredict the GPAs of females and overpredict the GPAs of males; that is to say, females obtain better GPAs than is predicted from their SAT scores, while males obtain lower GPAs. With regard to the ethnic differences, SAT scores underpredict the GPAs of the white group (who obtain better GPAs than would be predicted from their SAT scores), while among the other three ethnic groups SAT scores overpredict GPAs (these do less well on GPAs than would be predicted from their SAT scores). The greatest overprediction is present among blacks. Closely similar results to these were obtained by Young (1994). It appears that white students must have some advantages for GPAs not measured by the SAT, while Asian, Hispanic, and particularly black students must have some disadvantages. The reasons for these differential predictions among the four ethnic groups are not understood and should be the focus for further research in this area.

REFERENCES

- ARBEITER, S. (1985) *Profiles, college bound seniors*. New York: College Entrance Examination Board.
- FARMER, H. S. (1983) Career and homemaking plans for high school youth. *Journal of Counseling Psychology*, 30, 40-45.
- HYDE, J. S., & LINN, M. C. (1988) Gender differences in verbal ability: a meta-analysis. *Psychological Bulletin*, 104, 53-69.
- LINN, R. L. (1990) Admissions testing: recommended uses, validity, differential prediction and coaching. *Applied Measurement in Education*, 3, 297-318.
- LUZZO, D. A. (1994) An analysis of gender and ethnic differences in college students' commitment to work. *Journal of Employment Counseling*, 31, 38-45.
- NATIONAL CENTER FOR EDUCATIONAL STATISTICS. (1996) *Baccalaureate and beyond longitudinal study: 1993/94 first follow-up methodology report*. (Report No. NCES 96-149) Washington, DC: U.S. Department of Education.
- NATIONAL CENTER FOR EDUCATIONAL STATISTICS. (1997) *A note from the chief statistician: technical approaches to performing regression and other multivariable techniques on NCES survey data*. (Notes No. 3) Washington, DC: U.S. Department of Education.
- NEVILL, D. S., & PEROTTA, J. M. (1985) Adolescent perceptions of home and work in Australia, Portugal and the USA. *Journal of Cross-Cultural Psychology*, 16, 483-495.
- NEVILL, D. S., & SUPER, D. E. (1988) Career maturity and commitment to work in university students. *Journal of Vocational Behavior*, 32, 139-151.
- RAMIST, L., & ARBEITER, S. (1986) *Profiles, college bound seniors*. New York: College Entrance Examination Board.
- WATSON, M. B., & STEAD, G. B. (1990) Work role salience of South African adolescents. *Journal of Vocational Behavior*, 36, 249-257.
- YOUNG, J. W. (1994) Differential prediction of college grades by gender and ethnicity: a replication study. *Educational and Psychological Measurement*, 14, 1022-1029.

Accepted July 10, 2001.