SEX DIFFERENCES IN PREDICTABILITY OF SOME MEASURES OF COLLEGE ACHIEVEMENT FROM THE COLLEGE ENTRANCE TEST AND/OR SOME MEASURES OF HIGH SCHOOL ACHIEVEMENT

REMIGIO S. GUMBAN
Guidance and Testing Division
Fund for Assistance to Private Education

Tests of significance of the hypotheses that the predictabilities of college academic criteria (viz., subject-area averages in English, Mathematics, and Social Science, and; an overall measure, defined as average of the subject-area measures) are equal for male and female "populations" of college freshmen are presented.

The findings are as follows: (1) Non-significant sex differences in predictability of the college overall measure, and subject-area average in Social Science when predicted from (a) high school subject-area averages, (b) CET subtest scores, (c) combination of the high school and CET variables, and (d) overall measures in high school (i.e., average of the subject-area averages) and the CET; (2) Significant sex differences in predictabilities of subject-area average in Mathematics in all four prediction schemes above; (3) Significant sex differences in predictability of subject-area average in English in three prediction schemes, viz., (b)–(d) above.

Discussions on (1) the "fidelity" of single measures of initial college academic performance, and (2) the implications of the results for admissions purposes are also presented.

In a validity study on the College Entrance Test (CET) Gumban and Iledan (1971) compared the predictabilities of some measures of academic achievement in college from four sets of predictors for each "population"[1] of male and female college freshmen obtained from a relatively homogeneous group of schools. The criterion variables were first semester subject-area averages in English, Mathematics, and Social Science, and an approximate overall measure of initial college academic achievement (defined as average of the subject-area measures). The sets of predictors, from which optimal predictors for each criterion were selected using a stepwise multiple linear regression procedure, were as follows: (1) high school subject-area averages in English, Pilipino, Mathematics, Natural Science, and Social Science; (2) CET subtest scores; (3) combination of high school subject-area averages, and CET subtest scores, and; (4) overall measure of high school academic achievement (defined as average of the high school subject-area measures), and CET overall score.

The findings showed evidences that sex is a dimension in differentiating the population of college freshmen for more refined results in academic prediction.

The present paper gives a formal test of the effect of sex in academic prediction based on Gumban and Iledan's data: Tests of significance are presented of the hypotheses that the predictabilities of the academic criteria are equal for male and female "populations." Furthermore, a discussion on the implications of the findings for academic admissions policies is presented.

METHOD

Statistical Treatment

Within a population of male or female students the predictability of an academic criterion is appropriately measured by the multiple correlation coefficient (note that the ordinary correlation coefficient is a special case) for within-population comparative analysis. But if

[1]The term population is to be understood with reference to the cluster of schools where the results of the study hold. Whenever it is intended to convey this specific meaning it is written under quotes for clarity.

our interest is to compare the predictability of such a criterion between male and female populations of students, then such an index of predictability is not sufficient. Abelson (1952) indicated that sex difference in academic prediction must be tested using the standard error of estimate as a measure of predictability rather than the multiple correlation coefficient because the female criterion-variances are generally smaller. More recent reports in the psychological literature still give implicit evidences (i.e., indicated in the data reported but not discussed) to support this sex difference in variability of academic criteria (Holland and Richards, 1966; Richards and Lutz, 1967). The standard error of estimate, given by the relation $S_{\hat{c}} = S_c\sqrt{1-R^2}$ clearly embodies the combined effects of the multiple correlation coefficient, R, and the variance, S_c^2, of the criterion. It can also be deduced "prima facie" that R is an appropriate measure of predictability for within-population comparative analysis, S_c being constant.

We can test the simple hypothesis that the true standard errors of prediction are equal for two populations using the statistic,

$$G = (N_m + N_f)\log_e\left[\frac{N_m(S_{\hat{c}}^2)_m + N_f(S_{\hat{c}}^2)_f}{(N_m + N_f)}\right] - N_m\log_e(S_{\hat{c}}^2)_m - N_f\log_e(S_{\hat{c}}^2)_f$$

where N_m, N_f denote the sample sizes, and $(S_{\hat{c}})_m$, $(S_{\hat{c}})_f$ denote sample standard errors of estimate. G as given above is a simplification of the expression, $G = 2\log_e \lambda$ where λ is a Neyman-Pearson likelihood ratio for testing the hypothesis. For large sample sizes N_m, N_f, G is distributed as chi-square with 1 degree of freedom if the hypothesis is true (for more details see Gulliksen and Wilks, 1950).[2]

If G is significant (insignificant) at an assumed level of 0.05 we consider this as evidence that the hypothesis must be rejected (accepted), i.e., the population standard errors of estimate are unequal (equal).

Data

The data obtained from the study of Gumban and Iledan are presented in Table 1. Clearly, in three out of four cases the sample criterion-variances of the initial college academic performance measures (viz., college subject-area averages in English, Mathematics, and Social Science, and; the approximate overall measure) are smaller for females.

RESULTS AND DISCUSSION

The results of the statistical tests are summarized in Table 2.

[2]In this study G is approximately distributed only as chi-square.

Non-Significant Sex Differences in Predictability of Overall College Academic Performance

When the overall measure of academic achievement (COL-AVE) is predicted from high school achievement measures, {HSAA}, *or* CET subtest scores, {CET}, *or* their two combinations, {HSAA} U {CET} and {HS-AVE, CET-OVL}, G is invariably insignificant. Thus, the predictability of this criterion is not affected by sex as a sampling dimension in all prediction schemes considered, i.e., in this particular case the male and female samples can be lumped together.

Significant Sex Differences in Predictability of Subject-Area Measures in English and Mathematics

In all prediction schemes considered, the subject-area measure in Mathematics (COLSAA-M) is predicted with greater precision for the female "population" of college freshmen.

Except in one case, i.e., when the prediction scheme involves the set of high school subject-area measures, {HSAA}, as regressor set, the subject-area measure in English (COLSAA-E) is also predicted with precision favoring the female "population" of college freshmen.

It must be emphasized that in cases where sex difference are insignificant, the regression surfaces of the male and female "populations" of college freshmen *may* still fail to be coincident: The regression surfaces may *not* be parallel and/or they may *not* have equal intercepts. In the study of Gumban and Iledan the latter seems to hold because the optimal predictors for a fixed criterion and a fixed prediction scheme are impressionistically different.

These issues are important especially in connection with the problem of "bias," and can be subjected to formal tests. However, such operations are beyond the scope of this study.

An Implication for Academic Admissions Policies

In formulating college admissions policies the lack of sex differences in predictability of the overall measure of initial college performance within the "population" of college freshmen is of great practical value.

TABLE 1

MULTIPLE CORRELATIONS (R*) AND STANDARD DEVIATIONS OF CRITERIA (S_c) FOR EACH PREDICTION SCHEME AND FOR EACH SAMPLE OF MALES (N = 68) AND FEMALES (N = 151)

		MALE		FEMALE	
PREDICTOR SET	CRITERION	R*	S_c	R*	S_c
(HSSAA)	COLSAA-E	0.506	5.18	0.199	3.82
	COLSAA-M	0.296	8.25	0.195	6.61
	COLSAA-SS	0.411	5.95	0.210	6.08
	COL-AVE	0.441	4.77	0.230	4.08
(CET)	COLSAA-E	0.423	5.18	0.370	3.82
	COLSAA-M	0.335	8.25	0.278	6.61
	COLSAA-SS	0.363	5.95	0.377	6.08
	COL-AVE	0.464	4.77	0.366	4.08
(HSSAA) U (CET)	COLSAA-E	0.556	5.18	0.376	3.82
	COLSAA-M	0.355	8.25	0.325	6.61
	COLSAA-SS	0.492	5.95	0.417	6.08
	COL-AVE	0.552	4.77	0.391	4.08
(HS-AVE, CET-OVL	COLSAA-E	0.456	5.18	0.361	3.82
	COLSAA-M	0.290	8.25	0.218	6.61
	COLSAA-SS	0.388	5.95	0.267	6.08
	COL-AVE	0.497	4.77	0.369	4.08

*adjusted for df

REMARK: The codes of the predictor sets and criteria are interpreted as follows:

(HSSAA) – set of High School Subject-Area Averages
(CET) – set of CET subtest scores
(HSSAA) U (CET) – union of the sets of high school and CET predictors
COLSAA-E – College Subject-Area Average in English
COLSAA-M – College Subject-Area Average in Mathematics
COLSAA-SS – College Subject-Area Average in Social Science
COL-AVE – College Average Computed as mean of subject-area averages
(HS-AVE, CET OVL) – set of High School Average (computed as mean of subject-area averages) and CET overall score

TABLE 2

STANDARD ERROR OF ESTIMATES FOR EACH SAMPLE OF MALES (N = 68) & FEMALES (N = 151) AND FOR EACH PREDICTION SCHEME, AND VALUES OF G STATISTIC

PREDICTOR SET	CRITERION	$(S_{\hat{c}})_m^{\dagger}$	$(S_{\hat{c}})_f^{\dagger}$	G
(HSSAA)	COLSAA-E	4.568	3.770	3.672
	COLSAA-M	8.002	6.524	4.094*
	COLSAA-SS	5.503	5.982	0.474
	COL-AVE	4.345	3.998	0.538
(CET)	COLSAA-E	4.874	3.583	9.577*
	COLSAA-M	7.887	6.391	4.290*
	COLSAA-SS	5.670	5.727	0.046
	COL-AVE	4.321	3.859	1.276
(HSSAA) U (CET)	COLSAA-E	4.475	3.586	4.984*
	COLSAA-M	7.887	6.312	4.931*
	COLSAA-SS	5.384	5.660	0.449
	COL-AVE	4.130	3.818	0.561
(HS-AVE, CET-OVL)	COLSAA-E	4.714	3.587	7.412*
	COLSAA-M	8.010	6.517	4.260*
	COLSAA-SS	5.611	5.916	0.019
	COL–AVE	4.236	3.831	0.937

*$p \leqslant 0.05$.

†adjusted for df

Let us consider a cluster of colleges where the regression surfaces of the previous study are valid. And let us suppose that there is at least one coeducational school belonging to this cluster. Suppose further that one such a school has been screening its college freshmen applicants on the basis of high school grades as academic criteria used in a manner that is equivalent[3] (at least on a practical level) to the first prediction scheme presented in the previous study (i.e., the prediction scheme where the regressors are subject-area averages in high school) and this school opts to include CET measures in its academic criteria to be used in a manner that is similarly equivalent to the second prediction scheme (i.e., the prediction scheme where the regressors are the combined high school subject-area averages and CET subtest scores) in order to have a more predictive set of academic criteria. Then, the desired improvement in predictive efficiency will be *effected* for both sexes. Naturally, for male and female exclusive schools belonging to the cluster the CET will also be equally useful under the same usage schemes.

Content Coverage of Overall College Achievement Measures

It must be emphasized that the overall measure of initial college performance considered, defined earlier as average of the college subject-area measures (i.e., a linear combination of the three subject-area measures with all three coefficients equal to 1/3) is generally wider in span[4] than the usual grade-point average[5] (GPA), defined as weighted average of subject grades (i.e., a linear combination of the subject marks with the ith coefficient equal to $u_i / \sum_i^k u_i$ where u_i is the number of credit units for the ith subject, and k is the number of subjects taken for which a student was given a grade) in the college academic achievement space because it receives contributions from three areas, viz., English, Mathematics, and Social Science areas; whereas, the GPA is constituted from generally varying components from one student to another, tapping in most cases fewer areas of college achievement (e.g., for the first semester most students in some schools enroll in subjects that fall mostly under English, and Social Science areas.

Overall Measures as Crude Approximations of Initial College Academic Achievement

Both our overall measure of initial college performance and the initial GPA do not contain all the information latent in the respective sets of individual measures from which each overall measure is linearly constituted. In Gumban and Iledan's study it is indicated that initial college achievement is multidimensional from the viewpoint of factorial vector model. Therefore, what we have are unidimensional approximations of that space. It can be noted for instance that on the average each of the subject-area measures accounts only for 52.79% and 53.05%, respectively for males and females, of the variances of our overall measure. (Since correlation is a symmetrical relation this is equivalent to saying that the overall measure accounts for only about half of the variances of each of the component measures.)

Strictly speaking, our overall measure must be evaluated for "fidelity" by comparing it simultaneously with all its component measures in order to take the covariance structure of these

[3]By equivalent it is meant that some subject-area measures can be replaced one-to-one by, say some subject grades (e.g., high school subject-area average in Mathematics may be replaced by 4th year grade in Mathematics) provided each pair of variables are highly correlated and similar in content. The latter condition is necessary to safeguard against high correlations due to "contamination."

It is suggested that the replacement is practical if about 90% of the variance of the variable to be replaced is accounted for by the substitute variable.

A subset of the subject-area measures that is adequately homogeneous can also be replaced by a single measure under the same conditions for replacement.

[4]This was made possible in the previous study by including in the samples only those examinees with complete set of data.

[5]In many schools the direction of the GPA is opposite that of the overall measure considered, i.e., the lower the numerical value the higher is the level of performance indicated.

components into consideration. Yet our rather simple analysis above is enough to show that our overall measure of initial college academic performance, and possibly the GPA, is at most a crude approximation. The same arguments can be used against reducing the CET subtest scores into an overall score (see Felipe, A. I., 1970, and Gumban, R. S., 1971 concerning the factorial structure of the CET). There is at least one alternative that is theoretically more sound, i.e., the use of factor scores. Presently, however, we have yet to generate sufficient information regarding this issue.

Relative Independence of the "Behaviors" of the Overall Measure and its Components

It becomes clear from the preceding discussion why our overall measure of initial college academic achievement is equally predictable for both male and female college freshmen "populations" despite the unequal predictabilities of its two components (viz., subject-area averages in English, and Mathematics): The "behavior" of the former is relatively independent of the "behaviors" of the latter measures.

For admissions purposes this finding is not so much a practical disadvantage because presumably only border cases (with respect to overall initial performance) would probably be evaluated on the basis of these subject-area measures as academic criteria. An applicant whose initial overall performance in terms of achievement grades is not likely to be a "success" might still be admitted if his predicted performance in the area of Social Science is likely to be a "success", provided that the student intends to specialize in the same area.

CONCLUSION

This paper presents further analysis on the effect of sex in academic prediction based on the data reported by Gumban and Iledan in a validity study on the College Entrance Test (CET). Specifically, tests of significance are presented of the hypotheses that the predictabilities of the college academic criteria considered are equal for the male and female "populations" of college freshmen.

The following are the salient findings:

1. When the overall measure and subject-area measure in Social Science are predicted from the high school achievement subject-area measures (viz., subject-area averages in English, Pilipino, Mathematics, Natural Science, and Social Science) or CET subtest scores *or* their two combinations (viz., subject-area measures combined with CET subtest scores, and overall high school achievement measure combined with CET overall score) there is invariably insignificant sex differences in the predictability of these criteria. These findings make the CET equally useful for male and female college applicants, i.e., CET-user schools need only to formulate one set of academic criteria for admissions purposes that will apply equally for both sexes, assuming that these schools use the CET for the purposes of gaining efficiency in academic prediction. The possibility of bias against one subgroup of applicants may still exist, however, because for one thing the regression surfaces may still fail to be coincident, i.e., parallel and equal in intercepts.

2. When the subject-area measures in English, and Mathematics are predicted from the sets of predictors given in (1) the following findings are obtained: (a) In all three prediction schemes the subject-area in Mathematics is predicted with greater precision for the female "population" of college freshmen; (b) except in one case (i.e., when the prediction scheme involves the set of high school subject-area measures as regressors) the subject-area measures in English is similarly predicted with greater precision for the female "population" of college freshmen. Presumably, only border cases (with respect to overall performance) in coeducational schools, thus constituting a minority group in many instances, will be evaluated for admissions purposes based on these measures. Therefore, these findings do not offer much practical constraints to the usefulness of the College Entrance Test (CET).

REFERENCES

ABELSON, R. P. Sex differences in predictability of college grades. *Educational and Psychological Measment,* 1952, *12,* 638–644.

FELIPE, A. I. *The development and structure of the college entrance test. Research Bulletin GTD-RB-71–3,* Makati, Rizal: Fund for Assistance to Private Education, 1971.

GULLIKSEN H., and WILKS, S. S. Regression tests for several samples. *Psychometrika,* 1950, *15,* 91–114.

GUMBAN, R. S. Preliminary findings on the predictive validity of the college entrance test. Paper read at the FAPE-GTD Sponsored Seminar on the CET held at Saint Joseph's College, 1971.

GUMBAN, R. S. and ILEDAN, B. R. *Some predictive validities of the college entrance test based on a relatively homogeneous group of schools. Research Bulletin GTD-RB-71–2,* Makati, Rizal: Fund for Assistance to Private Education, 1971.

HOLLAND, J. L., and RICHARDS, J. M. JR. Academic and non-academic accomplishment: Correlated or uncorrelated? *ACT Research Report No. 2.* Iowa City: American College Testing Program, 1965.

RICHARDS, J. M. JR., and LUTZ, S. W. Predicting student accomplishment in college from the ACT assessment. *ACT Research Report No. 21,* Iowa City: American College Testing Program, 1967.