ACT Research Report Series 96-8

Differential Prediction/Impact in Course Placement for Ethnic and Gender Groups

Julie Noble

Jill Crouse

Matt Schulz

ACT

December 1996

For additional copies write:
ACT Research Report Series
P.O. Box 168
Iowa City, Iowa 52243-0168

Differential Prediction/Impact in Course Placement for Ethnic and Gender Groups

Julie Noble
Jill Crouse
Matt Schulz

Abstract

This research investigated course placement decisions for ethnic and gender groups. Success in four standard college freshman courses was predicted from ACT scores or high school subject area grade averages (SGA). The number of institutions varied from 13 to 50, depending on the course. Mean between-subgroup differences in estimated probability of success and three course placement validity indices were evaluated by gender or ethnic group within course type and institution. All statistics were summarized across institutions by course type and ethnic or gender group.

The results of this study showed that, from a statistical perspective, both ACT scores and SGA slightly overpredict course success for blacks and males relative to whites and females. From a practical perspective, the differences between ethnic and gender groups were small.

Differential Prediction/Impact in Course Placement for Ethnic and Gender Groups

Research on differential prediction based on standardized test scores and high school grades (e.g., Sawyer, 1985; Young, 1994) has shown slight overprediction of the college GPAs of African-Americans relative to Caucasian-Americans, and of males relative to females. A limitation of these studies is that the criterion variables are either pooled subject area course grades (e.g., course grades from all English courses) or first-year GPA. Though generally more reliable than grades in specific courses, pooled grades or GPAs are less informative as measures of academic success in the first year in college. A second limitation is that the research does not consider or control for prior selection in admissions or course placement and resulting restriction of range problems (Linn, 1983). Finally, Linn (1994) states that such overprediction, although consistent across studies, may not be an indication of "bias," but of imperfect reliability. He suggests that the magnitude of group differences in prediction found in current research often do not represent "true" group differences.

In course placement, the typical decision is whether a student should be placed into a standard-level course or into a lower-level course (e.g., standard freshman English vs. developmental English or intermediate algebra vs. elementary algebra). For this use of placement variables (e.g., test scores or high school grades), traditional statistics, such as R^2 or χ^2 values, standard errors of estimate (SEE), or differences in linear regression slopes, appear less informative. A more meaningful approach is to determine how differential prediction affects the outcomes of placement or admissions decisions for specific subgroups. Sawyer (1993) developed an approach to placement validation that focusses on estimating the percentage of correct placement decisions made about an unselected group of students (students for whom a placement decision is to be made, referred to here as the "placement group"). The estimates are based on logistic regression models developed for students who completed a given course.

This research investigated the differential impact of course placement decisions on male, female, African-American, and Caucasian-American students. Using Sawyer's approach would show, in practical terms, the implications of course placement decisions for these population subgroups. It would also correct for restriction of range problems identified by Linn (1984) by estimating the impact of course placement decisions on an unselected group of students (i.e., the placement group).

Though research has shown that using test scores in combination with either high school grades, subject area grade averages (SGA), or high school GPA results in differential prediction, it has not compared the differential impact of using SGA in course placement with that of test scores. Therefore, test scores and SGA were used separately as predictor variables; the differential impact of course placement decisions based on SGA or on test scores was then compared.

Data

The data for this study consisted of students' grades from over 80 institutions for 11 different college courses, ACT Assessment scores (in English and Mathematics), self-reported high school English and mathematics grade averages, and gender and ethnicity information. The ACT Assessment consists of four academic tests (in English, Mathematics, Reading, and Science Reasoning) and a Composite score, a Student Profile

Section, an Interest Inventory, and the Course Grade Information Section (CGIS). Test scores are reported on a scale of 1 to 36. The CGIS collects information about students' grades in 30 specific high school courses. Self-reported grades collected by the CGIS have been found to be accurate, relative to information provided on students' transcripts (Sawyer, Laing, & Houston, 1988).

To help insure statistical stability and consistency of population subgroups across institutions, only data for the courses from each institution that had sample sizes of at least 50 and subgroup sample sizes of at least 25 were used. The sample for each course was also limited to students with the relevant ACT Assessment score (ACT English for English courses, and ACT Mathematics for mathematics courses), high school subject area grade average (English grade average for English courses and mathematics grade average for mathematics courses), and college course grade. These sample size constraints restricted the number of course types and racial/ethnic subgroups that could be examined. For the gender analyses, four courses were investigated: English composition, intermediate algebra, college algebra, and calculus. For the racial/ethnic analyses, English composition was the only course type for which there were sufficient numbers of African-Americans and Caucasian-Americans within each institution.

Some institutions provided grades for more than one course within a given course type. For these institutions, each course was analyzed separately. Summary statistics were based on courses across institutions within a given course type.

Method

Descriptive Statistics

For each course type and institution, the following descriptive statistics were computed: mean course grade, mean ACT subject area score, mean high school subject area grade average, the percentage of students with a B or higher grade in the course, and the percentage of students with a C or higher grade in the course. Distributions of these statistics were then summarized across institutions within course type using median, minimum, and maximum values.

Linear Regression

Linear regression analyses were performed to determine whether, as prior research has shown, there was differential prediction of course grades for females or males, or for African-Americans or Caucasian-Americans. Separate regression models were developed for each course using either ACT subject area score or SGA. The typical mean difference between observed and predicted course grade across institutions, based on a total group linear regression model, was used to indicate the presence and direction of differential prediction. Squared multiple correlations (R^2) and standard errors of estimate (SEE) were also developed using subgroup-specific regression analyses. The median values across institutions were used to examine differences between subgroups in the amount of variance explained and accuracy of the predicted course grades.

Logistic Regression

Three logistic regression models were developed for each course, by institution, for predicting course outcomes (B or higher, or C or higher grade):

- Total group regression model, consisting of a single prediction equation for all gender or ethnic subgroups. The only predictor was the relevant ACT Assessment test score or SGA, thus modeling the typical usage of one cutoff score for all students by an institution.

- Common slope, subgroup-specific model, consisting of the total group regression model to which a dummy variable denoting subgroup membership (e.g., male or female) was added.

- Separate slope, subgroup-specific model, consisting of the total group regression model, a dummy variable denoting subgroup membership, and the interaction between subgroup membership and ACT score (or SGA) as a predictor.

These models would illustrate the differential impact on population subgroups of using either a single cutoff score for all students or subgroup-specific cutoff scores.

Differential Prediction

For each course type, subgroup-specific probabilities of success (B or higher, or C or higher grade) were calculated using the subgroup-specific ACT score or SGA prediction model. The logistic regression weights from the models were applied to the ACT scores or SGA of all students at each institution with valid predictor data (i.e., the placement group), resulting in an estimated probability of success for each student. Then, for each course type, a mean between-subgroup difference in probability of success was computed. The difference at each ACT score or SGA was weighted by the number of females (in the gender analysis) or African-Americans (in the ethnic analysis) in the placement group for the course. The median, minimum and maximum mean differences across institutions, within course type, were calculated.

Differential Impact

For each subgroup within course type and institution, two optimum cutoff scores were identified--one based on the total group regression model (referred to as the *total group* cutoff score) and the other based on the common slope, subgroup-specific model (referred to as the *subgroup-specific* cutoff score). (This model was judged to be the optimum model on the basis of statistical significance criteria (see Results section)). Optimum cutoff scores correspond to a .50 probability of success for a given model, and maximize the estimated percentage of correct placement decisions for a given course. Using the two cutoff scores for each course type and institution, the following impact statistics for a given subgroup were estimated: 1) the percentage of students that would be placed into a lower-level course, 2) the percentage of successful students among those who would be placed into the course (success rate), and (3) the percentage of correct placement decisions (accuracy rate). Optimum cutoff scores and differential impact statistics were summarized across institutions using median, minimum and maximum values.

Estimates of the success rate and accuracy rate were calculated through the conditional probabilities of success for individual students in the placement group, as estimated by a logistic regression model (Sawyer, 1993). In order to assess the differential impact on success rates and accuracy rates of using a suboptimum cutoff score versus an optimum cutoff score, it is necessary to use the logistic regression model that best represents a student's true, conditional probability of success. Therefore, for both cutoff scores (total group and subgroup-specific), estimates of success rates and accuracy rates were based on the common slope, *subgroup-specific* model.

Results

The differential prediction and differential impact results for the C-or-higher success criterion were essentially the same as those for the B-or-higher success criterion. For some courses, however, very few students received course grades lower than a C; consequently, logistic regression equations for the C-or-higher success criterion could not be developed for these courses. The B-or-higher results are therefore reported here, to maximize the number of institutions and courses that could be studied. The C-or-higher results may be obtained from the authors.

Descriptive Statistics

The distributions of descriptive statistics are summarized, by course type, in Table 1. For each course type, and for each gender or ethnic subgroup, the number of institutions and the number of courses are reported, as well as median, minimum, and maximum mean ACT subject area score, SGA, and course grade. The last column provides the distribution of the percentages of students who received a B or higher grade.

For all courses except English composition, males typically had higher ACT scores and lower SGAs and college grades than females. A higher percentage of females than males typically had B or higher grades in English composition, but median percentage differences for the other course types did not exceed 5%. Caucasian-American students typically had higher ACT scores, slightly higher SGAs, and higher English composition grades than African-American students. Of Caucasian-American students, 60% typically had B or higher grades, compared to 39% of African-American students.

TABLE 1

Descriptive Statistics Summarized Across Institutions

Course type	No. of inst./ no. of courses	Subgrp.	Mean ACT subj. area score		Mean HS subj. area grade aver.		Mean course grade		Percentage with B or higher grade	
			Med	Min/max	Med	Min/max	Med	Min/max	Med	Min/max
English comp.	40/47	Females	20.3	14.1/25.7	3.18	2.61/3.56	2.70	2.07/3.54	65	27/95
		Males	19.7	14.1/24.7	2.88	2.29/3.28	2.40	1.70/3.42	51	22/89
Inter. algebra	13/13	Females	19.1	16.0/20.0	2.72	2.13/2.90	2.13	1.39/3.20	39	20/82
		Males	19.5	17.2/21.1	2.53	2.18/2.84	1.94	.88/3.03	35	13/78
College algebra	22/25	Females	20.9	18.2/25.4	3.14	2.46/3.60	2.28	1.05/2.78	46	18/65
		Males	21.7	18.8/25.2	2.97	2.30/3.41	2.08	.83/2.76	41	13/64
Calculus	12/16	Females	26.2	20.2/30.0	3.64	2.89/3.83	2.53	1.96/3.00	52	36/74
		Males	26.7	20.4/30.8	3.50	2.62/3.77	2.43	1.88/3.03	51	27/70
English comp.	8/11	Afr.-Am.	17.3	13.3/20.9	2.86	2.41/3.16	2.25	1.47/3.00	39	11/75
		Cau.-Am.	20.6	14.3/22.9	2.94	2.48/3.26	2.64	2.02/3.34	60	27/93

Linear Regression

The linear regression results were based on only those students who completed each course and who had valid predictor data. Therefore, the limitations of range restriction apply. Complete results can be found in Appendix A.

Overall, the results supported prior findings. The results for gender showed, based on a total group regression model, that ACT scores and SGAs slightly underpredicted the course grades of females relative to males for all course types. The differences using SGA were smaller than the differences found with ACT scores. Adding high school grades to ACT scores in a two-predictor regression model slightly reduced the underprediction of female grades for all course types.

The results for ethnic subgroups revealed that ACT scores overpredicted English composition grades of African-Americans relative to Caucasian-Americans by .15 grade units (-.14 vs. .01). SGA also overpredicted the English composition course grades of African-Americans, relative to Caucasian-Americans, by .21 grade units (-.19 vs. .02). Adding ACT scores to SGA in a two-predictor regression model slightly reduced the overprediction of English composition grades for African-Americans (to -.10).

Logistic Regression

Differential Prediction

The results showed that the total group regression models based on ACT Assessment score or SGA were statistically significant ($p<.05$) for 92% of the models based on ACT scores and 85% of the models based on SGA. The gender dummy variable added to the total group ACT or SGA regression models was statistically significant for 42% and 27%, respectively, of the courses studied; and the ethnic subgroup dummy variable was statistically significant for 23% and 63% of the courses. The interaction terms (ACT score or SGA by gender or ethnicity) were not statistically significant over and above the dummy variable models for nearly all of the courses (> 90%). The interaction terms were therefore dropped from all subsequent analyses.

Table 2 contains the median, minimum, and maximum weighted average gender differences in the estimated probability of success by course type. For every course type, females had a slightly higher median probability of success than males, based on ACT scores (.08 to .10). The median differences between gender subgroups based on SGA (.02 to .06) were slightly smaller than those based on ACT scores (.08 to .10). However, the range of gender differences in probability of success across institutions was larger for all course types when based on SGA than on ACT score.

TABLE 2

Gender Differences in Probability of Success
Using B-or-higher Success Criterion
(Female probability minus male probability)

Course type	Weighted average gender difference in probability of success					
	ACT			SGA		
	Median	Minimum	Maximum	Median	Minimum	Maximum
English composition	.08	-.04	.26	.06	-.08	.27
Intermediate algebra	.10	.02	.18	.05	-.03	.17
College algebra	.08	-.03	.22	.03	-.17	.17
Calculus	.08	-.07	.17	.02	-.20	.16

The results by ethnic subgroup showed that African-Americans had a lower median probability of success in English composition than Caucasian-Americans, whether based on ACT English score (median difference = -.11) or SGA (median difference = -.14).

Differential Impact

Gender subgroups. For every course type except English composition, using a total group ACT cutoff score would generally result in a slightly higher percentage (median difference = 4 to 14%) of females than males placed into the lower-level course. For English composition courses, the median percentage placed into lower-level courses, based on an ACT English cutoff score of 17, was 35% for females and 46% for males (see Table 3).

Using a total group SGA cutoff score would generally result in placing more males than females into lower-level English and mathematics courses. The one exception was calculus, where slightly more females than males would be placed into lower-level courses.

As shown in Table 3, among students placed into a course using a total group ACT cutoff score, the typical percentage of females who would be successful (estimated success rate) was higher (by 8 to 15%) than that for males for all courses. The largest differences were found for English composition (15%) and intermediate algebra (13%). The typical success rates based on a total group SGA cutoff were also higher for females than for males for English composition (10%) and intermediate algebra (8%). Success rates of females in college algebra and calculus were higher than those of males, but the differences were small.

The differences in estimated percentages of males and females correctly placed (accuracy rate) based on total group ACT cutoff scores were relatively small and varied across course types. The differences between medians was no greater than 3 percentage points. The accuracy rate differences based on SGA were very similar to those based on ACT scores.

TABLE 3

Differential Impact of Using Total Group or Subgroup-specific Optimum Cutoffs Across Gender Subgroups (Medians)

Course	Subgroup	ACT score				SGA			
		Opt. cutoff score	Percent placed in lower-level course	Percent successful (success rate)	Percent correctly placed (accuracy rate)	Opt. cutoff score	Percent placed in lower-level course	Percent successful (success rate)	Percent correctly placed (accuracy rate)
Total group cutoff									
English composition	Females	17	35	74	69	2.67	23	72	69
	Males		46	59	66		38	62	67
Intermediate algebra	Females	21	61	74	68	3.34	65	61	66
	Males		50	61	70		68	53	69
College algebra	Females	22	71	68	68	3.26	61	63	66
	Males		61	60	69		65	61	68
Calculus	Females	25	88	65	73	3.49	74	64	69
	Males		78	56	73		72	63	71
Subgroup-specific cutoff									
English composition	Females	16	26	71	69	2.55	18	72	70
	Males	19	55	64	67	2.82	47	64	67
Intermediate algebra	Females	19	50	67	68	3.17	63	60	66
	Males	23	63	67	72	3.54	82	56	69
College algebra	Females	21	58	64	68	3.19	60	62	67
	Males	23	66	64	70	3.36	69	61	68
Calculus	Females	25	79	64	76	3.49	73	64	69
	Males	26	79	62	74	3.39	70	62	71

Compared to the total group optimum cutoff scores, gender-specific optimum cutoff scores across institutions were slightly lower for females and slightly higher for males for every course type except calculus, as shown in Table 3. Gender-specific optimum cutoff scores were generally 1-2 scale score units lower for females and 1-2 score units higher for males than the corresponding total group cutoff scores. For calculus, the median optimum cutoff score for males was 1 scale score unit higher than the total group cutoff score.

Using gender-specific ACT cutoff scores, rather than a total group ACT cutoff score, would generally decrease the percentages of females (by 9 % to 13% and increase the percentages of males (by 5% to 13%) placed in lower-level courses for all course types except calculus. Conversely, it would decrease the typical success rates for females (by 1% to 7%) and slightly increase the success rates for males (by 4 to 6%) for all course types. Accuracy rates for gender-specific optimum ACT cutoff scores were comparable to those obtained using a total group optimum cutoff score for all course types.

Gender-specific optimum SGA cutoffs were generally lower by .2-.3 grade units for females than for males for all course types except calculus. Using gender-specific SGA cutoffs would result in more males (47% vs. 38%) and slightly fewer females (23% vs. 18%) placed into lower-level English courses. Success rates and accuracy rates would typically be comparable to those obtained using a total group SGA cutoff for all course types.

Ethnic subgroups. Using a total group ACT English cutoff score for English composition courses would typically result in a higher percentage of African-Americans than Caucasian-Americans placed into a lower-level course, as shown in Table 4 (63% of African-Americans and 35% of Caucasian-Americans). The estimated percentage of African-Americans who would be successful in English composition courses, given a total group optimum ACT cutoff score, was lower than that for Caucasian-Americans (59% vs. 68%).

TABLE 4

Differential Impact of Using Total group and Subgroup-specific Cutoffs for English Composition Across Ethnic Subgroups (Medians)

	ACT score				SGA			
Subgrp.	Opt. cutoff score	Percent placed in lower-level course	Percent successful	Percent correctly placed	Opt. cutoff score	Percent placed in lower-level course	Percent successful	Percent correctly placed
Total group cutoff								
Afr.-Am.	18	63	59	62	2.40	52	52	63
Cau.-Am.		35	68	63		29	69	70
Subgroup-specific cutoff								
Afr.-Am.	20	69	62	64	2.79	71	60	65
Cau.-Am.	18	35	68	70	2.21	26	66	70

Using a total group optimum SGA cutoff would also result in a higher percentage of African-Americans than Caucasian-Americans placed into a lower-level course; the median percentage was 52% for African-Americans and 29% for Caucasian-Americans. Compared to using a total group optimum ACT cutoff score, however, using a total group optimum SGA cutoff would result in a lower percentage of successful African-American students than Caucasian-American students. The median SGA success rate was 52% for African-Americans and 69% for Caucasian-Americans.

The typical percentage of correct decisions for African-Americans and Caucasian-Americans, based on a total group optimum ACT English cutoff score (18), was similar for Caucasian-Americans (63%) and African-Americans (62%). Using a total group optimum SGA cutoff score (2.40), an higher percentage of Caucasian-Americans than African-Americans (70% vs. 63%) would be correctly placed.

Subgroup-specific optimum ACT English cutoff scores for African-Americans were generally slightly higher than the corresponding total group optimum ACT cutoff scores. Compared to using a total group optimum ACT cutoff score, using subgroup-specific optimum ACT cutoff scores would typically result in higher percentages of African-American students placed into lower-level courses (median = 69% vs. 63%), but higher percentages of African-American students who would be successful (median = 62% vs. 59%), and slightly higher percentages of African-American students who would be correctly placed (median = 64% vs. 62%).

Median subgroup-specific optimum SGA cutoff scores were higher for African-Americans than for Caucasian-Americans, and would results in correspondingly higher percentages of African-American students placed into lower-level courses (median = 71% vs. 52% for African-Americans). Using subgroup-specific optimum SGA cutoffs, rather than total group optimum SGA cutoffs, would typically increase the percentages of correct placement decisions by 2%, and would slightly increase the percentages of African-American students, but not Caucasian-American students, who would be successful (60% vs. 52%).

Discussion

The results of this study were consistent with prior research (e.g., Sawyer, 1985), showing that <u>both</u> ACT Assessment scores and high school subject area grade averages slightly overpredict college English composition and mathematics course grades of males relative to those of females, and English composition grades of African-Americans relative to those of Caucasian-Americans. Overprediction was consistent across ACT scores and subject area grade averages; the predictor by subgroup interaction terms were not statistically significant for all course types. As a result, for any cutoff score used, somewhat different percentages of African-American and Caucasian-American students, or males and females, would be placed into standard courses.

From a practical perspective, ethnic and gender differences in the prediction of course outcomes and the impact course placement decisions were very small. Differential prediction was slight for both ethnic and gender subgroups, corresponding to the difference between a B and a B- grade. Placement accuracy was fairly consistent across ethnic and gender subgroups. Using subgroup-specific optimum cutoff scores did not change placement accuracy, but decreased the percentages of successful females for some courses (because the total group optimum cutoff score was lower than the female-specific optimum cutoff score), and increased differences between gender groups and ethnic groups in the percentages of students placed into lower-level courses.

There may be factors influencing these findings. One such factor concerns the validity and reliability of using college course grades as measures of educational achievement. Both ACT Assessment scores and high school subject area grade averages showed overprediction for males and African-Americans, which suggests that the overprediction might be attributable, at least in part, to problems with college grading practices. Grading standards differ substantially across teachers, fields, and disciplines, and often include (formally or informally) factors other than academic achievement

(Goldman, Schmidt, Hewitt, & Fisher, 1974; Shea, 1994; Stockard, Lang, & Wood, 1985). Grade inflation has been shown at both the high school (Ziomek & Svec, 1995) and college (Duke, 1983; Kapel, 1980; Shea, 1994) levels.

If college grades measure characteristics other than academic achievement, and if these characteristics are directly or indirectly related to ethnicity or gender, their inclusion or exclusion in the regression models will affect overprediction. The regression models used in this study consisted of one predictor variable, exclusive of ethnicity or gender. Including other important variables in the regression models could alter the results (Cronbach & Schaeffer, 1981; Novick, 1992). Further research on other factors related to gender or ethnicity and college success (e.g., social support, values concerning education, aspirations) would likely reduce overprediction, particularly for gender groups (Stricker, Rock, & Burton, 1993).

Another factor to consider is the reliability of ACT Assessment scores, high school grades, and college grades. Linn (1984) cautioned that over- or underprediction may be influenced by predictor or criterion unreliability, and therefore should not necessarily be interpreted as proof of predictor bias. College grades have been found to be relatively unreliable, compared to ACT Assessment scores, with reliabilities ranging from .30 to .44 for freshman course grades (Etaugh, Etaugh, and Hurd, 1972) and .39 to .76 for GPAs in 12 specific curricular areas (Schoenfeldt & Brush, 1975). Reliabilities of ACT Assessment scores generally range from .85 to .96 (ACT, in press); no reliability estimates were found for high school grades.

A final factor to consider is that the results shown here estimate the results for unselected groups of students using restricted data. Studies by Schiel and Noble (1992) and Houston (1993) have shown that severe restriction of range (greater than 25%) in test scores will adversely affect the accuracy of these estimates. As a result, if prior selection differed by ethnic or gender group, it could differentially impact the accuracy of the estimates. It would be very difficult to determine the degree of restriction occurring in the sample for this study, however.

Implications

Institutions should, where possible, investigate how their placement systems affect educationally important subgroups of students. The ACT Course Placement Service provides a convenient means to do this. Institutions that do find differential prediction or impact might consider using subgroup-specific cutoffs, or adjusting their placement requirements to balance success rates or subgroup representation in standard- or lower-level courses. However, beyond the political and legal defensibility of instituting such placement policies, there would likely be consequences in terms of either the percentages of students placed into courses who would be successful, or the disproportionate representation of population subgroups in lower-level courses.

REFERENCES

ACT (in press). *ACT Assessment Technical Manual.* Iowa City, IA: American College Testing.

Cronbach, L. J. & Schaeffer, G. A. (1981). *Extensions of personnel selection theory to aspects of minority hiring* (Project Rep. No. 81-A2). Stanford, CA: Stanford University. Institute for Research on Educational Finance and Governance.

Duke, J. D. (1983). Disparities in grading practices, some resulting inequities, and a proposed new index of academic achievement. *Psychological Reports,* **53**, 1023-1080.

Etaugh, A.F., Etaugh, C. F., & Hurd, D. E. (1972). Reliability if college grades and grade point averages: Some implications for predictions of academic performance. *Educational and Psychological Measurement,* **32**, 1045-1050.

Goldman, R. D., Schmidt, D. E., Hewitt, B. N., & Fisher, R. (1974). Grading practices in different major fields. *American Educational Research Journal,* **11**, 343-357.

Houston, W. M. (1993). *Accuracy of validity indices for course placement systems.* paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Kapel, D. E. (1980). A case history of differential grading: Do teacher education majors really receive higher grades? *Journal of Teacher Education,* **31**, 43-47.

Linn, R. L. (1983). Pearson selection formulas: Implications for studies of predictive bias and estimates of educational effects in selected samples. *Journal of Educational Measurement,* **20**, 1-16.

Linn, R. L. (1984). Selection bias: Multiple meanings. *Journal of Educational Measurement,* **21**, 33-47.

Novick, M. R. (1992). Educational testing: Inferences in relevant subpopulations. *Educational Researcher,* **11**, 4-10.

Sawyer, R. L. (1985). *Using demographic information in predicting college freshman grades.* ACT Research Report No. 87). Iowa City, Iowa: American College Testing.

Sawyer, R. (April, 1993). *Decision theory models for validating course placement systems.* A paper presented at the annual meeting of the American Educational Research Association in Atlanta.

Sawyer, R. L., Laing, J., & Houston, W. M. (1988). *Accuracy of self-reported high school courses and grades of college-bound students.* (ACT Research Report No. 88-1). Iowa City, IA: American College Testing.

Schiel, J., & Noble, J. (1992). *The effects of data truncation on estimated validity indices for course placement.* (ACT Research Report No. 92-3). Iowa City, IA: American College Testing.

Shea, C. (1994). Grade inflation's consequences. *The Chronicle of Higher Education, **40**,* A54-A46.

Schoenfeldt, L. F. & Brush, D. H. (1975). Patterns of college grades across curricular areas: Some implications for GPA as a criterion. *American Educational Research Journal,* **12**, 313-321.

Stockard, J., Lang, D., & Wood, J. W. (1985). Academic merit, status variables, and students' grades. *Journal of Research and Development in Education,* **18**, 12-20.

Stricker, L. J., Rock, D. A., & Burton, N. W. (1993). Sex differences in predictions of college grades from Scholastic Aptitude Test scores. *Journal of Educational Psychology,* **85**, 710-718.

Young, J. W. (1994). *Differential prediction of college grades by gender and by ethnicity: A replication study.* A paper presented at the annual meeting of the National Council on measurement in Education in New Orleans.

Ziomek, R. L. & Svec, J. C. (1995). *High school grades and achievement: Evidence of grade inflation.* (ACT Research Report No. 95-3). Iowa City, IA: American College Testing.

Appendix A

Linear Regression Results by Gender and Ethnic Subgroup

Course type	Subgrp.	ACT Observed - predicted* Med	ACT Observed - predicted* Min/Max	ACT R^2 Med	ACT R^2 Min/Max	ACT SEE Med	ACT SEE Min/Max	SGA Observed - predicted* Med	SGA Observed - predicted* Min/Max	SGA R^2 Med	SGA R^2 Min/Max	SGA SEE Med	SGA SEE Min/Max
English comp.	Females	.09	-.07/.36	.09	.01/.35	.84	.50/1.25	.05	-.09/.38	.09	.01/.46	.83	.43/1.26
	Males	-.11	-.40/.06	.06	.00/.26	.97	.53/1.40	-.06	-.42/.09	.10	.00/.30	.94	.53/1.36
Inter. algebra	Females	.10	.03/.33	.12	.03/.20	1.07	.97/1.37	.05	-.01/.26	.08	.01/.34	1.11	.88/1.43
	Males	-.18	-.43/-.06	.06	.00/.13	1.10	.90/1.40	-.09	-.31/.01	.09	.01/.16	1.13	.87/1.38
College algebra	Females	.13	-.12/.32	.16	.01/.29	1.07	.78/1.34	.05	-.22/.14	.12	.00/.38	1.12	.82/1.34
	Males	-.16	-.29/.18	.10	.03/.44	1.09	.81/1.31	-.05	-.18/.32	.11	.03/.42	1.11	.79/1.28
Calculus	Females	.15	-.12/.30	.19	.05/.42	.99	.76/1.26	.02	-.20/.20	.12	.00/.34	1.00	.76/1.40
	Males	-.09	-.27/.09	.13	.02/.23	1.03	.77/1.25	-.02	-.17/.15	.12	.02/.31	1.02	.81/1.33
English comp.	Afr.-Am.	-.14	-.27/.10	.07	.03/.28	.96	.62/1.21	-.19	-.38/.24	.07	.00/.18	.97	.57/1.22
	Cau.-Am.	.01	-.01/.04	.06	.00/.20	.88	.57/1.17	.02	-.02/.08	.09	.00/.21	.87	.56/1.16

*Based on a total group regression model; R^2 and SEE values are based on subgroup-specific models.