

The Effects of Alternative Interpretations of Incomplete and Withdrawal Grades on Course Placement Validity Indices

**Cheng H. Ang
Julle P. Noble**

August 1993

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

©1993 by The American College Testing Program. All rights reserved.

**THE EFFECTS OF ALTERNATIVE INTERPRETATIONS OF INCOMPLETE AND WITHDRAWAL
GRADES ON COURSE PLACEMENT VALIDITY INDICES**

Cheng H. Ang and Julie Noble

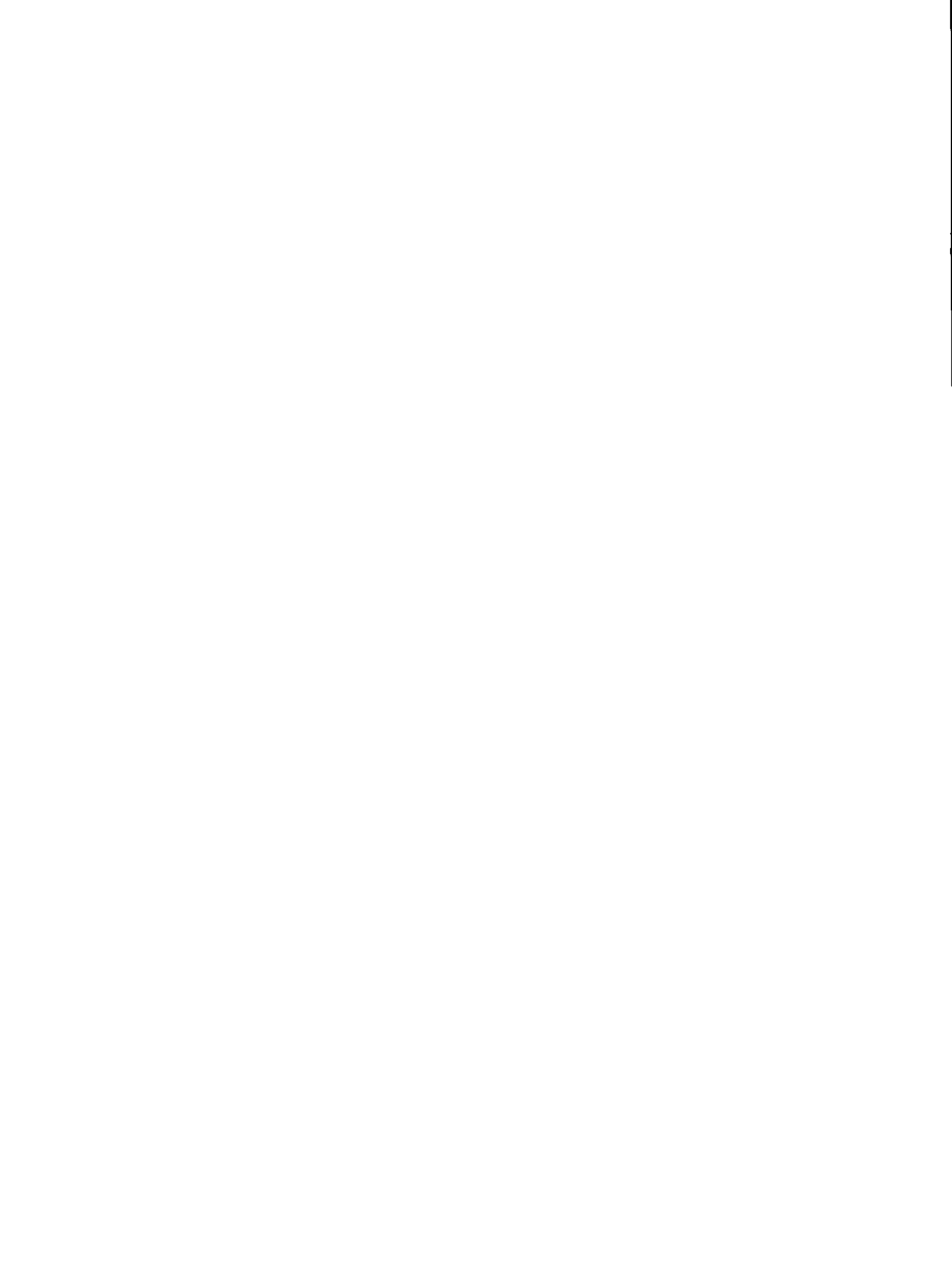


ABSTRACT

In validating tests for course placement in college, the criterion variable is usually defined in terms of the grade earned in a particular standard course. For example, success may be defined as completing the standard course with a grade of C or higher. This study considered the issue of interpreting incomplete (I) and withdrawal (W) grades either as unsuccessful outcomes or as missing data. The effects of either type of interpretation on placement indices and optimum cutoff scores were studied. The data for the study were obtained from four two-year colleges through their participation in a pilot study of the ACT Course Placement Service. Courses included were mathematics and English/reading courses. ASSET test scores were used as predictor variables.

The results of this study showed that interpreting I and W grades as unsuccessful outcomes, rather than as missing data, generally resulted in lower conditional probabilities of success, higher optimal cutoff scores, and higher estimated delta accuracy rates.

The manner in which I and W grades are interpreted should depend on an institution's policy or philosophy on those grades. Ideally, grades of I should be changed to grades of A through F or S/U before being included in the analyses of course placement criteria, and only students who received grades of W because of academic reasons should be classified as unsuccessful. If course placement accuracy indices were determined using these criteria, then the resulting optimum cutoff scores would be more accurate and appropriate.



THE EFFECTS OF ALTERNATIVE INTERPRETATIONS OF INCOMPLETE AND WITHDRAWAL GRADES ON COURSE PLACEMENT VALIDITY INDICES

Although grading policies may differ from one institution to another, most postsecondary institutions use a grading scale of A to F to denote excellent to failing performance in a course, and grades of I and W to denote incomplete work and withdrawal from a course. Some institutions may also use combinations of grades, such as withdrawal pass (WP) or withdrawal fail (WF). Students may receive grades of I or W for academic problems, such as the inability to cope with the demands of a course; or for nonacademic problems, such as illness, emotional distress, or financial difficulty. Institutions may also differ in how grades of I are maintained. Grades of I may be permanently kept on transcripts, or the grades may be changed to failure (F) if supplementary assignments (or course requirements) are not completed by a prespecified deadline.

Many colleges, particularly two-year community colleges, have an open admission policy for all high school graduates. The implementation of this policy, and attempts by colleges to provide opportunities for students to succeed, has resulted in course placement policies and developmental/remedial instruction. Because course placement cutoff scores are often developed from statistical analyses of placement test scores and course grades, the interpretation of I and W grades as either unsuccessful outcomes or missing data (deleted from the analysis) may affect the cutoff scores selected for course placement and the resulting course placement decisions made for future students.

If course placement validity indices and optimum cutoff scores differ because of the way in which I and W grades are interpreted, then thoughtful consideration must be given to how these grades are interpreted when evaluating the accuracy of course placement cutoff scores. Correct course placement decisions promote student success and foster persistence among students. Incorrect course placement decisions, however, waste students' time in school and educational expenses, as well as institutions' personnel allocations and costs.

The purpose of this paper was to compare optimal placement cutoff scores and placement validity indices resulting from two interpretations of I and W grades, where I and W grades were interpreted as unsuccessful outcomes or as missing data. The placement validity indices and cutoff

scores were developed based on the logistic regression of placement variables (i.e., test scores) on dichotomous (successful or unsuccessful) course outcomes in a particular course of interest (Sawyer, 1989).

It should be noted that because the grades assigned in a course depend on the policies of the institution and instructor, when and how I and W grades are assigned may vary from institution to institution and from instructor to instructor. Thus, the perception of what these grades mean varies among college educators and administrators. The use of I and W grades in evaluating course placement cutoff scores is therefore best determined by individual institutions and/or instructors, and cannot be addressed solely through statistical methodology.

Logistic regression can be used to estimate the conditional probability that a student would be successful in a course (e.g., a grade of C-or-higher), given the student's score on a predictor variable (e.g., placement test). The conditional probability of success estimates are based on the test scores and course grades of students in a particular course of interest. Placement validity indices can then be estimated from the conditional probability of success and the distribution of test scores for a larger group of students, those who could have taken the course (the "placement group"). These validity indices can provide information about cutoff scores used to place students into particular courses, and about the probable results of modifying such cutoff scores.

Consider the following potential outcomes for a given cutoff score:

- A. *True positive*: the student is placed in the standard-level course and is successful (Correct decision).
- B. *False positive*: the student is placed in the standard-level course and is unsuccessful (Incorrect decision).
- C. *True negative*: the student is placed in a lower-level course and would have been unsuccessful in the standard-level course (Correct decision).
- D. *False negative*: the student is placed in a lower-level course, but would have been successful in the standard-level course (Incorrect decision).

The sum of outcomes A and C is the number of students who could have taken the course and for whom correct decisions would have been made using the corresponding cutoff score and success criterion. This ratio of $A+C/A+B+C+D$ is referred to as the accuracy rate (AR). The value of AR

depends on the cutoff score, the distribution of scores, and the statistical relationship between the test score and the success criterion. The AR attains a maximum value at or around a probability of success of .50, which corresponds to the optimum cutoff score.

The delta accuracy rate (Δ AR) is an indicator of the effectiveness of the predictor variable for placing all students scoring above a specific cutoff score, and not others, in the standard course, compared to placing all students in the course. This statistic is equal to the difference between the maximum AR value and the "base line" AR value, which is the proportion of correct decisions associated with using the lowest possible score as a cutoff score.

The success rate (SR) is the estimated proportion of students in the placement group who would be placed in the standard-level course and who would be successful, given the corresponding cutoff score and success criterion. This statistic is equal to the ratio of $A/(A+B)$.

The lower-level course placement rate (LPR) refers to the proportion of all students in the placement group who would not be admitted to the standard-level course, given the corresponding cutoff score and success criterion. This statistic is equal to the ratio of $C+D/A+B+C+D$.

Data for the Study

Criterion Variables

The data for the study were obtained from four, two-year community colleges through their participation in the ACT Course Placement Service Pilot Study. The criterion variables were grades in mathematics, English, and reading courses. The course grades were scaled from A(4) to F(0) and I and W; courses graded as satisfactory or unsatisfactory (S/U) were not included in the study. Two definitions of course success were studied: B-or-higher and C-or-higher. Students were considered as successful if they achieved the specified success criterion. I and W grades were interpreted either as unsuccessful outcomes (i.e., below the success criterion) or as missing data (not included in the logistic regression analyses). For each institution, only those courses with sample sizes of at least 25 were included.

Predictor Variables

ASSET test scores were used as predictor variables in this study. The ASSET Basic Skills and Advanced Mathematics tests are designed to measure important and essential academic skills and knowledge needed for success in specific two-year college freshman courses. ASSET test scores are reported on a score scale of 23 to 55. For mathematics courses, ASSET test scores included as predictor variables were Numerical Skills, Elementary Algebra, Intermediate Algebra, and College Algebra scores. For English and reading courses, ASSET test scores included as predictor variables were ASSET Reading Skills and Writing Skills scores. Institutions submitted to ACT student records containing course grades; these records were matched with ASSET test scores from the ASSET history files using students' Social Security numbers.

The Estimation Sample and Placement Groups

Two types of samples are needed to estimate validity indices: the estimation sample and the placement group. The estimation sample for each course is used to develop the logistic regression models. In this study, the estimation sample consisted of students who completed the course of interest with a grade of A-F or who received an I or W, and who had the relevant ASSET test scores. Because I and W grades were interpreted in two ways, there were two estimation samples for each predictor variable and course: one for the analysis where I and W grades were interpreted as unsuccessful outcomes and the other where I and W grades were interpreted as missing data.

The placement group is the population of students for whom a placement decision must be made; the course placement validity indices pertain to this group. The placement group used in this study included all ASSET-tested students from an institution, regardless of course grades. The same placement group was used for both interpretations of I and W grades.

Method

Descriptive Statistics

Means and standard deviations were computed for courses grades (I & W = missing data) and

test scores, by institution. Simple correlations were also calculated between test scores and course grades of A - F.

I and W grades could be interpreted as unsuccessful outcomes, but could not be appropriately interpreted as grades of F: unsuccessful outcomes in this study included grades of C - F or D - F, depending on the success criterion. Therefore, for this interpretation, descriptive statistics were calculated for test scores only.

The distributions of institutional statistics were then summarized across the four institutions using minimum, median, and maximum values. Only test scores that were statistically significantly correlated with course grades (I & W = missing data) or course grades (I & W = missing data) that were statistically significantly correlated with test scores were included in the summaries.

Logistic Regression

For each institution, all courses with grades A - F that were statistically significantly ($p < .05$) correlated with test scores were included in the logistic regression analyses. For the logistic regression analyses in this study, course grades were dichotomized into successful or unsuccessful outcomes, based on either a B-or-higher or a C-or-higher success criterion. I and W grades were interpreted either as unsuccessful outcomes or as missing data. For all predictor models, the logistic regression equation, the regression weight for each predictor model, and the probability of success were computed for students who completed each course (estimation sample). If the regression models were statistically significant ($p < .05$), the parameter estimates and probabilities of success were applied to the placement group to compute estimated placement validity indices.

Optimum Cutoff Scores and Lower-level Course Placement Rates

Using the estimated probabilities of success from the statistically significant logistic regression models, optimum cutoff scores and lower-level course placement rates were calculated for every course and institution. The optimum cutoff values corresponded to approximately the .50 probability of success. The cutoff scores and the corresponding lower-level course placement rates were determined using the two interpretations of I and W grades and two definitions of course success

(B-or-higher and C-or-higher grades). Minimum, median, and maximum optimum cutoff scores were then calculated across the four institutions for the mathematics and English/reading courses.

Optimum cutoff scores and lower-level course placement rates for the two interpretations of grades I and W were then compared.

Placement Validity Indices

For every statistically significant prediction model ($p < .05$), estimated accuracy rates (AR), success rates (SR), and delta accuracy rates (Δ AR) were identified for the optimum cutoff score. These validity indices were computed for the B-or-higher and C-or-higher criteria, using both interpretations of I and W grades. Minimum, median, and maximum AR, SR, and Δ AR were then calculated across the four institutions, and differences in validity indices for the two interpretations of I and W grades were examined.

Although all statistically significant ($p < .05$) models were included in the computation of placement validity indices, models that yielded minimum probabilities of success greater than .50 or maximum probabilities of success less than .50 were not included in the study. This was because Δ ARs could not be computed for models with these probabilities of success. In addition, differences in placement validity indices for the two interpretations of I and W grades could be examined only when the results using both interpretations met the above probability requirements.

Results

Descriptive Statistics

Tables 1 and 2 contain the distributions of descriptive statistics and correlations for mathematics (Table 1) and English (Table 2) course grades and the relevant predictor variables. The quantile column in each table shows the minimum, median, and maximum values across all courses and institutions for a particular ASSET test. The minimum and maximum values represent the range of values obtained across courses, and the median values represent the typical values for a course.

For the results for individual institutions, please contact the authors.

Although only test scores that were statistically significantly correlated with course grades ($p < .05$) or course grades that were statistically significantly correlated with test scores were included in the summaries in Tables 1 and 2, an exception was made for a test score/English course grade correlation whose p-value was at .0515. The corresponding logistic regression analysis between test scores and grades for this course was statistically significant ($p < .05$). This course was therefore included in these statistics and in the subsequent placement validity indices computations.

Tables 1 and 2 also illustrate the results of interpreting I and W grades as missing data or as unsuccessful outcomes. The two interpretations are identified in Column 3 of the tables. When I and W were interpreted as unsuccessful outcomes, mean ASSET test scores for both subject areas were typically about the same or lower and the standard deviations were typically about the same or larger than when Is and Ws were interpreted as missing data.

Logistic Regression Analysis

Test scores that statistically significantly correlated ($p < .05$) with course grades of A - F were chosen for the logistic regression analyses. Statistically significant test scores from the correlational analysis also showed statistically significant logistic regression models ($p < .05$) for both interpretations of I and W grades. The total number of courses with statistically significant logistic regression models were the same as those reported for the correlational analyses in Column 1 of Tables 1 and 2.

For the B-or-higher success criterion, all models had minimum probabilities of success less than .50. For the C-or-higher success criterion, however, many of the models had minimum probabilities of success greater than .50. For each predictor, there were more models with minimum probabilities of success greater than .50 when Is and Ws were interpreted as missing (13 to 64% of the models) than when they were interpreted as unsuccessful (13 to 29%). Because the ΔAR could not be computed when the minimum probabilities of success were greater than .50, only those models with a minimum probability of success less than .50 were used for the comparison of placement validity indices.

Cutoff Scores and Lower-level Course Placement Rates

The Cutoff score and Lower-level course placement rate columns in Tables 3 and 4 show the minimum, median, and maximum optimum cutoff scores and their associated lower-level course placement rates for mathematics and English courses. As expected, the B-or-higher success criterion resulted in higher cutoff scores than the C-or-higher success criterion for both mathematics and English courses. In addition, interpreting Is and Ws as unsuccessful outcomes resulted in substantially higher minimum, median, and maximum cutoff scores than interpreting Is and Ws as missing data. As a result, when Is and Ws were interpreted as unsuccessful outcomes, the lower-level course placement rate at the optimum cutoff score was typically larger than when Is and Ws were interpreted as missing data (e.g., a median value of .81 versus .58 for Numerical Skills score and mathematics course grade).

Placement Validity Indices

Tables 3 and 4 also show the minimum, median, and maximum placement validity indices (accuracy rate, delta accuracy rate, and success rates) associated with optimal cutoff scores across all institutions. For mathematics courses, using the B-or-higher success criterion, the median ARs and Δ ARs were generally higher when I and W grades were interpreted as unsuccessful outcomes than when they were interpreted as missing data. Median SRs, however, were generally lower when I and W grades were interpreted as unsuccessful outcomes.

The results for the C-or-higher success criterion were similar to those for the B-or-higher success criterion, except median ARs were lower when I and W grades were interpreted as unsuccessful outcomes than when they were interpreted as missing data.

Median ARs and SRs for English/reading courses, using the B-or-higher success criterion, were generally lower when I and W grades were interpreted as unsuccessful outcomes than when they were interpreted as missing data. Median Δ ARs, however, were higher when I and W grades were interpreted as unsuccessful outcomes. The results for the C-or-higher success criterion were generally similar to those for the B-or-higher success criterion.

Discussion

Interpreting I and W grades as unsuccessful outcomes rather than as missing data resulted in lower conditional probabilities of success and higher optimum cutoff scores. Logistic regression analyses that included students with grades of I and W also resulted in larger sample sizes and therefore less sampling error in estimating the conditional probability of success function and corresponding placement accuracy indices.

Students might receive grades of I or W for different reasons, some academic and others nonacademic. If a large proportion of students received I and W grades for nonacademic reasons, but I and W grades were interpreted as unsuccessful when evaluating course placement results, the optimum cutoff scores could be overestimated and higher rates of false negatives could result. Similarly, if a large proportion of students received I and W grades for academic reasons, but I and W grades were not included in the analyses, the optimum cutoff scores could be underestimated and higher rates of false positives could result.

The manner in which I and W grades should be interpreted depends on an institution's policy or philosophy about these grades. Ideally, grades of I should be changed to grades of A through F or S/U before being included in the analyses of course placement criteria, and only students who received grades of W because of academic reasons should be interpreted as unsuccessful. If course placement accuracy indices were developed based on these criteria, then optimum cutoff scores would be more accurate and appropriate.

The results of this study were based on data from four two-year community colleges, and thus cannot be generalized to all community colleges with ASSET course placement systems.

REFERENCES

- Noble, J. (1991). Predicting college grades from ACT Assessment scores and high school course work and grade information. (ACT Research Report No. 89-4). Iowa City, Iowa: American College Testing.
- Peterson, A. (1980). Community college grading policies. Bethesda, MD: (ERIC Document Reproduction Service No. ED 187 374).
- Sawyer, R. (1989). Validating the use of ACT Assessment scores and high school grades for remedial course placement in college. (ACT Research Report No. 89-4). Iowa City, Iowa: American College Testing.

Table 1

Distributions of Descriptive Statistics for Mathematics Course Grades and Predictor Variables, Across All Institutions

ASSET test	Quantile	Interpretation of I and W	N	r	p	Test score				Course grade			
						Min	Max	Mean	SD	Min	Max	Mean	SD
Numerical Skills (8 courses)	Max	Unsuccessful	2831			33	55	45.9	5.6				
		Missing	2189	.33	.0113	36	55	46.2	5.6	0.0	4.0	2.8	1.5
	Median	Unsuccessful	375			28	55	42.2	4.4				
		Missing	272	.22	.0003	29	55	42.4	4.3	0.0	4.0	2.4	1.4
	Min	Unsuccessful	147			23	46	35.9	3.1				
		Missing	116	.16	.0001	23	46	36.2	2.9	0.0	4.0	1.9	1.2
Intermediate Algebra (2 courses)	Max	Unsuccessful	354			25	54	39.5	6.0				
		Missing	305	.47	.0149	25	54	39.8	5.9	0.0	4.0	2.6	1.5
	Min	Unsuccessful	54			23	49	33.8	5.8				
		Missing	49	.14	.0006	23	49	33.5	5.8	0.0	4.0	2.5	1.4

Note. Quantile column refers to the minimum, median, and maximum values across all courses and institutions for a particular ASSET test.

(continued on next page)

Table 1 (cont.)

Distributions of Descriptive Statistics for Mathematics Course Grades and Predictor Variables, Across All Institutions

ASSET test	Quantile	Interpretation of I and W	N	r	p	Test score				Course grade			
						Min	Max	Mean	SD	Min	Max	Mean	SD
Elementary Algebra (3 courses)	Max	Unsuccessful	836			23	55	44.2	6.9				
		Missing	725	.22	.0077	23	55	44.6	7.2	0.0	4.0	2.6	1.4
	Median	Unsuccessful	457			23	51	36.5	6.0				
		Missing	387	.18	.0001	23	51	36.8	5.8	0.0	4.0	2.5	1.3
	Min	Unsuccessful	186			23	50	29.7	5.8				
		Missing	145	.14	.0001	23	50	30.1	5.5	0.0	4.0	2.3	1.3
College Algebra (2 courses)	Max	Unsuccessful	240			24	47	33.7	5.0				
		Missing	215	.40	.0149	24	47	33.9	5.1	0.0	4.0	2.6	1.3
	Min	Unsuccessful	47			23	38	28.0	4.9				
		Missing	37	.32	.0001	23	38	28.3	5.0	0.0	4.0	2.5	1.2

Note. Quantile column refers to the minimum, median, and maximum values across all courses and institutions for a particular ASSET test.

Table 2

Distributions of Descriptive Statistics for English/Reading Course Grades and Predictor Variables, Across All Institutions

ASSET test	Quantile	Interpretation of I and W	N	r	p	Test score				Course grade			
						Min	Max	Mean	SD	Min	Max	Mean	SD
Writing Skills (14 courses)	Max	Unsuccessful	7078			33	54	46.2	5.4				
		Missing	5954	.26	.0515	34	54	46.4	5.4	0.0	4.0	2.9	1.4
	Median	Unsuccessful	581			23	54	42.1	4.6				
		Missing	489	.19	.0003	24	54	42.3	4.5	0.0	4.0	2.5	1.2
	Min	Unsuccessful	145			23	47	33.7	2.6				
		Missing	126	.12	.0001	23	47	34.3	2.6	0.0	4.0	2.3	1.1
Reading Skills (9 courses)	Max	Unsuccessful	6699			27	53	43.9	5.6				
		Missing	5647	.21	.0350	28	53	43.9	5.5	0.0	4.0	2.8	1.4
	Median	Unsuccessful	1013			23	53	42.2	4.5				
		Missing	871	.14	.0001	24	53	42.3	4.4	0.0	4.0	2.6	1.2
	Min	Unsuccessful	338			23	43	32.5	2.4				
		Missing	254	.10	.0001	23	41	32.7	2.2	0.0	4.0	2.4	1.1

Note. Quantile column refers to the minimum, median, and maximum values across all courses and institutions for a particular ASSET test.

Table 3

Optimum Cutoff Scores and Associated Placement Indices for Mathematics Courses

Predictor variable	Interpretation of I and W	Cutoff score			Lower-level placement rate			Accuracy rate			Delta accuracy rate			Success rate		
		Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.
B-or-higher success criterion																
Numerical Skills (8 courses)	Unsuccessful	42	46	49	.06	.81	.96	.60	.67	.73	.13	.31	.46	.54	.58	.64
	Missing	34	42	43	.17	.58	.71	.61	.63	.68	.02	.16	.23	.60	.61	.68
Elementary Algebra (3 courses)	Unsuccessful	36	38	47	.42	.78	.84	.58	.64	.64	.06	.25	.26	.54	.57	.59
	Missing	30	30	41	.14	.59	.62	.57	.58	.60	.01	.09	.09	.56	.60	.61
Intermediate Algebra (2 courses)	Unsuccessful	32	--	42	.44	--	.66	.59	--	--	.08	--	.14	.57	--	.68
	Missing	31	--	36	.36	--	.36	.57	--	--	.04	--	.10	.58	--	.74
College Algebra (2 courses)	Unsuccessful	30	--	35	.18	--	.43	.72	--	--	.05	--	.14	.76	--	.78
	Missing	27	--	33	.11	--	.29	.72	--	--	.02	--	.10	.74	--	.85
C-or-higher success criterion																
Numerical Skills (7 courses)	Unsuccessful	35	40	43	.19	.51	.71	.61	.64	.67	.03	.14	.24	.56	.62	.69
	Missing	29	33	36	.02	.15	.29	.65	.66	.76	.00	.01	.06	.67	.68	.77

Table 4

Optimum Cutoff Scores and Associated Placement Indices for English/Reading Courses

Predictor variable	Interpretation of I and W	Cutoff score			Lower-level placement rate			Accuracy rate			Delta accuracy rate			Success rate		
		Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.	Min.	Med.	Max.
B-or-higher success criterion																
Writing Skills (14 courses)	Unsuccessful	36	41	49	.17	.41	.89	.59	.61	.71	.02	.08	.41	.54	.62	.69
	Missing	28	36	46	.01	.19	.69	.59	.63	.73	.00	.04	.21	.58	.67	.75
Reading Skills (9 courses)	Unsuccessful	35	39	51	.13	.32	.92	.55	.60	.72	.01	.05	.20	.51	.58	.74
	Missing	29	34	39	.01	.10	.41	.57	.64	.77	.00	.01	.08	.58	.65	.78
C-or-higher success criterion																
Writing Skills (5 courses)	Unsuccessful	31	34	42	.04	.21	.63	.60	.64	.72	.00	.02	.20	.61	.63	.73
	Missing	24	30	37	.00	.00	.15	.68	.78	.86	.00	.00	.02	.69	.78	.86



