CollegeBoard
inspiring minds™

Research Report
No. 2009-2

Predictive Validity of ACCUPLACER® Scores for Course Placement: A Meta-Analysis

Krista D. Mattern and
Sheryl Packman

www.collegeboard.com

College Board Research Report No. 2009-2

Predictive Validity of ACCUPLACER® Scores for Course Placement: A Meta-Analysis

Krista D. Mattern and Sheryl Packman

The College Board, New York, 2009

Krista Mattern is an associate research scientist at the College Board.

Sheryl Packman is an assistant research scientist at the College Board.

The College Board

The College Board is a not-for-profit membership association whose mission is to connect students to college success and opportunity. Founded in 1900, the College Board is composed of more than 5,600 schools, colleges, universities and other educational organizations. Each year, the College Board serves seven million students and their parents, 23,000 high schools, and 3,800 colleges through major programs and services in college readiness, college admission, guidance, assessment, financial aid, enrollment, and teaching and learning. Among its best-known programs are the SAT®, the PSAT/NMSQT® and the Advanced Placement Program® (AP®). The College Board is committed to the principles of excellence and equity, and that commitment is embodied in all of its programs, services, activities and concerns.

For further information, visit www.collegeboard.com.

Printed in the United States of America.

Contents

Tables

Abstract

A disconnect between the educational requirements of secondary institutions and postsecondary institutions often results in a large percentage of first-year college students requiring remediation (Moss & Bordelon, 2007). As such, postsecondary institutions administer tests to incoming students for placement into courses of the appropriate difficulty level. However, research on the efficacy of placement testing has revealed mixed results. Accordingly, the purpose of this study was to meta-analyze a large sample of ACCUPLACER® placement studies in order to estimate its true validity of placement decisions by correcting for statistical artifacts. Data included all ACCUPLACER placement validity studies completed through the Admitted Class Evaluation Service™ (ACES™) between 2001 and 2006. Forty-seven studies were examined, which comprised data from 17 unique institutions. When course success was defined as obtaining a grade of "B or higher," the average sample-size-weighted observed correlation between test scores and course success was 0.28; and when course success was defined as obtaining a grade of "C or higher," the average sample-size-weighted observed correlation was 0.24. After statistical artifacts were corrected, those estimates increased to 0.47 and 0.38, respectively. Furthermore, when these cut scores were applied, the percentage of students that were placed correctly ranged from 58 to 84 percent across ACCUPLACER tests. Overall, results indicated a moderate to strong relationship between ACCUPLACER scores and course success, demonstrating that ACCUPLACER test scores provide utility in terms of placing students into courses in which they are likely to succeed.

Introduction

Performance in the first few courses of postsecondary education forms the basis for predicting the probability of students' future success in college, specifically at the institution where the introductory courses are taken (Armstrong, 2000). This is often complicated by the significant disconnect found between high school performance and college readiness (Berkner & Chavez, 1997; Moss & Bordelon, 2007; Schuemann, 1998). To address this gap without denying admission, institutions often utilize placement testing as a heuristic to assess the knowledge and skills of first-year students for placement in courses that are appropriate for their current knowledge level. Remedial course work can then be utilized to ensure that students are prepared for college-level work.

The main goal of placement testing is to enroll students in courses that are aptly challenging to their current knowledge level so as not to bore or frustrate, which can lower motivation to perform. For this process to work, placement testing policies (e.g., established cut scores for entrance into specific courses) need to be continuously reviewed and evaluated to ensure that students are in fact being placed into courses that will maximize the probability of their success. However, research on the topic is largely limited and has produced mixed results, which makes it difficult to make definitive statements regarding the utility of placement tests and subsequent course placement.

Placement Testing Research

The validity of the use of placement test scores is typically researched through the correlation of test scores and course grades, which are occasionally dichotomized into successful completion (e.g., B or higher, C or higher) versus unsuccessful completion, or through the percentage of students who were correctly placed, again in terms of obtaining grades of C or higher or grades of B or higher. The majority of studies utilize the former method. For example, Armstrong (2002) examined the predictive validity of placement scores with lettered course grades (i.e., on the A through F grade scale) in English and mathematics courses. Results indicated that test scores and subsequent course grades correlated weakly. For English courses, the observed correlation was approximately 0.25 for full-time and part-time instructors, and the correlations for mathematics were 0.14 for full-time and 0.19 for part-time instructors. Additionally, demographic and situational variables (i.e., employment hours, support for attending school, income, financial aid, part- or full-time attendance, and family responsibilities) were found to account for more of the variance in course grades than test scores. Armstrong attributed the low-to-modest predictive validity to the high variability in the dependent variable of course grade. That is, course performance is dependent on not only a student's ability, but also other factors not measured by placement tests such as motivation, perseverance, and attendance.

A similar study was conducted by the Educational Testing Service, where data from 50 postsecondary schools were analyzed (College Board, 2003). Findings indicated moderate correlations between ACCUPLACER test scores and the parallel course grades for Reading

Comprehension (*r* ranged between 0.18 and 0.38), Sentence Skills (*r* ranged between 0.15 and 0.34.), Arithmetic (*r* ranged between 0.31 and 0.38), Elementary Algebra (*r* ranged between 0.19 and 0.38), and College-Level Mathematics (*r* ranged between 0.25 and 0.53). Similar to the conclusions reached by Armstrong (2002), the low-to-modest predictive validity was attributed to the high variability in the dependent variable of course grade as well as diversity in placement policies and course curricula. Furthermore, these results were not corrected for statistical artifacts, even though test scores were used to place students into courses, resulting in range restriction and therefore underestimated correlation coefficients. Additional research on the validity of placement testing has reached similar conclusions (e.g., Cohen & Brawer, 1987; College of the Canyons, 1994; Gabe, 1989; Hargis, 1990; Hughes & Nelson, 1991; Isonio, 1991, 1992; Rasor & Barr, 1993; Smittle, 1995).

Studies that examined the validity of placement tests for course placement by identifying the percentage of students correctly placed have found more positive support for the validity of such tests. For example, Saunders (2000) examined the relationship between placement into an entry-level writing course at a community college based on ACCUPLACER Sentence Skills scores and course success, which was defined as obtaining a C or higher in the course. Findings showed that 82 percent of students were accurately placed into an entry-level writing course and that 64 percent of those students passed the course (i.e., obtained a grade of C or higher). Given that this study contradicts the findings of other studies that analyzed the utility of placement testing and subsequent course placement, further investigation is necessary to corroborate or explicate either conclusion.

Overview of Current Study

While past research on placement testing suggests that placement test scores are not strongly related to subsequent course performance, these results may be misleading for two reasons. First, the test scores used to predict course success are often the same criterion used to place students into those courses. Therefore, restriction of range in the predictor suppresses the relationship between test scores and course performance. Second, additional statistical artifacts of sampling error and measurement error may also lead to downwardly biased estimates. Accordingly, the purpose of this study was to meta-analyze a large sample of placement studies to estimate the true validity of placement decisions by correcting for these aforementioned statistical artifacts. Specifically, data from all ACES ACCUPLACER placement validity studies completed between 2001 and 2006 were analyzed to determine the overall relationship between ACCUPLACER test scores and success in postsecondary courses across institutions.

ACCUPLACER® Tests

The College Board's ACCUPLACER program, which is a computer-adaptive placement testing system used to assess students' knowledge and skills in a variety of subject areas, is one of the most commonly employed test batteries for placement purposes. For example, in 2008, more than 1,300 institutions used ACCUPLACER tests and nearly seven million exams were administered. ACCUPLACER tests measure the following six dimensions: reading comprehension, sentence skills, arithmetic, elementary algebra, college-level mathematics, and writing. These tests are composed of multiple-choice items, except for the writing test, which requires students to compose a writing sample. For the College-Level Math, Reading Comprehension, and Sentence Skills tests, students are administered 20 items, whereas there are 17 items on the Arithmetic test and 12 on the Elementary Algebra test. The scores for these tests range from 20 to 120. WritePlacer®, the ACCUPLACER writing test, is scored electronically using IntelliMetric, an artificial intelligence-based, writing-sample scoring tool, and has a score scale of 2 to 12. For more information about the content and psychometric properties of the ACCUPLACER tests, refer to the *ACCUPLACER OnLine Technical Manual* (College Board, 2003).

Institutions that use ACCUPLACER scores for course placement are encouraged to conduct placement validity studies through ACES to determine whether their placement policies are appropriate for the students and courses at their institutions. Specifically, the validity of cut scores, or the test scores required for placement into one course (e.g., English 101) over another course (e.g., English 102), is examined (College Board, 2007). Through ACES, institutions receive information on the relationship between ACCUPLACER scores and subsequent course performance at their institutions. However, data on individual institutions do not provide information on the generalizability of these results to other institutions, which was an additional goal of the proposed study.

Method

From 2001 through 2006, 47 ACES ACCUPLACER placement validity studies were conducted for 17 unique institutions, which comprised the data used in the current analyses. Of the 17 institutions, 14 (82.4 percent) were two-year community colleges. Across studies, institutions

examined the validity of ACCUPLACER tests for placement for a variety of courses. For the math-related tests (i.e., Arithmetic, Elementary Algebra, and College-Level Mathematics), courses examined ranged from basic mathematics to precalculus. For the verbal-related tests (Reading Comprehension, Sentence Skills, and Writing), institutions mainly examined composition or reading courses. For each placement report, ACES provided the correlation between ACCUPLACER test scores and course success, the percentage of students correctly placed, and the probability of success in a given course given different ACCUPLACER scores. Course success was operationalized two ways: (1) obtaining a "B or higher" in the course and (2) obtaining a "C or higher" in the course. The percentage of students correctly placed was determined by each individual's probability of success. Students who had a 50 percent chance or higher were predicted to succeed, whereas students who had a probability that was less than 50 percent were predicted not to succeed. ACES uses a cut value of 50 percent because this value maximizes the percentage of cases correctly classified; however, if a false positive or a false negative is more severe than the other, a different cut value may be used in order to minimize the more grave misclassification. This classification was compared to actual success in the course to determine the percentage correctly placed.

Hunter and Schmidt's (1990, 2004) meta-analytic method was used to quantitatively aggregate results from the previous ACES studies. Given that most of the ACES studies did not include data on every ACCUPLACER test, the number of studies included per analysis varied by ACCUPLACER test, ranging from 1 to 35 with an average of 11 studies per analysis. For each ACCUPLACER test and test combination, the sample-size-weighted average correlation (r_{obs}) and standard deviation (SD_{obs}) between ACCUPLACER test scores and course success, aggregating across different courses, were calculated. The sample-size-weighted average of the percentage of students correctly placed (percent CP) was also computed.

Using a computer program developed by Schmidt and Le (2004), meta-analytic procedures were used to correct for statistical artifacts of sampling error, range restriction, and measurement unreliability. To correct for range restriction, the population of interest included all students who were at least 18 years of age and who had taken the ACCUPLACER test the same year the study was conducted. Correlations were also corrected for unreliability of the predictor and criterion. The *ACCUPLACER Online Technical Manual* provided estimates of the internal consistency of the ACCUPLACER tests examined in the current study (see Table 1). Replicating the method used by Kuncel, Hezlett, and Ones (2004) to correct for measurement error of the criterion (i.e., course grades), the estimates of 0.84, 0.84, and 0.80, which are based on studies by Reilly and Warech (1993), Barritt (1966), and Bendig (1953), respectively, were used to create an artifact distribution for criterion unreliability.

Table 1

Reliability Estimates of ACCUPLACER Test Scores Obtained from the *ACCUPLACER Online Technical Manual*

ACCUPLACER Test	*α*
Arithmetic	0.92
Elementary Algebra	0.92
College-Level Math	0.86
Reading Comprehension	0.87
Sentence Skills	0.91
WritePlacer	0.71

Note: For WritePlacer, the reliability estimate is an inter-rater reliability estimate based on the correlation between the scores obtained by the computer and the mean score of the human graders. The 95 percent CI is 0.39–0.88. The reliability estimates for the other tests are based on item response theory. Simulation studies were conducted to estimate the conditional standard error of measurement (CSEM) along various point of the score scale. For each 10-point interval on the score scale (e.g., 20–30), the CSEM was computed based on 200 simulated examinees (1,800 total). Next, the weighted average CSEM for the entire score scale was computed. Finally, this value was then squared, divided by the total test variance, and then subtracted from 1 to obtain the reliability.

To estimate the true correlation (ρ) between scores on ACCUPLACER exams and subsequent performance in college courses, correlations were corrected for all statistical artifacts discussed above. This value represents the best estimate of the true population correlation by removing sampling error, range restriction, criterion unreliability, and predictor unreliability. However, for practical purposes, correlations corrected only for sampling error, restriction of range, and criterion unreliability, referred to as operational validity $r_{(op)}$, provide useful information for course placement because decisions are made with imperfectly reliable measures. Both values were computed and presented below.

The residual standard deviation (SD_{res}), which is the standard deviation of the observed correlations after corrected for study artifacts, and the standard deviation of true score validities (SDρ), were also computed. The magnitude of SDρ is an indication of the presence of moderator variables, whereby larger values suggest that variability in true score validities may be due to uncorrected statistical artifacts, other methodological differences, and unidentified moderators. If substantial variability remained, the effect of potential moderator

variables, such as course difficulty/level and institution type (i.e., two- or four-year institution), would be explored to explain the findings.

Results

For both measures of success (i.e., obtaining a "B or higher" and obtaining a "C or higher" in the course), sample-size-weighted correlations were calculated for individual ACCUPLACER tests: Arithmetic, Elementary Algebra, College-Level Mathematics, Reading Comprehension, Sentence Skills, and WritePlacer. Wherever data were available, combinations of ACCUPLACER tests were also analyzed. Combinations of ACCUPLACER tests included: Elementary Algebra and Arithmetic Elementary; Algebra and College-Level Mathematics; Reading Comprehension and Sentence Skills; Reading Comprehension and WritePlacer; Sentence Skills and WritePlacer; and Reading Comprehension, Sentence Skills, and WritePlacer.

On average, the results showed higher correlations for combinations of tests in comparison to individual tests (see Tables 2 and 3). Specifically, when success was defined as obtaining a "B or higher," the observed biserial correlations (r_{obs}) for individual ACCUPLACER tests ranged across institutions from 0.16 to 0.36 (M = 0.24), and correlations for combinations of tests ranged from 0.23 to 0.62 (M = 0.32). The percentage of students correctly placed using an individual test ranged from 58.4 to 66.5 percent (M = 63.0 percent) and the percentage of students correctly placed using a combination of tests ranged from 58.0 to 81.0 percent (M = 63.8 percent). When success was defined as obtaining "C or higher," the observed correlations (r_{obs}) ranged from 0.10 to 0.32 (M = 0.19) for individual tests and 0.15 to 0.58 (M = 0.26) for combinations of tests. The percentage correctly placed using individual tests ranged from 73.0 to 83.7 percent (M = 77.1 percent) and using combinations of tests ranged from 69.0 to 81.0 percent (M = 74.0 percent). Based on Cohen's (1992, 1988) general guidelines, the observed correlations constitute a small-to-medium effect and parallel what has been found in previous work.

Table 2

Overall Relationship Between Individual ACCUPLACER Tests and Course Success

Test	*N*	*k*	r_{obs}	SD_{obs}	SD_{res}	$r_{(op)}$	ρ	SD_{ρ}	*80 percent cred.*	*Percent CP*
B or Higher										
ARIT	1,824	13	.29	.08	.00	.45	.46	.00	.46–.46	66.4
EA	7,307	34	.27	.12	.10	.48	.50	.00	.50–.50	64.5
CLM	2,135	9	.36	.11	.10	.57	.62	.09	.51–.73	66.5
RC	12,699	25	.17	.08	.07	.34	.36	.09	.24–.48	62.4
SS	12,485	21	.19	.05	.03	.34	.36	.00	.36–.36	59.1
WP	3,408	8	.16	.06	.03	.35	.42	.00	.42–.42	58.4
C or Higher										
ARIT	1,824	13	.23	.12	.09	.35	.37	.08	.26–.47	83.7
EA	7,307	34	.25	.13	.11	.45	.47	.10	.34–.60	73.0
CLM	2,135	9	.32	.13	.11	.51	.55	.13	.38–.72	75.1
RC	12,699	25	.10	.09	.08	.21	.22	.15	.03–.41	80.3
SS	12,485	21	.13	.05	.03	.23	.24	.00	.24–.24	74.5
WP	3,408	8	.13	.07	.05	.29	.34	.07	.26–.43	75.0

Note: ARIT = Arithmetic; EA = Elementary Algebra; CLM = College-Level Math; RC = Reading Comprehension; SS = Sentence Skills; and WP = WritePlacer. N = sample size; k = number of studies; r_{obs} = sample-size-weighted average correlation; SD_{obs} = standard deviation of observed correlations; SD_{res} = residual standard deviation; $r_{(op)}$ = estimated validity for applied use of ACCUPLACER scores; ρ = estimated true score validity; SD_{ρ} = standard deviation of estimated true score correlations; 80 percent cred. = 80 percent credibility interval; percent CP = percentage of students correctly placed.

Table 3

Overall Relationship Between Combinations of ACCUPLACER Tests and Course Success

Test	*N*	*k*	r_{obs}	SD_{obs}	SD_{res}	$r_{(op)}$	ρ	SD_{ρ}	*80 percent cred.*	*Percent CP*
B or Higher										
EA + ARIT	868	6	.36	.10	.07	.61	.62	.00	.62–.62	66.6
EA + CLM	36	1	.62	-	-	-	-	-	-	81.0
RC + SS	5,371	5	.24	.02	.00	.38	.39	.00	.39–.39	59.5
RC + WP	2,857	4	.24	.03	.00	.44	.47	.00	.47–.47	59.9
SS + WP	1,704	1	.23	-	-	-	-	-	-	59.0
RC + SS + WP	1,693	1	.24	-	-	-	-	-	-	58.0
C or Higher										
EA + ARIT	512	5	.36	.16	.14	.64	.66	.09	.54–.78	79.0
EA + CLM	36	1	.58	-	-	-	-	-	-	81.0
RC + SS	5,668	7	.17	.06	.04	.28	.29	.02	.27–.31	73.5
RC + WP	2,690	3	.16	.02	.00	.29	.32	.00	.32–.32	72.4
SS + WP	1,704	1	.15	-	-	-	-	-	-	69.0
RC + SS + WP	1,693	1	.15	-	-	-	-	-	-	69.0

Note: ARIT = Arithmetic; EA = Elementary Algebra; CLM = College-Level Math; RC = Reading Comprehension; SS = Sentence Skills; and WP = WritePlacer. N = sample size; k = number of studies; r_{obs} = sample-size-weighted average correlation; SD_{obs} = standard deviation of observed correlations; SD_{res} = residual standard deviation; $r_{(op)}$ = estimated validity for applied use of ACCUPLACER scores; ρ = estimated true score validity; SD_{ρ} = standard deviation of estimated true score correlations; 80 percent cred. = 80 percent credibility interval; percent CP = percentage of students correctly placed.

However, after correcting correlations for statistical artifacts, the mean operational validity coefficient for individual ACCUPLACER tests was 0.42 when success was defined as obtaining a "B or higher," and 0.34 when success was defined as obtaining a "C or higher," suggesting a moderate relationship in general between ACCUPLACER scores and course success. Additionally, the mean correlation for the population estimates was slightly higher, with a value of 0.45 and 0.37 for "B or higher" and "C or higher," respectively. Where data were available, the same corrections were applied to combinations of tests. The mean operational validity for combinations of ACCUPLACER tests was 0.48 when success was defined as obtaining a "B or higher" and 0.40 when success was defined as obtaining a "C or higher," which supports a moderate-to-strong relationship between ACCUPLACER scores and course success. The mean correlation was slightly higher for the population estimates, with respective values of 0.49 and 0.42 for "B or higher" and "C or higher," respectively. These results indicate that there is a large effect for test combinations, especially when course success is defined as obtaining a "B or higher."

In sum, there was a moderate-to-strong relationship between test scores and subsequent course performance. However, there was variability in the magnitude of the relationship across tests and test combinations. Tests that included math content correlated more strongly with course success than tests that included verbal content. When course success was defined as obtaining a "B or better," the population estimates for the three math tests — Arithmetic, Elementary Algebra, and College-Level Mathematics — were 0.46, 0.50, and 0.62, respectively; whereas, for the three tests assessing verbal content — Reading Comprehension, Sentence Skills, and WritePlacer — the population estimates were 0.36, 0.36, and 0.42, respectively.

Discussion

This study supports the placement validity of ACCUPLACER scores as a measure for deciding the appropriate college course enrollment for students. The results of this study support a moderate-to-strong relationship between test scores and subsequent course performance. Furthermore, the percentage of students correctly placed was also quite high, with an average of 70 percent for all tests and combinations of tests, which provided additional support for the validity of ACCUPLACER test scores for placement purposes.

To assess the predictive validity of ACCUPLACER test scores for placement purposes, corrections for statistical artifacts of measurement error, range restriction, and unreliability were conducted, which provided an estimate of the true validity. Contrary to some past research (e.g., Armstrong, 2002), the results revealed a substantial relationship between placement test scores and course success. Operational validity was also computed and provided, which is the most practical finding for test users because this estimate was not corrected for predictor unreliability. Given that students are placed into courses with imperfect measures, this estimate provides information on the utility of using test scores for placement purposes.

Overall, this study provides valuable information about the placement validity of ACCUPLACER tests. Unlike an admission test, such as the SAT®, which has thousands of articles devoted to analyzing its psychometric quality (Camara & Echternacht, 2000), most notably its predictive validity (e.g., Bridgeman, McCamley-Jenkins, & Ervin, 2000; Hezlet et al., 2001), relatively little is known about placement tests, specifically ACCUPLACER tests. Furthermore, the studies that have examined the placement validity of various measures often fail to correct for statistical artifacts that attenuate the results. Thus, it seems plausible that if studies examining the predictive validity of the SAT only reported the uncorrected observed correlation, a similar modest relationship would be found. That is to say, the problem is not the tests used to make placement decisions but rather the methods used to validate their use and, specifically, not using methods to correct for the effects of statistical artifacts. By aggregating over multiple studies and correcting for statistical artifacts, the results of this study provide a more accurate representation of the validity of one commonly used placement testing system, ACCUPLACER.

In the current study, no moderator analyses such as by course content or difficulty were performed because the results for the total sample did not support these analyses. Generally, the standard deviations of the true validities were quite small. In fact, of the 18 analyses, half had a value of 0.00 for SDρ. Furthermore, the 80 percent credibility interval for all tests and test combinations were not very large and did not include zero, which also suggests the absence of moderating variables. When additional data become available, moderator analyses may be necessary; however, given the relatively small number of studies and low variability in true validities, additional analyses were not warranted.

There were three main limitations to this study, which restricted the analyses and generalizability of the results. First, the criterion "college success" was dichotomized, which eliminated meaningful variation in course performance and deflated estimates of the relationship between test scores and course success. If the criterion had been operationalized as course grade on a 100-point scale, or even a lettered grading scale, the results would have revealed a stronger relationship between test scores and course performance. Second, and most important, only 17 schools conducted an ACES ACCUPLACER placement validity study during the specified time period. In addition, 14 (82.4 percent) of these schools were two-year institutions, which may limit the applicability of the results to four-year institutions. Finally, only ACCUPLACER tests were examined in the current study; however, there are other tests on the market that are used nationally, such as ACT's COMPASS and ASSET, for placement into college courses. Thus, a larger investigation of the relationship between various test scores and course performance from a variety of institutions is necessary to extend these findings to other placement tests.

Future studies should aim to address these methodological limitations. In addition, future research should examine the validity of placement scores in predicting other dimensions of college success, such as retention and graduation. For example, researchers should investigate whether students who are placed correctly are more likely to stay at that institution and more likely to graduate than their counterparts. Additional research should also examine data at the institutional level to determine whether institutions that use placement testing have higher retention rates and graduation rates than institutions that do not use placement testing.

When additional data become available, analyses should also be conducted by course type to determine which test scores are appropriate for placement into which courses. It seems logical to use a test of reading comprehension for placement into an English course; however, further investigation is needed to determine utility for other courses that require a lot of reading for success, such as psychology. Finally, future studies need to determine whether test scores result in differential validity for various subgroups. For example, it is important to investigate whether placement decisions based on test scores are equally valid for males and females, and for students of different racial and ethnic groups.

References

Armstrong, W.B. (2000). The association among student success in courses, placement test scores, student background data, and instructor grading practices. *Community College Journal of Research and Practice, 24*, 681–695.

Barritt, L.S. (1966). Note: The consistency of first-semester college grade point average. *Journal of Educational Measurement, 3*, 261–262.

Bendig, A.W. (1953). The reliability of letter grades. *Educational and Psychological Measurement, 13*, 311–321.

Berkner, L., & Chavez, L. (1997). *Access to postsecondary education for the 1992 high school graduates* (Report No. NCES-98-105). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. (ERIC Document Reproduction Service No. ED413854)

Bridgeman, B. McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning Test* (College Board Research Report No. 2000-1). New York: The College Board.

Camara, W.J., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college* (College Board Research Report No. RN-10). New York: The College Board.

Cohen, A.M., & Brawer, F. (1987). *The collegiate function of community colleges*. San Francisco: Jossey-Bass.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

College Board. (2003). *ACCUPLACER OnLine Technical Manual*. New York: The College Board.

College Board. (2007). *ACCUPLACER*. Retrieved July 2, 2007, from http://professionals.collegeboard.com/higher-ed/placement/accuplacer

College Board. (2007). *ACES*. Retrieved July 2, 2007, from http://professionals.collegeboard.com/higher-ed/validity/aces

College of the Canyons. (1994). *Predictive validity of the APS writing and reading tests and validating placement rules for the APS writing test*. Santa Clarita, CA: College of the Canyons, Office of Institutional Development.

Gabe, L.C. (1989). *Relating college-level course performance to ASSET placement scores*. (Institutional Research Report Abstract RR 89–22). Fort Lauderdale, FL: Broward Community College.

Hargis, C.H. (1990). *Grades and grading practices: Obstacles to improving education and to helping at-risk students*. Springfield, IL: Charles C Thomas.

Hezlett, S.A., Kuncel, N., Vey, M.A., Ahart, A.M., Ones, D.S., Campbell, J.P., & Camara, W. (2001, April). *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle.

Hughes, R.E., & Nelson, C.H. (1991). Placement scores and placement practices: An empirical analysis. *Community College Review, 19*(1), 42–46.

Hunter, J.E., & Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications.

Hunter, J.E., & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousands Oaks, CA: Sage Publications.

Isonio, S. (1991). *Implementation and initial validation of the APS English writing test at Golden West College: Evidence for predictive validity*. Huntington Beach, CA: Golden West College.

Isonio, S. (1992). *Implementation and initial validation of the MDTP tests at Golden West College*. Huntington Beach, CA: Golden West College.

Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*, 148–161.

Moss, B., & Bordelon, S. (2007). Preparing students for college-level reading and writing: Implementing a rhetoric and writing class in the senior year. *Reading Research and Instruction*, *46*(3), 197–221

Rasor, R.A., & Barr, J. (1993). *Refinement in assessment validation: Technicalities of dealing with low correlations and instructor grading variation*. Sacramento, CA: American River College.

Reilly, R.R., & Warech, M.A. (1993). The validity and fairness of alternatives to cognitive tests. In L. C. Wing & B. R. Gifford (Eds.), *Policy issues in employment testing* (pp. 131–224). Boston: Kluwer.

Saunders, P.I. (2000). *Meeting the needs of entering students through appropriate placement in entry-level writing courses*. Forest Park, MO: Saint Louis Community College. (ERIC Document Reproduction Service No. ED447505)

Schmidt, F.L., & Le., H. (2004). Software for the Hunter-Schmidt meta-analysis methods. Iowa City, IA 42242. University of Iowa, Department of Management and Organizations.

Schuemann, C.M. (1998). *An investigation into the comparable effectiveness of two tests of writing administered to college-level English as a second language students: The computerized placement test and holistically scored essay exams*. Unpublished doctoral dissertation, Florida International University, Miami.

Smittle P. (1995). Academic performance predictors for community college student assessment. *Community College Review, 23*(2), 37–46.


CollegeBoard
inspiring minds™