DOCUMENT RESUME

ED 345 781 JC 920 261

AUTHOR Isonio, Steven
TITLE Implementation and Initial Validation of the APS English Test and The APS English-Writing Test at Golden West College: Evidence for Predictive Validity.
INSTITUTION Golden West Coll., Huntington Beach, Calif.
PUB DATE 91
NOTE 39p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Statistical Data (110)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College English; Community Colleges; Correlation; Grades (Scholastic); Predictive Measurement; Predictive Validity; Predictor Variables; *Reading Tests; *Student Evaluation; *Student Placement; Tables (Data); Test Results; Test Use; *Test Validity; Two Year Colleges; Writing (Composition); *Writing Tests
IDENTIFIERS Assessment and Placement Services for Comm Coll

ABSTRACT

In May 1991, Golden West College (California) conducted a validation study of the English portion of the Assessment and Placement Services for Community Colleges (APS), followed by a predictive validity study in July 1991. The initial study was designed to aid in the implementation of the new test at GWC by comparing data on APS use at other community colleges and by testing placement rules. The follow-up study examined the relationship between the APS and eventual course success as measured by students' grades. The initial sample for both studies consisted of 381 students from 23 classes offered in spring 1991 who had taken the reading and writing portions of the APS. Three standard criterion groups were identified, consisting of native English speakers placed in English 009, 010, or 100. Selected findings included the following: (1) the APS appeared able to differentiate among course levels and was significantly related to standard group membership, making a clear placement recommendation possible for the majority of students; (2) ratings of writing samples could aid in definitive placement of those students earning unclear APS scores; (3) the gain in predictability from using a combination of reading and writing score information over using writing alone was insignificant; (4) the analysis of the relationship of grades to APS scores revealed a low to moderate-sized predictive validity coefficient, a value within the expected range; and (5) a significant difference in APS scores was found between successful and unsuccessful students. Data tables, references and short appendixes, providing a sample student survey and a correction for restricted range of a predictive validity coefficient, are included. (JSP)

Implementation and Initial Validation of the APS English Test

and

The APS English-Writing Test at Golden West College: Evidence for Predictive Validity

GOLDEN WEST COLLEGE

1991

Steven Isonio, PhD

Advisor, Matriculation Research/Assessment

Implementation and Initial Validation of the APS English Test

GOLDEN WEST COLLEGE
May 1991

Steven Isonio, PhD
Advisor, Matriculation Research/Assessment

Implementation and Initial Validation of the APS English Test

A question and answer format is used to present the background, procedures and results of an assessment validation study conducted in Spring 1991. The primary intention of the study was to establish placement rules for the new test at Golden West College. An overview of necessary followup research is presented.

1. What steps are involved in the adoption and implementation of a new measure of writing skills for placing students into English courses at Golden West College?

The first steps, namely, recognizing the need for a better test, identifying a test for consideration, discussing its merits, and finally formally adopting it, are complete. A proposal to adopt the English portion of the Assessment and Placement Services for Community Colleges (APS), developed and distributed by the Educational Testing Service (ETS) was made in December 1990. The test was reviewed and recommended by a subcommittee of the GWC English department. In February 1991, the English Department full-time faculty voted 14-0 (with one abstention) to adopt the new test.

A pilot study was designed to gather information that would help with the establishment of placement rules as well as begin the process of documenting the relationship between the APS and relevant criteria. Such assessment validation research is necessarily ongoing. Adjustments in placement rules are likely, and an asssortment of validation studies are required.

2. Is the APS a better test than the Stanford TASK for placing students into courses at Golden West?

While both tests have face validity, there is a consensus that the APS taps a broader range of higher level skills than the TASK (Stanford Test of Academic Skills), thereby making it more appropriate for placing students into writing courses. Both instruments were thoroughly reviewed by the English department assessment subcommittee. This review resulted in a strong recommendation that the APS replace the TASK as the objective English placement test at Golden West. The TASK was characterized as "simplistic, concentrating on discrete items" and evaluating "a low level of ability". Conversely, the APS was deemed more appropriate in that it "features a higher level of writing and language skills and has a more global scope."

Further, unlike the TASK, the APS was developed specifically for placing of community college students into basic skills courses. Additionally, the APS is now the most widely used instrument for placing students into English courses in California Community Colleges.

3. What guidelines for use of the APS are offered by the test publisher?

ETS provides extensive information about the construction and norming of the test in the technical manual (College Entrance Examination Board, 1990). References to pertinent research are also provided. Additionally, the manual contains a discussion of selected psychometric properties of the test. However, this instrument, like any placement test, does not come with universally applicable placement rules. Local placement rules must be developed.

4. What type of placement rules are used at other schools that have adopted the APS?

Information about how the instrument is used at a sample of other colleges was gathered. This information suggested models that would be tested as a part of the pilot study. Table 1 presents placement rules implemented at a sample of four colleges that use the APS for placing students into three levels of English. To facilitate comparison, course numbers for the transferable composition course have been changed to "100"; similarly, the two courses below this level are referred to as "10" and "9", respectively, although these may not be the actual course numbers at the other colleges.

Two schools, American River College and Rancho Santiago College use only the objective writing section of the APS to place students directly into the three courses. Moorpark College and Santa Barbara City College incorporate information from an essay to clarify placements in certain marginal ranges. Specifically, in the case of Moorpark College, students who scored in the upper portion of the range for placement into the English 10 course can challenge this placement by having their essay evaluated. The student's placement is either changed to the English 100 course, or remains English 10. Santa Barbara City College routinely evaluates essays written by students in either of three ranges: those students in a band between English 10 and English 100, those students in a band between English 9 and English 10, and those students scoring in the very low ranges of the distribution.

There appears to be much consistency in placement rules used by these schools. This information was useful in helping to generate reasonable placement rules for consideration in the study.

Table 1

Three-tiered Placement Rules
With APS Writing Test at Sample Colleges

Score	MOORPARK	AM.RIV.	SANTA BARBARA	RANCHO SANTIAGO
40	100	100	100	100
39	100	100	100	100
38	100	100	100	100
37	100	100	100	100
36	100	100	100	100
36	100	100	100	100
35	100	100	100	100
33	100	100	100	100
32	100	100	100	100
31	100	100	10 OR 100	100
30	100	100	10 OR 100	100
29	100	100	10 OR 100	10
28	100	100	10	10
27	10+	10	10	10
26	10+	10	10	10
25	10+	10	10	10
24	10+	10	10	10
23	10	10	10	9
22	10	10	10 OR 9	9
21	10	10	10 OR 9	9
20	10	9	10 OR 9	9
19	10	9	10 OR 9	9
18	10	9	9	9
17	10	9	9	9
16	9	9	9	9
15	9	9	9	9
14	9	9	9 OR LOWER	9
13	9	9	9 OR LOWER	9
12	9	9	9 OR LOWER	9
11	9	9	9 OR LOWER	9
10	9	9	9 OR LOWER	9
9	9	9	9 OR LOWER	9
8	9	9	9 OR LOWER	9
7	9	9	9 OR LOWER	9
6	9	9	9 OR LOWER	9
5	9	9	9 OR LOWER	9
4	9	9	9 OR LOWER	9
3	9	9	9 OR LOWER	9
2	9	9	9 OR LOWER	9
1	9	9	9 OR LOWER	9

Note. In cases where two placements are indicated, definitive placement is made based upon an evaluation of the essay. "10+" indicates placement into the English 10 level course, but students may request that their essay be read for possible placement into the English 100 course.

5. Describe the design and procedures of the pilot study. What was expected?

In the Spring 1991 semester, a representative sample of 23 English 9, 10, and 100 classes was identifed. A total of 381 students from these classes were administered the reading and writing sections of the APS (see Table 2). Additionally, information about the students' placement based upon the TASK test (previously used objective English placement test), a writing sample produced the first day of instruction, and instructors' recommendations was gathered.

Three "standard" criterion groups were identified. These groups consisted of native English speakers who were placed into English 9, 10, or 100 and for whom these indicators were in agreement. Certainly this composite criterion is not perfect, but its reliability was enhanced by the convergence of these various indicators. Finally, only native speakers of English were included in the analysis because the APS was designed specifically for placement of such students. Primary language was determined from a supplemental survey administered after the test (see Appendix A).

Distributions of scores were obtained for the reading and writing sections of APS, and their combination. The relationship between APS scores and course level was determined with the multiserial correlation coefficient (Jaspen, 1946). To the extent that a relationship exists between the APS scores and standard group membership, the APS can be said to have criterion-related validity. Further, the central points of the three distributions should be at different locations along the abscissa, ideally with a only a moderate degree of overlap.

Finally, a variety of placement rules were evaluated. The ability of these various rules to differentiate among the three groups (maximizing correct "placements" while minimizing incorrect ones) was assessed. The score ranges for the most effective model would constitute the recommended initial cut scores.

6. Is the APS capable of differentiating the three levels of English courses?

The process of implementing a test and establishing placement rules entails comparing it to a standard. The difficulty arises in that there is no ideal criterion to serve as the standard. Therefore, as noted above, existing information (placement based on the previous English placement test (TASK), ratings of the first day placement writing samples (PWS), and students' self-evaluation implicit in their course selection were used to define three standard comparison groups.

At a minimum, APS mean scores for the three standard groups should be clearly different from each other. The standard groups do indeed differ significantly in terms of mean APS writing score [F (2,187) = 79.09, $p < .001$]. Figure 1a depicts these group mean differences. The reading and writing scores were combined

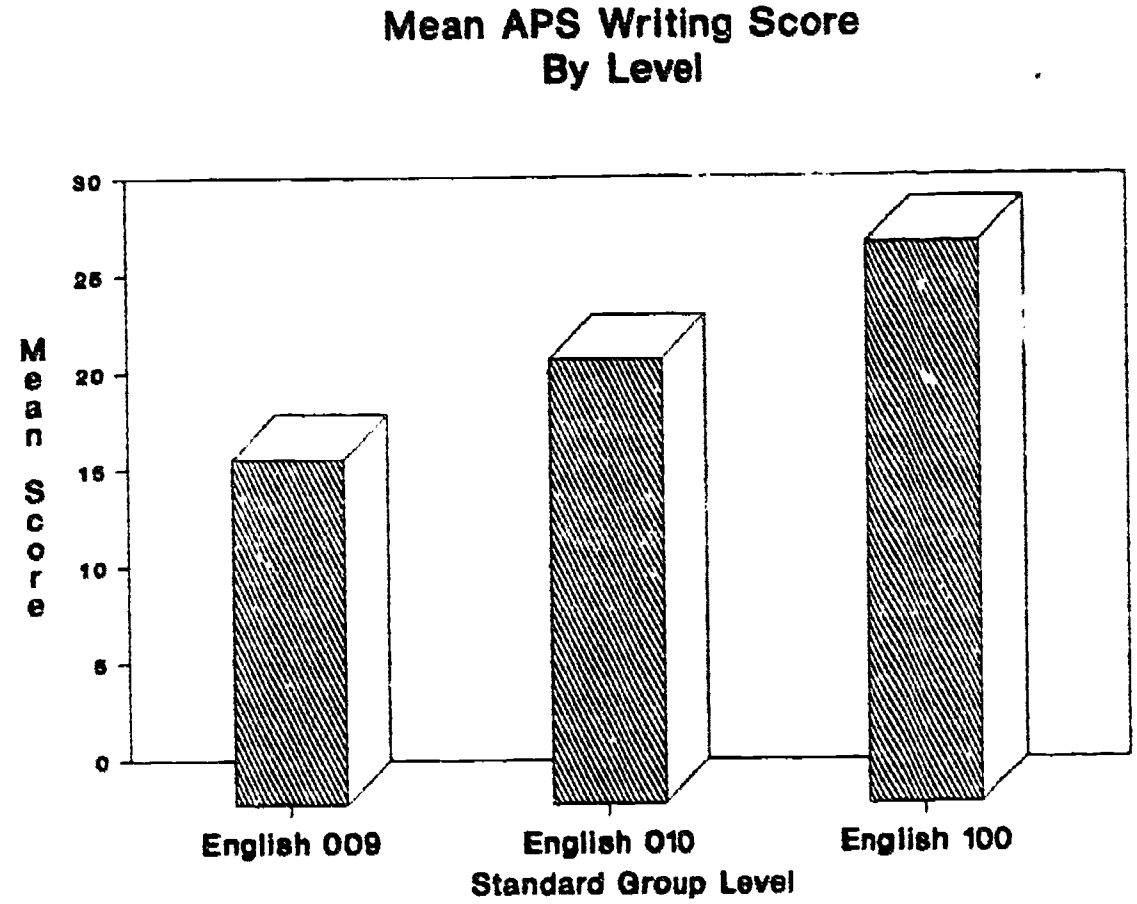
Table 2

Number of Students Tested by English Course Level

Course Level	Number of Classes	Students Tested	Students in the Analysis*
English 9	4	54	44
English 10	8	156	60
English 100	11	171	87
Total	23	381	191

*Only Native Speakers of English who were placed into their enrolled class with the TASK-EPT and the Placement Writing Sample were included in the analysis. These students comprise the standard criterion groups.

Figure 1a



APS Writing Scores

Descriptive Statistics

Level	Mean	Standard Deviation	Range
English 009	17.82	5.13	7 - 27
English 010	22.92	4.87	12 - 33
English 100	28.87	4.75	15 - 38

by summing the two values for each student (Schmidt, 1971). The mean combined score also differed significantly across the three groups [F (2,185) = 94.50, $p < .001$]. These means are presented in Figure 1b. Thus, for both the mean APS writing score alone and in combination with the reading score, there are very striking differences among the three groups.

Differences among the group means such as those just described, however, indicate only that the central points of the three distributions differ. Therefore, it was necessary to examine the full distributions of scores to determine the extent to which the distributions overlap. Figures 2a, 2b, and 2c depict the APS writing score distributions for English 009, 010, and 100, respectively. The ordinate was scaled to reflect relative frequency to facilitate direct comparisons across the three different-sized samples. The striking differences among group mean writing scores noted above, are reflected by the differences among the distributions. For the English 009 sample, the scores ranged from 7 to 27, whereas the ranges were 12 to 33 for English 10, and 20 to 38 for English 100.

The differences are also reflected in the semi-interquartile ranges, the middle 50% of the distribution. These ranges are 15 to 21 for English 9, 20 to 27 for English 10, and 27 to 32 for English 100. The advantage of the semi-interquartile range is that it reflects the "core" of the distribution, excluding extreme scores which may be attributable to very low motivation or luck. Therefore, these values are particularly stable.

Similarly, Figures 3a, 3b, 3c depict the distributions of combined reading and writing scores, for students in the standard English 9, 10 and 100 groups, respectively. The central points of the distributions are clearly located at different places on the APS score axis. However, there appears to be slightly more overlap between distributions for adjacent course levels. The semi-interquartile range for English 9 is 25-38. For English 10 this value is 33-50; for English 100 the range is 48-61.

In short, the APS writing section alone appears somewhat able to differentiate among the three levels of English. The combination of writing and reading scores seems slightly less able to do so.

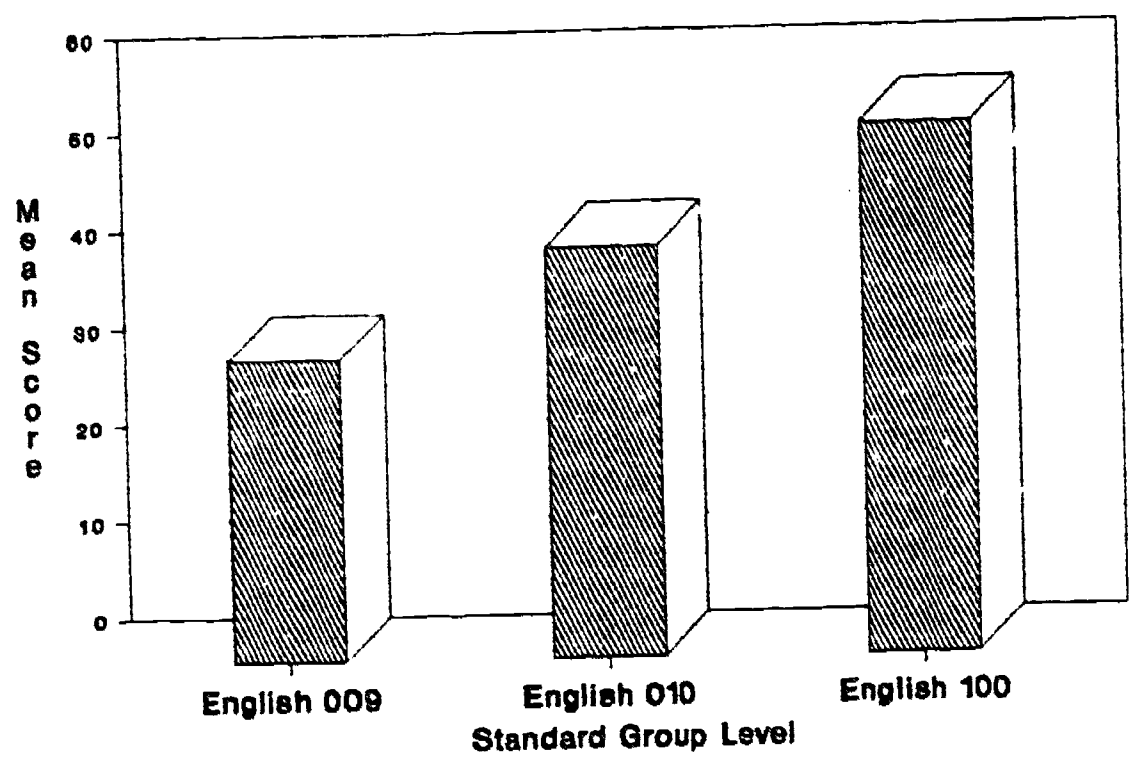
7. How can a writing sample be best utilized in conjunction with APS test score information (that is, maximizing accuracy of placement while maintaining efficiency)?

As noted earlier, there is no consensus among the sample APS schools regarding the use of writing samples. Some schools do not use writing samples. Other schools use them to supplement the APS scores, typically to make definitive placement recommendations for students in marginal ranges.

Since the APS appears able to differentiate among course levels and is significantly related to standard group membership, a clear placement recommendation is possible for the majority of students. Ratings of writing samples can contribute to a definitive placement for those students who earn APS scores that

Figure 1b

Mean APS Writing Plus Reading Score By Level



APS Pilot, Spring 1991

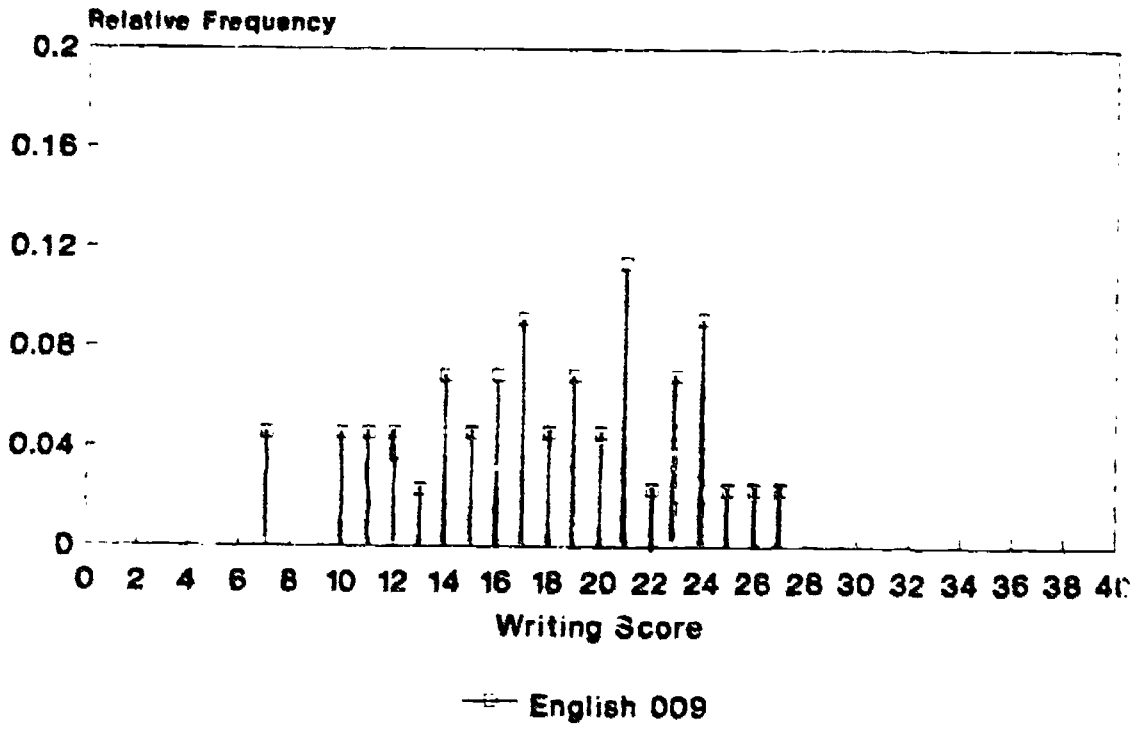
APS Writing Plus Reading Scores

Descriptive Statistics

Level	Mean	Standard Deviation	Range
English 009	31.21	9.46	11 - 49
English 010	42.33	10.59	19 - 63
English 100	54.51	7.92	33 - 69

Figure 2a

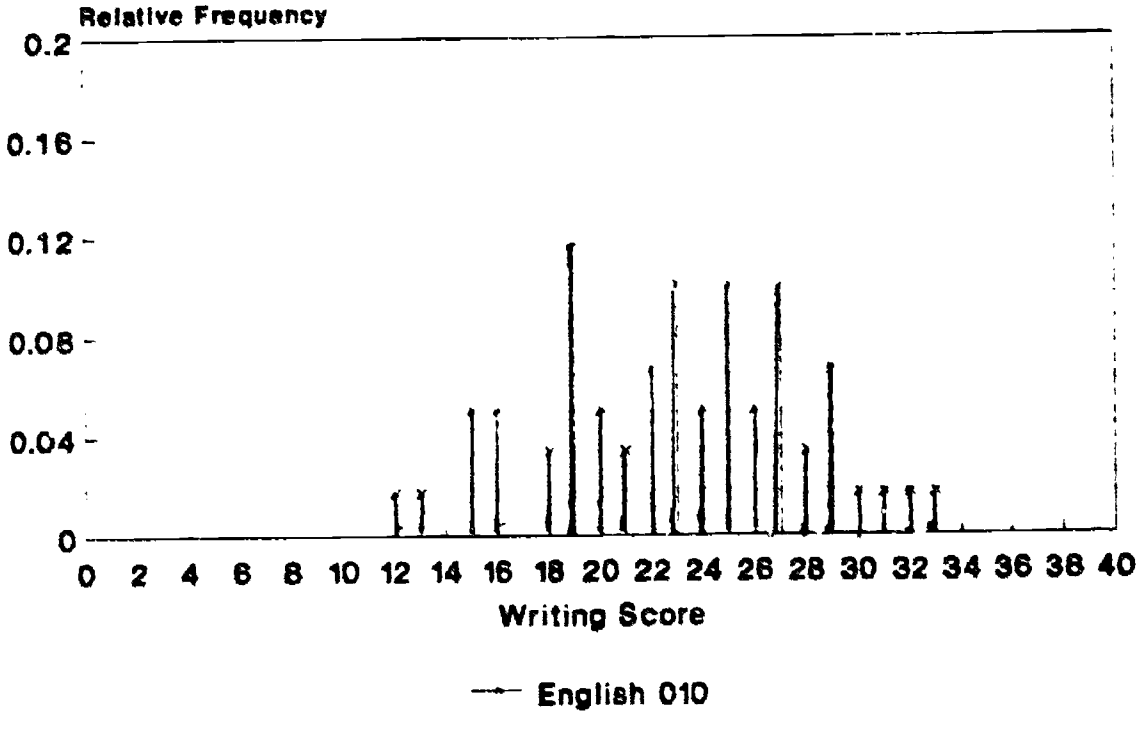
APS Scores: Writing Only (English 9 Standard Group)



Spring 1991 Pilot

Figure 2b

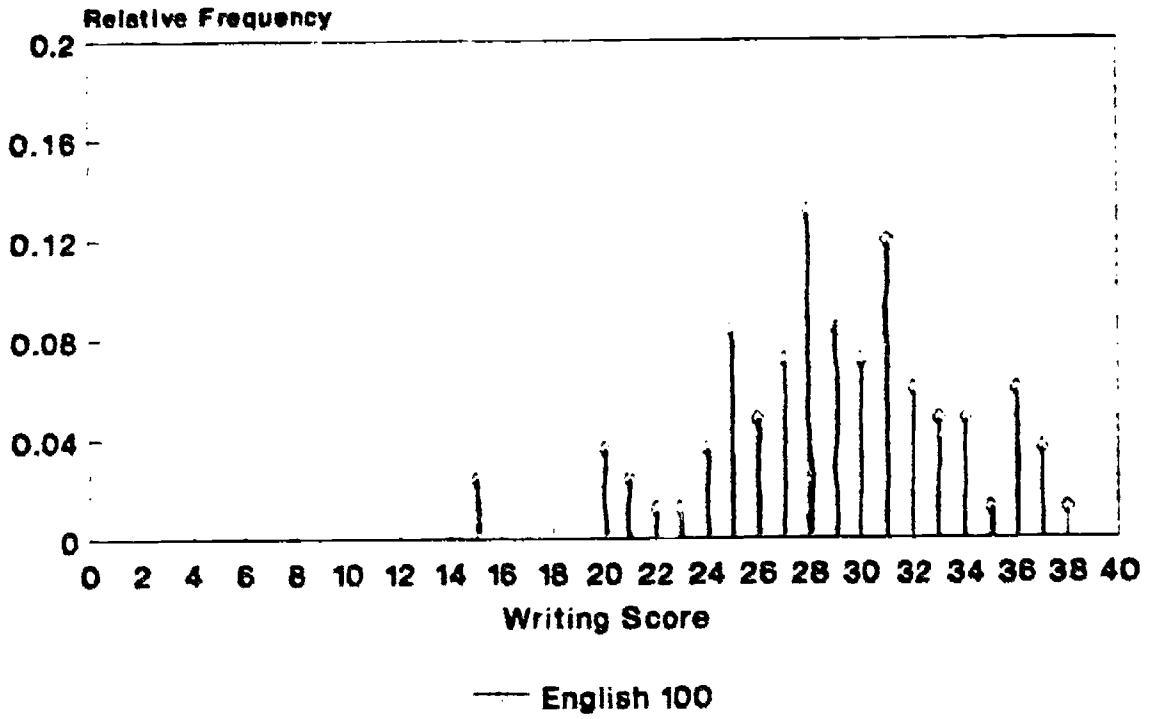
APS Scores: Writing Only
(English 10 Standard Group)



Spring 1991 Pilot

Figure 2c

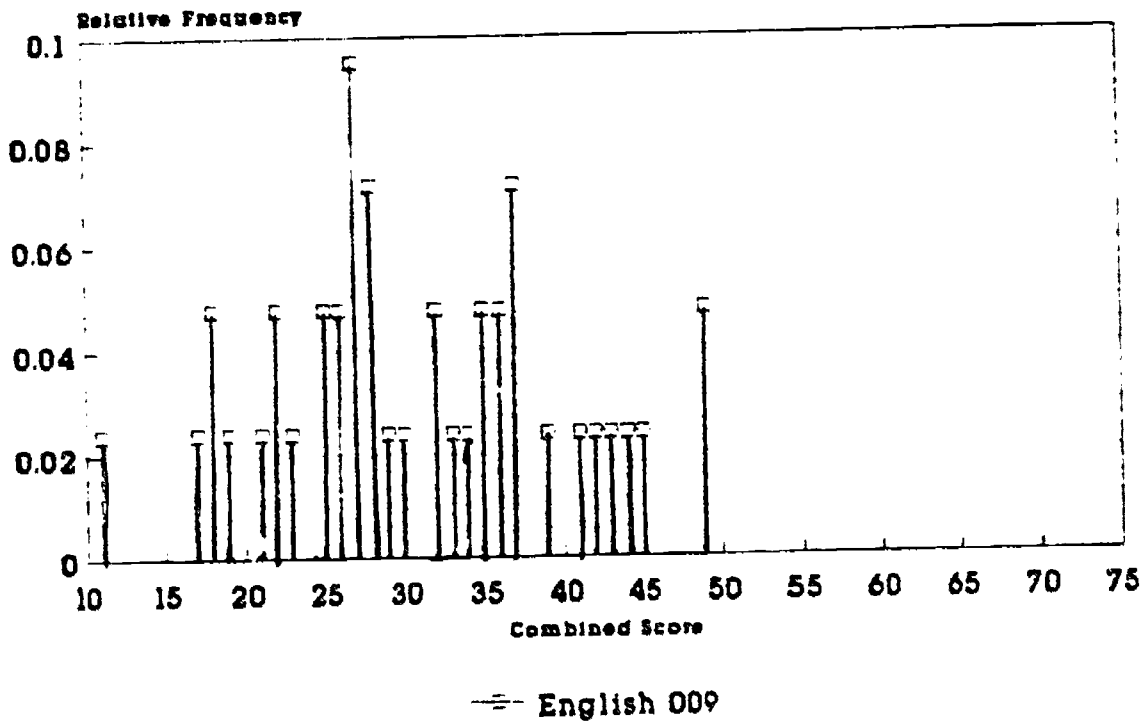
APS Scores: Writing Only
(English 100 Standard Group)



Spring 1991 Pilot

Figure 3a

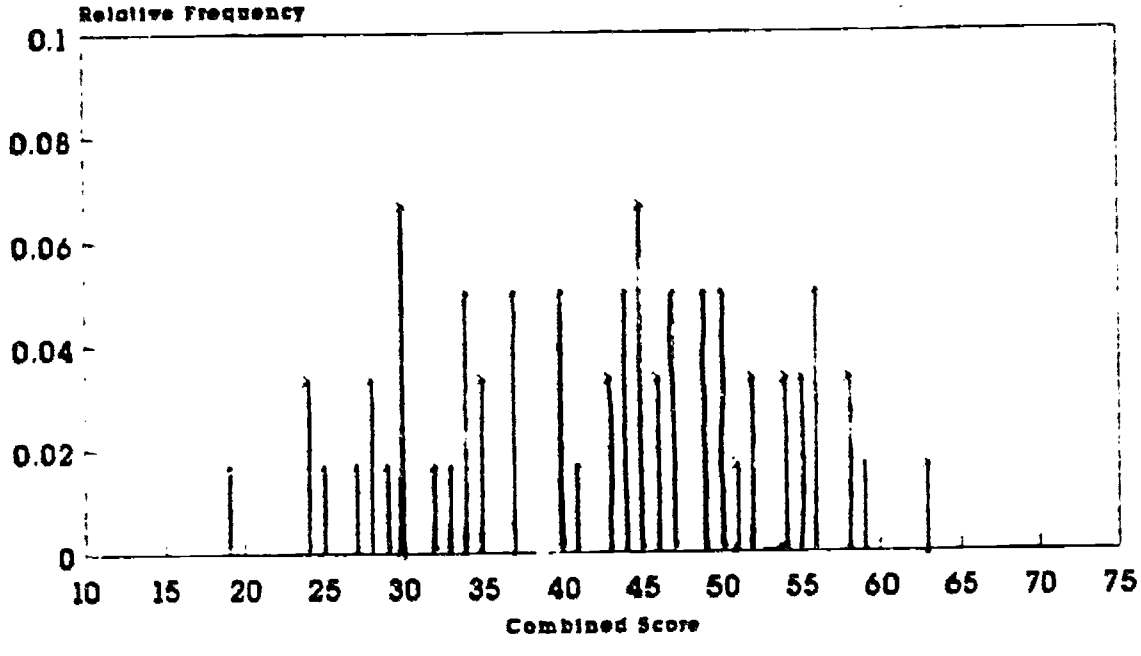
APS Scores: Reading plus Writing (English 9 Standard Group)



Spring 1991 Pilot

Figure 3b

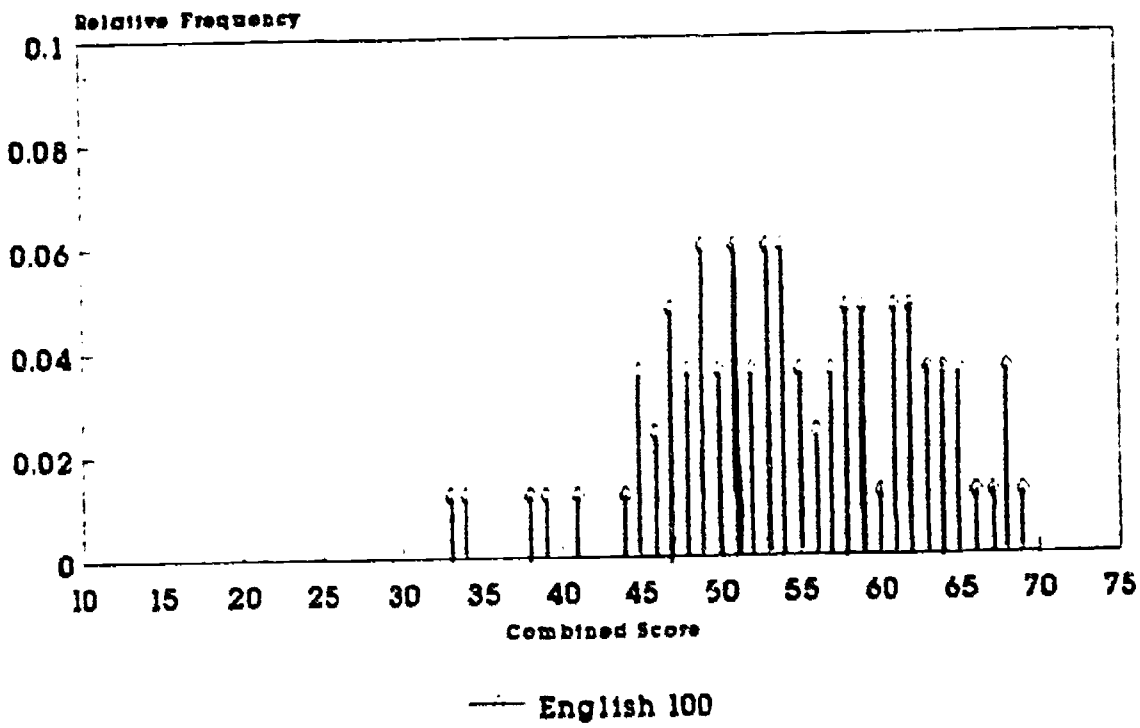
APS Scores: Reading plus Writing (English 10 Standard Group)



Spring 1991 Pilot

Figure 3c

APS Scores: Reading plus Writing (English 100 Standard Group)



Spring 1991 Pilot

do not suggest a clear placement. Therefore, the placement rules that were evaluated as a part of this study all entail using writing sample information only for students in such marginal ranges. Once again, the goal was to maximize both accuracy and efficiency.

8. Should information from both the reading and writing sections be combined in the placement rule?

This question involves the relative ability of the writing section alone compared with both the writing section and the reading section to predict standard group membership. Relevant evidence entails assessing the relationship between these APS scores and course level. The multiserial correlation coefficient (Jaspen, 1946) is appropriate for this case. The relationship between APS writing score and course level was significant [r (185) = .67, $p < .001$). The combination of reading and writing scores was also significantly related to course level [r (185) = .71, $p < .001$]. The difference between the two values was not significant (Z = .38, p ns). That is, the gain in predictability from using a combination of reading and writing score information over using writing alone is nonsignificant.

The test is to be used for recommending placement into either English 9 (Beginning Writing), English 10 (Writing Essentials), or English 100 (Freshman Composition). Therefore, the writing test is clearly more appropriate in terms of its content. A content validation study to document this obvious fact is currently being planned. Further, GWC does not have a comprehensive reading program into which to place students based upon their reading scores.

9. Based upon the data obtained from the standard groups, which APS score ranges are most closely associated with the English 009, 010, and 100 levels?

Since the distributions do overlap, there are some APS score values that would indicate placement into more than one of the levels. The challenge is to set specific cuts for "placing" students into the levels so that the proportion of "correct placements" are maximized. Three-tiered placement rule systems in use at other colleges helped guide the process. Three alternative candidate models for placement rules based upon scores on the writing section, as well as three that entail a combination of the writing and reading sections were considered.

The top part of Tables 3a through 3c present the three placement rules that rely only upon writing scores. Each scheme entails a two-point band between adjacent placement ranges, and a bottom range, from which the writing samples will be read. Since they are based on placement models in use at other schools and upon an examination of the actual distribution, differences among these schemes are slight--usually with comparable ranges varying by no more than one point. The bottom portion of each Table

Table 3a

Placement Rule #1 (Writing Section Only):
Percent Within Each Range and
Percent Identified

Writing Score Range	Placement	Percent of Students
>= 29	English 100	28.3%
27-28	Essay range	13.6%
22-26	English 10	25.1%
20-21	Essay range	8.9%
15-19	English 9	16.2%
<= 14	Essay range	7.9%

	Course Level			
	E009	E010	E100	Mean
Direct	31.8%	36.7%	52.9%	40.5%
Direct + Range	75.0%	58.3%	72.4%	68.6%

Table 3b

Placement Rule #2 (Writing Section Only):
Percent Within Each Range and
Percent Identified

Writing Score Range	Placement	Percent of Students
>= 29	English 100	28.3%
27-28	Essay range	13.6%
21-26	English 10	29.8%
19-20	Essay range	9.4%
14-18	English 9	12.6%
<= 13	Essay range	6.3%

	Course Level			
	E009	E010	E100	Mean
Direct	31.8%	40.0%	52.9%	41.6%
Direct + Range	63.7%	70.0%	72.4%	68.7%

Table 3c

Placement Rule #3 (Writing Section Only):
Percent Within Each Range and
Percent Identified

Writing Score Range	Placement	Percent of Students
>= 30	English 100	22.5%
28-29	Essay range	12.6%
22-27	English 10	31.9%
20-21	Essay range	8.9%
15-19	English 9	16.2%
<= 14	Essay range	7.9%

	Course Level			
	E009	E010	E100	Mean
Direct	31.8%	46.7%	44.8%	41.1%
Direct + Range	75.0%	65.0%	65.5%	68.5%

presents the proportions of correct identifications of standard group membership for English 9, 10, and 100. The row labeled "Direct" contains the percentage of students correctly identified by the specific placement range. The "Direct + Range" row presents the percentage of students properly identified by the specific placement range and the adjacent band. The means for both of these are presented in the far right column. Finally, the percentage of students that would have received each of the various placements is also presented.

Since the three schemes are very similar, their accuracy levels are also very similar. Nevertheless, there is some indication that placement rules two and three (Tables 3b and 3c) are superior to placement rule one. With slightly fewer writing samples being read, these two models maintain an accuracy rate in the high 60% range. Placement rule three (Table 3c) maintains the standard of a minimum score of 30 for direct placement into English 100, and since it may be best to have more stringent criteria at the outset, this model may be preferred as an initial scheme. While placement rule one has a high level of accuracy of identifying English 9 and English 100 students, it is somewhat less able to identify English 10 students.

Three models that combine writing and reading scores were also examined. Tables 4a through 4c present these models and data concerning their accuracy. Combining these two scores produces a much broader range of possible scores and a greater degree of variability among the scores. As such the essay range was extended to 3 score units and the direct placement ranges were also broadened. The proportions of correct identifications for the three models are all lower than for the models involving only the writing scores. Clearly no significant increase in accuracy of identification of course level results from combining reading and writing score information. This fact, coupled with the nonsignificant contribution in predictability described above suggests that placement rule three is still the preferred model.

10. What are the next steps in the implementation/validation process?

It is important to understand that validation of an instrument is necessarily a multifaceted, continual process. Further, validity is not an inherent attribute of an instrument; rather, it is a characteristic of a test used in a particular way at a particular place and at a given time. Thus, the validation process never truly ends. The Matriculation Local Research Options Committee consisting of representatives of the major California Community College research organizations, under the direction of the Chancellor's office, has very recently published and distributed a document entitled "Assessment Validation Project Local Research Options" that details eleven assessment research designs intended to meet the minimum standards for the evaluation of assessment instruments used at community colleges. These designs, therefore, essentially constitute the requirements for a <u>minimally acceptable</u> agenda for assessment research.

Table 4a

Placement Rule #4 (Writing Plus Reading):
Percent Within Each Range and
Percent Identified

Writing + Reading Score Range	Placement	Percent of Students
>= 52	English 100	34.5%
49-51	Essay range	8.9%
36-48	English 10	25.1%
33-35	Essay range	3.7%
23-32	English 9	14.7%
<= 22	Essay range	13.1%

	Course Level			
	E009	E010	E100	Mean
Direct	36.4%	35.0%	60.9%	44.1%
Direct + Range	68.2%	46.7%	72.4%	62.4%

Table 4b

Placement Rule #5 (Writing Plus Reading):
Percent Within Each Range and
Percent Identified

Writing + Reading Score Range	Placement	Percent of Students
>= 51	English 100	38.2%
48-50	Essay range	4.7%
37-47	English 10	22.5%
34-36	Essay range	3.7%
23-33	English 9	16.2%
<= 22	Essay range	14.7%

	Course Level			
	E009	E010	E100	Mean
Direct	38.6%	35.0%	66.7%	46.8%
Direct + Range	77.2%	45.0%	73.6%	65.3%

Table 4c

Placement Rule #6 (Writing Plus Reading):
Percent Within Each Range and
Percent Identified

Writing + Reading Score Range	Placement	Percent of Students
>= 52	English 100	34.6%
49-51	Essay range	8.9%
38-48	English 10	20.9%
35-37	Essay range	5.2%
25-34	English 9	17.3%
<= 24	Essay range	13.1%

	Course Level			
	E009	E010	E100	Mean
Direct	38.6%	30.0%	60.9%	43.2%
Direct + Range	79.5%	45.0%	72.4%	65.6%

Perhaps the primary requirement for validity information involves the need to evaluate the capability of the test to predict success in class. This entails evaluating the overall relationship between test scores and a criterion (usually course grades). This extension of the present study will be possible when the Spring 1991 grades are available. The form of that phase of the study will be similar to Assessment Validation Design 10 (specific designs are documented in the Assessment Validation Project Local Research Options). Validation of the placement rules, an issue separate from predictive validity, is also required (Assessment Validation Design 11). This entails gathering evidence that the particular cutscores result in a greater proportion of "hits" (correct predictions) and fewer "misses" (incorrect predictions) than other possible cutscores. Additionally, disproportionate impact and differential prediction, both which involve the issue of fairness, must be assessed (Assessment Validation Designs 12 and 13). The basic question in these case is whether the test is systematically biased against certain key subgroups of test takers.

Another source of validity evidence is students' and instructors' perceptions of the accuracy of placements (Assessment Validation Design 14). This involves conducting a survey near the midpoint of the semester to measure the perceived appropriateness of placements. An advantage of this approach is that it avoids some of the difficulties arising from instructor variability in grading (unreliability of the criterion) and contamination arising from a full semester of instruction. Of course, placement rates must be monitored and regularly reported.

SUMMARY:

1. In February 1991, the GWC English Department voted to adopt the APS as the objective test to be used for placing students into English 9, 10, and 100. The vote was 14 favoring adoption with 0 opposing. This followed a few months of discussions and reviews of the test.

2. A study was conducted to provide a foundation for implementation decisions and to obtain preliminary validation data. Students in 23 Spring 1991 semester classes provided data.

3. A sample of colleges that currently use the APS for placing students into a similar course structure were surveyed. Information regarding placement rules in use at these schools was gathered. There is a high degree of correspondence among the schools.

4. A number of candidate placement rules were tested. The primary comparison involved the rate of accurate identification of standard group membership. The percentage of students that would have received each of the various placement recommendations was also determined.

5. A model for advisory placement is recommended. It entails using information from the writing section of the APS for placing students into English 9, 10, and 100. For certain ranges where the APS placement is unclear, information from placement writing samples should be used to make a definitive placement.

6. The results are only preliminary. Ongoing monitoring of placement rates and the results of future validation studies may suggest modifications in the placement rule. All placement recommendations are advisory.

RECOMMENDATIONS:

1. Initial rules for placement of students into English 9, 10, and 100 should be based upon the writing section of the APS.

2. In cases where the APS is unable to make a definitive placement, information from writing samples produced as a part of placement testing should be used to make a clear placement. This approach maintains accuracy of placement without sacrificing efficiency.

3. The initial placement rule should be:

Score range	Placement
>= 30	English 100
28-29	Essay range
22-27	English 10
20-21	Essay range
15-19	English 9
<= 14	Essay range

4. The study should be extended to include course grades when this information becomes available. This would constitute predictive validity information that is required by matriculation regulations.

5. Placement rates should be monitored and regularly reported. Modifications in the placement rule should be made if evidence suggests that they are needed.

6. The content validity study currently in the early planning stages should be undertaken.

7. Studies of disproportionate impact and differential validity, specifically required by matriculation research regulations, should be undertaken.

References

Assessment Validation Project Local Research Options (1991). Developed by the Matriculation Local Research Options Committee. February, 1991.

College Entrance Examination Board (1990). Assessment and Placement Services for Community Colleges--Using and Interpreting Scores.

Jaspen, W. (1946). Serial correlation. Psychometrika, 11, 23-30.

Schmidt, F. L. (1971). The relative efficiency of regression and simple unit predictor weights in applied differential psychology. Educational and Psychological Measurement, 31, 699-714.

Appendix

SECTION #_____

ASSESSMENT & PLACEMENT SERVICES: PILOT STUDY
Student Survey (Short Form)

Please indicate your response to each of the following questions by circling the letter of your choice.

1. How many years of English did you complete in High School?

   a. less than one
   b. one
   c. two
   d. three
   e. four

2. How many English courses have you completed in college?

   a. none
   b. one
   c. two
   d. three
   e. four or more

3. What grade did you receive in the last English class you completed?

   a. A
   b. B
   c. C
   d. D
   e. Withdrawal or incomplete
   f. F

4. What is your predominant racial/ethnic background?

   a. Alaskan/Native American
   b. Asian
   c. Black
   d. Filipino
   e. Hispanic
   f. Pacific Islander
   g. White
   h. Other: _____

5. What is the first (primary) language that you learned?

   a. English
   b. Chinese
   c. Farsi
   d. Japanese
   e. Spanish
   f. Vietnamese
   g. Other: _____

6. What language is typically spoken in your home?

   a. English
   b. Chinese
   c. Farsi
   d. Japanese
   e. Spanish
   f. Vietnamese
   g. Other: _____

Your Social Security Number: ☐☐☐-☐☐-☐☐☐☐

The APS English-Writing Test at Golden West College: Evidence for Predictive Validity

GOLDEN WEST COLLEGE
July 1991

Steven Isonio, PhD
Advisor, Matriculation Research/Assessment

The APS English-Writing Test at Golden West College: Evidence for Predictive Validity

Steven Isonio, PhD
Advisor, Matriculation Research/Assessment

Background

The typical criterion by which placement instruments are judged is their relationship to course grades. That is, course grades usually serve as the yardstick against which placement tests are compared. If there is at least a moderately-sized positive relationship between test scores and course grades, then the test is said to have predictive validity.

An implementation study described elsewhere (Isonio, 1991), recommended an initial placement rule for use with the APS writing test (College Entrance Examination Board, 1990). That study was extended by incorporating course grades earned by English 100 students in the sample, thereby providing a means to assess the predictive validity of the test.

Method

The writing portion of the APS was previously administered to students in a sample of English 009, 010, and 100 classes early the Spring 1991 semester. An initial placement model based primarily upon APS scores was recommended and adopted. The next phase of the study, reported here, entailed determining the predictive validity of the test. An overview of this design (Design 10) is presented in the Local Assessment Validation Research Options (1991). The design simply involves determining the correlation between test scores and course grades.

The correlation between the APS Writing scores and course grades was computed for native English-speaking students in English 100 classes in the pilot sample. (The test was developed and normed on native English speakers, and it will be used with this population at Golden West College.) For this analysis, course grades were compiled from the District database. Each student in the sample was identified and his/her course grade was recorded. Grades were assigned the traditional point values (A=4.0, B=3.0, etc.). Withdrawals (W's) were not included in the analysis because of the difficulty interpreting their meaning. Further, since the distribution of scores was restricted because students had previously been placed into classes by another instrument, it was appropriate to use the statistical correction for restriction of range. This correction yields a better indication of the relationship between scores and grades by adjusting the value of the correlation coefficient to what it would be if the distribution of the predictor (APS test score) was unrestricted. (The formula for the correction for restricted range is presented in Appendix 1, along with the calculations involved in the present analysis).

Results

The primary analysis involved the predictive validity of the APS. A Pearson correlation coefficient, corrected for restriction of range, indicated that there is a moderate linear relationship between the placement test scores and course grades [r-corrected (129) = .29, p < .01]. Descriptive statistics for distributions of APS test scores attained by students earning various course grades are presented in Table 1.

A one-way analysis of variance was performed to test for differences in performance on the APS as a function of the grade earned in the course. This analysis indicated that there are significant differences among these groups [F (4,114) = 2.66, p < .05]. Tukey's HSD was used for multiple comparisons to isolate specific significant differences between pairs of means. This revealed that the only significant pairwise comparisons involved the three passing grades (A, B, and C) with the grade of F. A summary of the results of these comparisons is presented in Table 2.

A similar question, namely, whether there is a difference between the APS scores of successful and unsuccessful students was also addressed. In this case, a t-test was used to compare scores for students earning either grades of A, B, or C (successful) with those for students earning grades of D or F (unsuccessful students). The difference between these groups is significant [t (1, 117) = 6.32, p < .01].

Discussion

While the predictive validity coefficient value of .29 may appear somewhat small, it is statistically significant. Statistical significance, an index of the reliability of the finding, has been put forth as a minimal criterion for predictive validity (Assessment Validation Local Research Options, 1991). Further, the magnitude of the coefficient is within the range reasonably expected for placement tests. There is a broad range of factors that are not, and can not, be tapped by the test, but which are related to performance in class. Among these are an assortment of individual difference variables, motivation level, and differences in background and experiences. Because there is such a complex array of factors that interact to determine performance in classes, many of which are non-academic, it is not reasonable to expect placement tests to be very strongly related to course grades. Indeed, recent interpretations of minimal standards suggest that predictive validity coefficients of about .30 are adequate. Early informal reports of predictive validity studies at other community colleges in the state suggest that values as low as .00 to .20 are somewhat common.

While the obtained value of .29 indicates that the predictor (APS scores) accounts for 8.4% (the coefficient of determination, r-squared = .084) of the variance in the criterion, this is very likely an underestimate of the strength of the relationship between them. As such, there may be a tendency to

Table 1

Mean and Standard Deviation for APS-Writing Score and Grades Earned by English 100 Students in the Pilot Sample

	APS Scores		
Grade Earned	Mean	Standard Deviation	n
A	29.06	4.99	18
B	26.98	4.44	52
C	26.87	4.62	30
D	25.75	5.85	8
F	23.36	4.41	11

Table 2

Tukey-HSD Test for Pairwise Comparisons Between Means

Differences Among APS Mean Scores By Grade Earned:

	A	B	C	D	F
A	---				
B	2.08	---			
C	2.19	0.11	---		
D	3.31	1.23	1.12	---	
F	5.70**	3.62*	3.51*	2.39	---

* $p < .05$

** $p < .01$

Note. Group means are presented in Table 1.

inappropriately underrate the value of the APS as a placement tool. The reason that .29 probably underestimates the true strength of the relationship is that the statistical relationship between any two independent measures (e.g., placement test scores and course grades) is a function of an assortment of factors. In addition to the actual relationship between the underlying constructs tapped by the two measures, the empirically derived correlation also reflects the separate variabilities of the two constructs, as well as the reliabilities of the measures of the two constructs. The measure of the criterion in this case is not perfectly reliable, and both the distributions for APS scores and grades are restricted in range. These factors constrain the possible value for the predictive validity coefficient.

Summary and Conclusions

Course grades for English 100 students who participated in the APS implementation study (described elsewhere) were compiled and analyzed. Specifically, their relationship to scores on the APS writing section was determined. The analyses revealed a low to moderate-sized predictive validity coefficient, a value within the range to be reasonably expected for such tests. Further, a significant difference in APS scores was found between successful and unsuccessful students.

In light of the myriad of factors that impact course grades other than the specific skills measured by the test, and the unreliability of the criterion, the obtained value for the predictive validity of the APS is adequate to justify use of the test for placing students into writing courses at Golden West College.

References

Assessment Validation Local Research Options (1991). Handbook developed by the Matriculation Local Research Options Committee. February, 1991.

Collge Entrance Examination Board (1990). Assessment and Placement Services for Community Colleges--Using and Interpreting Scores.

Isonio, S. (1991). Implementation and initial validation of the APS English Test. Golden West College, Assessment & Research. May 1991.

Appendix 1

The Correction for Restricted Range of a Predictive Validity Coefficient

$$r\text{-corrected} = \frac{r\text{-obtained} \quad (S\text{-full} \,/\, S\text{-restricted})}{\sqrt{1 - (r\text{-obtained})^2 + (r\text{-obtained})^2 \,(S\text{-full} \,/\, S\text{-restricted})^2}}$$

Where:

r-obtained = .2532
S-full = 5.54
S-restricted= 4.84

r-corrected = .29