DOCUMENT RESUME

ED 275 761 TM 860 678

AUTHOR	Weber, Jerry; Twing, Jon
TITLE	Comparisons of Methods for Setting Placement Scores of Entering College Freshmen.
PUB DATE	Apr 86
NOTE	36p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Achievement Tests; Basic Skills; College Freshmen; *Comparative Analysis; *Cutting Scores; Higher Education; Measurement Techniques; Regression (Statistics); Research Methodology; *Scoring; Standardized Tests; *Student Placement; Withdrawal (Education)
IDENTIFIERS	California Achievement Tests; Descriptive Tests of Mathematics Skills; Percentile Ranks

ABSTRACT

This study examines the relative value of five empirical procedures for determining cutscores in placement of entering college freshmen. Placement test scores were obtained for all students entering in the fall, but tested in the summer of 1983 at a midwestern community college. Students who took at least one placement test in arithmetic, algebra, or English and who enrolled in at least one of the several freshmen courses were included in the study. The methods chosen were the lowest "C," the equal percentile, and three regression procedures. These procedures derive various cutscores, and this study looks at what placement distributions would have occurred if each procedure's cutscore had been rigidly applied. Relative value is determined by the percentage of true placements versus false placements. In addition, each method is compared with and without the inclusion of students who withdrew from courses. Results indicate limited utility for the lowest "C" and the equal percentile procedures. The two regression lines developed by regressing placement scores on grades (the criterion measure) and by regressing grades on placement scores were unstable across tests. The "average" regression line of these first two produced the most stable (across tests) cutscores and the most acceptable ratios of correct to incorrect placements. (Author/JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ERIC

ED275761

Comparisons of Methods for Setting Placement Scores of Entering College Freshmen

Jerry Weber
John Wood Community College

Jon Twing
University of Iowa

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Weber

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper Presented at the Annual Meeting of the American Educational Research Association

San Francisco
April 1986

TM 860 678

ERIC

Comparisons of Methods for Setting Placement Scores of Entering College Freshmen

Jerry Weber

John Wood Community College

Jon Twing

University of Iowa

ABSTRACT

This study examines the relative value of five empirical procedures for determining cutscores in placement of entering college freshmen. The methods chosen were the lowest "C", the equal percentile, and three regression procedures. These procedures derive various cutscores, and this study looks at what placement distributions would have occurred if each procedure's cutscore had been rigidly applied. Relative value is determined by placement accuracy (that is, percentage of true placements versus false placements). In addition, each method is compared with and without the inclusion of students who withdrew from courses. Results indicate limited utility for the lowest "C" and the equal percentile procedures. The two regression lines developed by regressing placement scores on grades (the criterian measure) and by regressing grades on placement scores were unstable across tests. The "average" regression line of these first two produced the most stable (across tests) cutscores and the most acceptable ratios of correct to incorrect placements.

STATEMENT OF PROBLEM

There has been a general and continued movement at community colleges toward placement testing of entering freshmen to establish the appropriate levels where a student's success would be most likely achieved. Typically placement tests will be administered for English, mathematics, and reading. On the basis of these measures, students will be placed into remedial coursework or the freshman entry courses. Decisions regarding placement rely on the establishment of cutscores (also referred to as decision scores or placement scores). Unfortunately, most of the research on setting cutscores is theoretical, and most applications at the college level have been attempted in foreign language courses. Little work has been done with empirical procedures to look at deriving cutscores for placement decisions regarding basic skills competencies.

This study examines the relative value of five empirical procedures for determining cutscores in placement of entering college freshmen. The methods chosen were the lowest "C", the equal percentile, and three regression procedures. These procedures derive various cutscores, and this study looks at what placement distributions would have occurred if each procedure's cutscore had been rigidly applied. Relative value is determined by placement accuracy (that is, percentage of true placements versus false placements). In addition, each method is compared with and without the inclusion of students who withdrew from courses.

In the past, research on placement tests has focused mainly on the validation process. In the context of basic skills assessment for entering freshmen, very little research has actually been directed toward procedures for setting decision scores. This may reflect a need to place the students according to factors such as available space or subjective judgements. In addition, the problem of setting cutscores for placement differs from the more widely researched problem of setting cutscores for admissions. In the former, students who fall below the cutscore usually receive an alternative treatment (remediation or learning assistance).

Attempts at applying empirical procedures in course placement have not met with much success. Mitchell and Smith (1980) attempted to apply the Rasch model (see also, Smith and Mitchell, 1980) with inconclusive results. Pearse, Agrella, and Powers (1982) attempted to apply the stochastic approximation technique of Dixon and Mood. They found this procedure to be a useful improvement over professional judgment, although they do not report findings that clearly demonstrate its utility (ease of use for example).

An example of a cumbersome, yet effective procedure can be found in the use of decision theory and contingency tables. This methodology has its theoretical basis in the work of Berk (1976) on setting cutscores for criterion-referenced measures. Hartman (1982) found this method useful for determining an optimum cutscore. For this procedure, "pass" (usually a "C" or higher grade) and "non-pass" performance levels are set. Then the percentages of "hits" are weighed against "misses" throughout a range of scores. For a given score, each student that would have been predicted to pass and did pass is a "hit"; likewise, each student predicted as not passing and who, in fact, did not pass is also a "hit." Converse criteria are used to establish "misses." The intersection of most hits with least misses becomes the optimum cut-score. Hartman also included the percentage of students who would be screened (that is, placed in remedial reading) as a factor in determining the cut-score. Klein and Whitney (1977) showed that this method could be combined with a decision procedure based on weighing the relative value of errors in placement. Appenzallar and Kelley (1981; Appenzallar, Dodd, Kelley, 1980), working with placement in foreign language courses, found that by using several methods for determining cutscores they could derive a range of cutscores, each of which produced different consequences for screening and percentanges of errors.

Klein and Whitney (1978) identify several other standard setting procedures that could be applied to course placement. Regression lines can be compared with regard to several possible cutscore criteria. The lowest "C" grade, the mean of "C" grades, the mean of other criteria (final exam scores), or the expected placement test scores could serve as this criteria. When matched with scores on the placement test, possible cutscores could be located. Another method involves the comparison of equivalent scores to some criterion measure (of success in the course).

This study will apply three of the Klein and Wh[illegible] (1978) procedures (Lowest "C", Equivalent/Equal Percentiles and the Regression procedures) to establish cutscores in entry level courses. The criterion of success will be chosen as the grade of "C". The cutscores (and the procedure) will be evaluated by comparison to contingency tables via successful and unsuccessful placements.

METHODS

SUBJECTS

Placement test scores were obtained for all students entering in the Fall, but tested in the summer of 1983 at a midwestern community college. Transcripts were examined for all students tested, a total of about 2,500 students. The students that made up this initial pool were both male and female community college students. The majority of students ranged in age between 18 and 25 years, but older students were also included. Professionals at the community college believed this group of enrollees to be representative of the "typical" student body in past attendance.

The data pool was reduced by eliminating students who did not fulfill the critera needed for this study. To remain in the study, students must have taken at least one placement test in arithmetic, algebra, or English. In addition to the placement test criteria, students must have been enrolled in at least one of several freshman entry level courses; both the course description and the placement test description are given in the next section.

The reduced data set contained all students remaining after the previous two criteria were imposed. What this reduced data pool (containing 1062 students) represented was a cross-classification of subjects in a placement test by course category. Table 1 displays a complete classification of subjects.

MEASURES

The placement tests used to assess students were selected by the institution prior to and independently from this investigation. The placement test in algebra (labeled ALGSCORE) is from the Cooperative Mathematics Test (1963), specifically the "Algebra III Test, Form A." The placement test in arithmetic (labled ARITHSCORE) is from the Descriptive Tests of Mathematical Skills (1978), specifically "The Arithmetic Skills Test, Form A." The placement test in English (labled ENGSCORE) is from the Comparative Guidance and Placement Program (1979), specifically the Test of Written English Expressions. The placement test in reading (labeled RDGSCORE) is from the California Achievement Test (1978), specifically Level 19, Form C.

The courses were selected as representative of freshmen entry level courses in English, math, psychology and sociology. Classes in general mathematics, subject areas like integers, fractions and operations were combined into one course classification referred to as MTH10X (it was later eliminated as an unjustifiable pooling of heterogeneous content areas). MTH120 is Intermediate Algebra, a course intended for students who had previously taken two years of high school algebra. MTH121 is Fundamentals of Mathematics, and was intended for students who had previously taken one year of high school Algebra. MTH122 is College Algebra, and was designed for the mathematically oriented student. Both SOC121, Introductory Sociology, and PSY121, Introductory Psychology, were courses which fulfill general education requirements. ENG105 is Introduction to College Writing, and is a pre-composition course. ENG121 is Introduction to English Composition. ENG125 is Speed Reading, and is designed to enhance comprehension. Although correlations were computed for every combination of course and test score, ENG125, ENG105, and MAT122 were eliminated from the rest of the study because the number of students in all meaningful comparisons was low (below 20). In addition, MTH10X was eliminated because further inve tion showed that the course structure, (variable credit) was unlike the courses. It was the researcher's intention to form a group that was rel ely homogeneous on arithmetic (labled ARITHSCORE) is from the Descriptive Tests of Mathematical Skills (1978), specifically "The Arithmetic Skills Test, Form A." The placement test in English (labled ENGSCORE) is from the Comparative Guidance and Placement Program (1979), specifically the Test of Written English Expressions. The placement test in reading (labeled RDGSCORE) is from the California Achievement Test (1978), specifically Level 19, Form C.

The courses were selected as representative of freshmen entry level courses in English, math, psychology and sociology. Classes in general mathematics, subject areas like integers, fractions and operations were combined into one course classification referred to as MTH10X (it was later eliminated as an unjustifiable pooling of heterogeneous content areas). MTH120 is Intermediate Algebra, a course intended for students who had previously taken two years of high school algebra. MTH121 is Fundamentals of Mathematics, and was intended for students who had previously taken one year of high school Algebra. MTH122 is College Algebra, and was designed for the mathematically oriented student. Both SOC121, Introductory Sociology, and PSY121, Introductory Psychology, were courses which fulfill general education requirements. ENG105 is Introduction to College Writing, and is a pre-composition course. ENG121 is Introduction to English Composition. ENG125 is Speed Reading, and is designed to enhance comprehension. Although correlations were computed for every combination of course and test score, ENG125, ENG105, and MAT122 were eliminated from the rest of the study because the number of students in all meaningful comparisons was too low (below 20). In addition, MTH10X was eliminated because further investigation showed that the course structure, (variable credit) was unlike the other courses. It was the researcher's intention to form a group that was relatively homogeneous on course content from the above domain of skills, but inspection of the course (MTH10X) indicated this was not possible. Finally, PSY 121 and SOC 121 were eliminated because of low correlation with the reading placement test.

STANDARD SETTING PROCEDURES

Each placement test was used to investigate the effects of standard setting when used to place students. A list of the placement test by course classification is presented in Table 1.

The arbitrarily determined criterion of success in a course was set as a grade of C. It was recognized that this would not be an acceptable criterion in all other studies and that any criterion of success could have been chosen without changing any of the standard setting procedures.

In actual practice, placement scores were used as information to assist in placement advisement. However, for this study the data generated were used to simulate the effects of mandatory placement of each student into the paired course using different cutscores. Various cutscores were calculated using the methods identified by Klein and Whitney (1978) in the following paragraphs.

The Lowest "C" Method

Using this procedure, students were ranked from high to low based on grades the students had received in a select course. After obtaining this ranking, the students who received a grade of C in the course were identified. The lowest placement test score associated with the "C students" was taken as the placement test cutscore. Figure 1 gives an example of how this procedure was performed. After plotting the courses by placement test, the C grade category was found by reading across the horizontal axis to the grade class. Then reading up the graph, the first score encountered (a score point is represented by an asterisk) was the lowest placement test score of the students with the "c" course grades. This score was then taken as the placement test cutscore.

For Klein and Whitney, the placement test score that was received by the "poorest C grade student" was taken as the cutscore. That is, the Klein and Whitney cited GPA and final exam scores as criteria thus students had varying magnitudes of those C grades. This study used grades only and, as such, the Lowest C method was modified.

The Equal/Equivalent Percentile Method

Using this method, the number of students receiving a grade below a C (D and F students) in each class, were counted. This number was then converted into a proportion. The percentile that corresponded to this proportion on the placement test was identified as the cutscore.

Regression Procedures

The regression procedure generated three regression lines in total. One regression line is generated from predicting the course grade from the placement test. The other regression line was generated from predicting the placement test score from the course grade. The point where these two lines intersect with the criterion of success in the course (an arbitrarily determined C) generated the two cutting scores respectively. This procedure is represented in Figure 2. Appenzallar and Dodd (1978) suggest a compromise in the above procedure by taking as the cutting score the point where the intersection of the "average" regression line and the criterion of success occurs. The average regression line is generated by taking the intersection of the angle between the previous two regression lines at their point of intersection.

The rationale for these three regression lines is explained by Willingham (1978). Using Figure 2 as a guide, it can be seen that by using the course grade regression line any slight change in the criterion of success will result in a large change in the cutscore and hence would have a serious impact on the distributions of successful and false placements. However, if the placement test regression line was used alone, a major change in the criterion of success would only have a minor influence on the cutscore and hance only a slight reflection on the distributions of successful and false placements (due to the smaller slope of this line). Therefore, Willingham suggests a compromise in determining the "average" regression line: that is, the line that intersects the angle between the previous two regression lines.

In most regression work, the problem of "deviant data" or data that seems to behave differently than the rest of the data pool is a major concern. Hence it is necessary to examine the bivariate plot of the standardized (studentized) residual correlations to ascertain how well each specific element of the data conforms to the representations of the entire group. By inspecting the plot of the residuals, these atypical scores or "outliers" can be identified and eliminated. In accordance with tradition, it was decided to eliminate from the data pool any score with a standardized (studentized) residual that fell two or more standard deviations away from the mean standardized (studentized) residuals (Pedhazur, 1978). After eliminating these outliers (33 scores total), and inspecting the type of score eliminated to be sure no one group of scores was treated differentially, the regression lines were recalcualted and plotted. The new cutting scores were obtained. Upon inspection of these new scores, only a few were modified by removing the outliers.

The few that were affected were influenced by only 2 or 3 points on the average (Median difference was approximately 2.0 points). It was therefore concluded for the sake of brevity to eliminate from analysis the cutscores obtained after eliminating outliers (see Figure 3 for an example of a residual plot).

WITHDRAWALS

It was recognized that students who for one reason or another dropped out of the data pool before a final grade was assigned might create a confounding of the study that would make any score interpretation difficult. Without a clear theoretical basis to guide how these "withdrawals" should be treated, the cutting scores for both the data pool which contained the students who had withdrawn and the data pool which treated the students who had withdrawn as failures ("W" were set equal to failing) were calcualted.

CONTINGENCY TABLES

For evaluating the procedures previously identified, a contingency table, establishing all possible cutscores was constructed. Using the strategy of Berk (1978), a measure of accuracy of placement was developed. People were placed into one of four possible categories by locating them in a two by two table according to their placement test score (generated from the contingency table) and course grade. Logical basis for labeling the cells of this table are given as follows:

> True Master: A student scoring above the placement test cutting score and above the course criterion.
>
> True Nonmaster: A student who scored below the placement test cutting score and below the course criterion.

False Master: A student who scored above the placement test cutting score, but below the course criterion.

False Nonmaster: A student who scored below the placement test cutting score, but above the course criterion.

An example of Berk's classification scheme is depicted in Figure 4. It was further stipulated that a successful placement would consist of all student's scores identified as either True Master of True Nonmaster, and would be labeled as "correct." Incorrect placements were then identified as False Master or False Nonmaster, a False Master would be referred to as "too high" and a False Nonmaster would be referred to as "too low.". A False Master represents a placement that is "too high" because this student would score above the cutscore yet receive an unsuccessful grade. Likewise, a "too low" placement represents someone who has scored below the cutscore yet has succeeded by virtue of a "C" or better grade. "Correct" placements are those for whom a given cutscore accurately predicts success or failure in receiving a "C" or better grade.

Given any cutscore, a simple counting of the numbers of high, correct, and low placements would reflect the value of a given standard setting procedure. For example, if 70% of the placed students were placed too high, by one procedure but only 10% were placed too high by another procedure, evidence for the support of the latter procedure as superior would be indicated.

RESULTS

The accuracy of a given cutscore depends on a judgement concerning the relative value of the percentages of high, correct, and low placements. Of particular importance is the relationship (ratio) between high and low placements. High placements represent students who were misplaced and unsuccessful. Low placements represent students who would be put in remediation. Students placed too low should have a high chance of success in the remedial courses. Thus, in examining the value of each cutscore procedure, the following judgements are made about cutscores:

1. A minimum acceptable cutscore would be one with 60% correct placements and a greater number of low placements than high placements.
2. An ideal cutscore would be one that yields 10% too high, 70% correct, 20% too low placements, respectively.

These judgements represent the criteria by which each procedure's results are evaluated. It is of course, recognized that these criteria are not empirical; however, these criteria represent for the authors reasonable professional judgement.

LOWEST C

Including the students who had withdrawn as either missing or as failures (equivalent to a grade of "F") made no difference in the Lowest C Method. Missing data or failures did not affect the number of students with a course grade of C.

The Lowest C Method consistently yielded lower cutting scores than the other procedures. In addition, this procedure was less accurate with regard to high, correct, and low placements than any other procedure since the Lowest C Method consistently placed too many students too high. This is readily seen by inspecting Table 2. A cutscore with a percentile rank of 1 will assuredly place too many students too high!

EQUAL PERCENTILE

The Equal Percentile Method yielded low cutting scores and a poorer high to correct to low ratio on the average than did the regression procedures. The Equal Percentile Method resulted in only a little better accuracy than the Lowest C Method as indicated by Table 3. For example, the lowest C method for Math 120 when withdrawals were treated as missing placed 67% correctly but 30% too high. Whereas the corresponding Equal Percentile method placed 69% correctly but 31% too high.

REGRESSION METHOD

Inspection of Tables 4 and 5 indicates that the "average" regression approach yielded more stable cutscores while containing an acceptable ratio of high, correct, and low placements. For example, consider the English placement test where students who had withdrawn are treated as failing. The cutscore for regressing the placement test on the course grade (ENG121) is 63. Its counterpart, regressing course grade on placement yields a cutscore of 25. Immediately this appears to be a large difference; as inspection of the placement ratios will confirm. The former cutscore would place 10, 54, and 36 percent of the students as too high, correct, and too low respectively. According to the criterion previously specified, this cutscore places too many students too low; hence the cutscore was too high. The latter cutscore fairs little better, placing 19, 74, and 6 percent of the students too high, correct, and too low respectively. Although these placement ratios may seem allowable, the "average" regression cutscore is 53 and places 14, 59, and 27 percent of the students as too high, correct, and too low respectively. These ratios, from the "average" regression, are seen as the best. Inspection of the rest of Tables 4 and 5 reveals the "average" regression cutscore as superior to cutscores generated by either of the other two regressions.

READING PLACEMENT TEST

Inspection of the correlation matrices of Tables 6 and 7 reveals some peculiarities in the underlying data structure. Observing the correlations between the placement tests (Percentile Ranks) and the standardized raw scores for these same tests yields correlations of .98, .95, .98, and .06 for the algebra, arithmetic, English, and reading placement tests respectively. The (incredibly) low correlation between the standardized raw scores for the Reading placement tests and its corresponding Percentile Rank can only indicate errors in the data. Taken at face value, this low correlation indicates that the standardized raw scores would rank order the subjects differently than the corresponding Percentile Ranks. This is nonsensical. Hence steps in the data analysis procedures were retraced to the original data source (college transcripts). After interviews with the testing personnel at the college, it was discovered that the reading placement test scores were hopelessly confounded. The actual reading placement test could have been one of not less than four previously used tests. These tests differed in difficulty, content, reliabilities, and number of items. With little if no apparent choice, all analyses resulting from the use of the Reading Placement test was eliminated. It should be noted that the community college had already changed its testing program to eliminate such anomalies as the above, before this investigation had begun.

CORRELATIONS AND RELIABILITIES

The utility of any placement test is related to its ability to predict successfully the criterion measure. The index of this ability is reflected by the correlation between the placement test scores and course grades, both of which are modified by their respective reliabilities. The reliabilities of the placement tests are relatively well established by virtue of their widespread use. The reliabilities of grades is highly questionable and are usually low. Nonetheless, the correlations are presented for inspection in Tables 6 and 7. The reliabilities (using K-R 20) of the placement test were the following: arithmetic, .87; algebra, .85; English, .83; and reading, .85.

DISCUSSION

Accuracy is easily defined as a successful placement. It is the definition of success that has caused many researchers grief. A successful standard setting procedure is not one which generates the most "true masters and true nonmasters!" There are two directions in which a placement can be inaccurate. We can set the cutting score too high and generate a disproportionate amount of misses in the direction of "false nonmaster." If this occurred people would be placed below their ability; being labeled as eventually failing when in fact they would pass. The risks of this type of error in placement are boredom and frustration. Time might be wasted, but it is submitted that very few people will know everything about a given subject regardless of how commonplace it may seem. It is also recognized that the student put into this position might be identified by classroom activities. If this occurs, some alternate form of advanced work could be proposed, or advancement to a higher level class could be arranged. On the other hand, we could set the cutting score too low, causing a greater influx of misses into the "false master" category. These are the people who are predicted to succeed in the course but ultimately fail. The risk here is not only frustration, but lower grades and wasted time and money on a course that may have to be taken again. In addition, these are students for whom the placement system has failed in the worst way: it has stacked the course content against their skill level. Given the above arguments, the authors maintain that the errors assoicated with the latter scenario are more "costly." Hence, it is better to set a cutscore a little higher and therefore place students too low (following on the above rationale). If it is taken into account that placement here is into an entry level freshman course, this argument takes on even more meaning. The student's entire future performance in an area will be guided or obstructed by the presence (or lack) of key basic skills prerequisites. Therefore, any judgements regarding the accuracy or utility of any standard setting procedure are made with this point of view in mind.

The results of this study are not as far from common sense as it first might seem at first inspection. Less than 10% of the cases were eliminated as outliers from the plots of the studentized residuals. This being so, little if any change in regression weights and hence cutscores was expected, and little occurred.

The differences in procedures based on including and excluding the cases of withdrawls from courses should have, and did, modify the procedures that used distribution specific information to establish the cutscore (Equal Percentile & Regression Methods). The issue of placement regarding withdrawals is far from resolved. This study may provide insight into investigation of this issue in the future, but is of secondary concern here.

It was no surprise to find that the Lowest C Method provided the poorest cutscores. This was likely for two reasons. First, this method is void of any distributional information at all. It solely depends upon the students who receive "C" course grades, and the cutscore rests entirely on the performance of one individual. Secondly, the present study uses a highly questionable modification of Klein and Whitney's original method; the lowest placement test score from the "C students" was taken as the standard and not the test score associated with the poorest student who had received a C course grade. Again, this was unavoidable due to the nature of the data, but still violates the underlying theme of an ordinal scale. Klein and Whitney's original method would order the subjects on the basis of the course grade and would then "fix" as the cutscore the placement test score associated with the student who had the minimally acceptable grade (the lowest C). The implication is that if the test measured the same ability as the course taught, then any score higher than the score associated with this minimally acceptable grade would also be acceptable. In the present study however, individual scores were not rank ordered, only groups of scores (A, B, C etc.) were. Hence one cannot tell if the lowest placement test score for the category of "C" students would have been the lowest "C" student's actual score. This compromise was necessary however because the practicality of using something other than grades in the course (as reported) is limited.

The "Average" Regression method produced cutscores that were the most acceptable (based on the criterion previously established) and the most consistent across courses. This method is affected by the strength of the relationship between test scores and grades; and higher correlations should produce a better placement ratio (high, correct, low). In addition, using a different criterion measure such as the average "C" grade might produce better ratios. Of course, a truly continuous criterion measure of ability is the most preferred.

All results, and any decisions for that matter are strapped to the reliabilities of the measures at hand. The placement tests were all nationally available standardized tests, with evidence for validity and reliability available as well as national norms. The reliabilities were reported earlier, with all other information at the test users ready. Hence issues of reliability should not be a large concern regarding the placement tests. However, this does not mean the tests are guaranteed to be used properly for placement information.

The evidence of reliability for course grades is always suspect. Only the personnel at the grass roots level can shed light on this issue. Again, a different criterion measure may not be plagued by reliability problems, but a trade off with utility is a cost that must be dealt with.

In conclusion, this study's findings must be taken as preliminary. While the "Average" Regression appears by the authors' definition the most accurate, further testing of these procedures on other data is necessary to make any generalizations. All these procedures have some utility in that each supplies different cutscores with varying consequences and each can be produced without the cumbersome procedure of establishing contingency tables; although these tables are an essential part of evaluating these same said cutscores. It is noted however that no procedure alone is an adequate replacement for good sound professional judgement, nor do the procedures cited here purport to be. Judgement on the other hand is the direct result of good informational input combined with expertise and wisdom. Acquiring more information to make these judgements better is all any empirical procedure can do, at least at this time.

REFERENCES

Appenzallar, A. B. & Kelley, H. P. (1982) Re-evaluation of decision scores used in course placement and credit by examination in Spanish at U. T. Austin: the Spanish validity study of 1981. (Report No. RB-82-4). Austin, TX: Measurement and Evaluation Center, University of Texas at Austin.

Appenzallar, A. B., Dodd, B. G., & Kelley, H. P. (1980). Validity study of two College Board German tests for use in course placement and credit by examination at U. T. Austin, 1979-1980. (Report No. RB-80-9). Austin, TX: Measurement and Evaluation Center, University of Texas at Austin.

Berk, Ronald (1976). Determination of Optional Cutting Scores in Criterion Referenced measurement. Journal of Experimental Education, 45, 4-9.

Bray, D. (1983). Assessment placement: a design for student success and institutional planning for the 1980's (Management Report, 1983 -3/4). Association of California Community College Administrators. (ERIC Reproduction Service No. ED 230 234)

California Achievement Tests (1978). Level 19, Form C. CTB/McGraw Hill; Monterey, California.

Comparative Guidence and Placement Program (1979). Written English Expressions Test. The College Entrance Examination Board; Princeton, New Jersey.

Cooperative Mathematics Test (1964). Algebra III, Form A. Educational Testing Service; Princeton, New Jersy.

Descriptive Tests of Mathematical Skills (1978). The Arithmetic Skills Test, Form A. College Entrance Examination Board; Princeton, New Jersey.

Frisbie, D. A. (1982). Methods of evaluating course placement systems. Educational Evaluation and Policy Analysis, 4, 133-140.

Glass, G. V. (1978). Standards and Criteria. Journal of Educational Measurement, 4, 237-261.

Hartman, N. (1981). Maximizing the effectiveness of reading tests in the community college. St. Louis, MO: St. Louis Community College. (ERIC Document Reproduction Service No. ED 237 121).

Klein, A. C. & Whitney, D. R. (1977). A discussion and comparison of selected methods for determining cutoff scores for proficiency and placement tests (Placement and Proficiency Testing Report No. 6). Iowa City, Iowa: University of Iowa, Evaluation and Examination Service.

REFERENCES (cont.)

Livingston, S. A. (1982). Passing Scores. A Manual for setting Standards of Performance on Educational and Occupational Tests. Princeton, NJ: Educational Testing Service.

Mitchell, V. P. & Smith, R. M. (1980). Determining cut-off scores using Rasch abitity estimates. Paper presented at the annual meeting of the National Council on Measurement in Education, Boston. (ERIC Document Reproduction Service No. ED 193 235).

Pearse, C., Agrella, R., & Powers, S. (1982). Identification and placement of college students in developmental education programs. Paper presented at the annual meeting of the California Educational Research Association, Sacremento. (ERIC Document Reproduction Service No. ED 231 415)

Pedhazur, E. J., (1982). Multiple Regression in Behavioral Research: Explanation and Prediction. Holt, Rinehart and Winston; New York, New York.

Reed, M. D. (1972). Studies of student performance on the Comparative Guidance Placement Program. Lynchburg, VA: Central Virginia Community College (ERIC Document Reproduction Service No. ED 063 930)

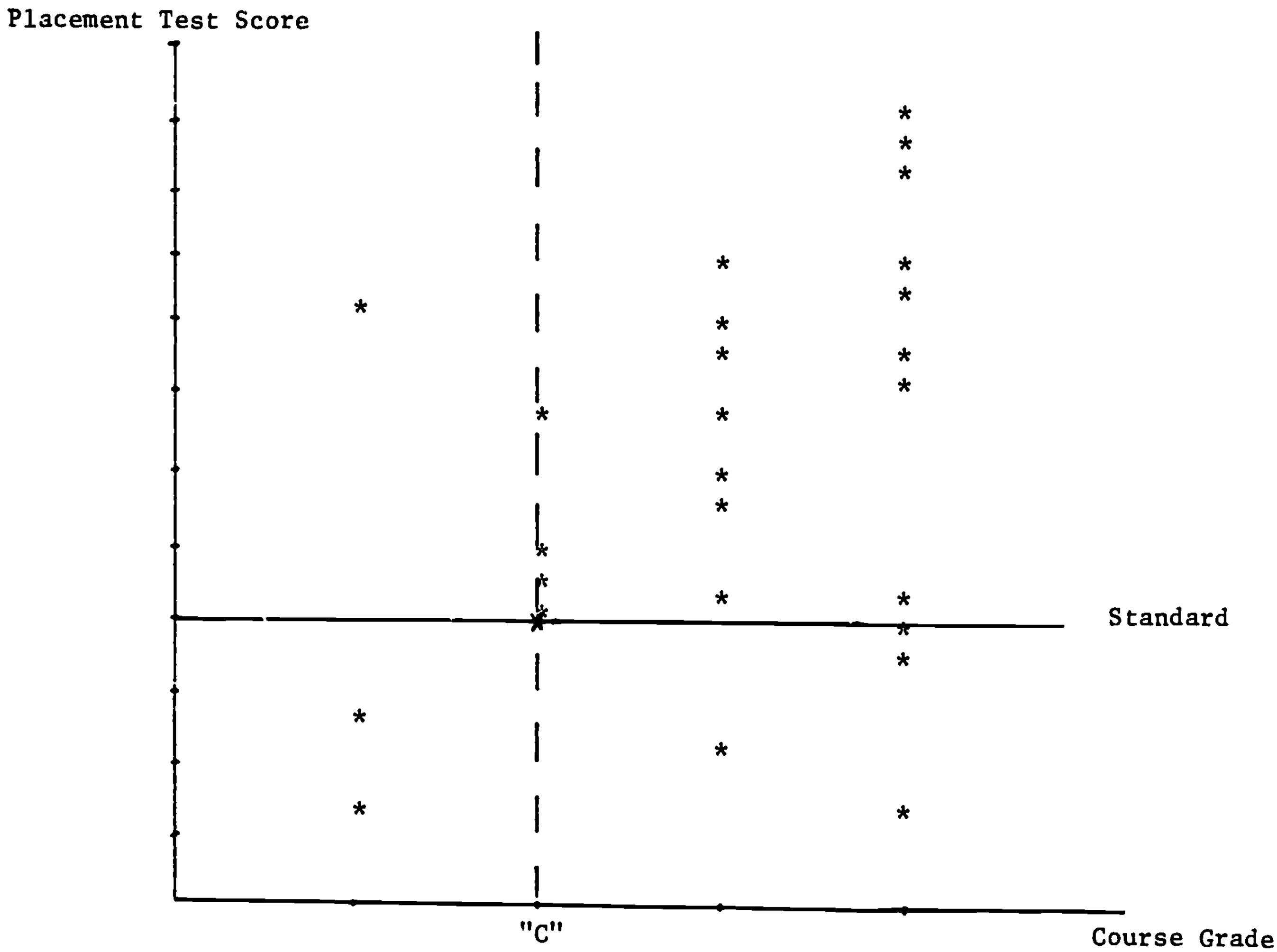
Roueche, J., Baker, G., & Roueche, S. (1984). College responses to low-achieving students: a national study. Orlando, FL: HBJ Media Systems Corporation.

Smith, R. M. & Mitchell, V. P. (1980). An analysis of Rasch person fit statistics in college placement tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Boston. (ERIC Document Reproduction Service No. ED 189 174).

Willingham, W. (1974). College placement and exemption. New York: College Entrance Examination Board. New York: (ERIC No. ED 100 225).

FIGURE 1

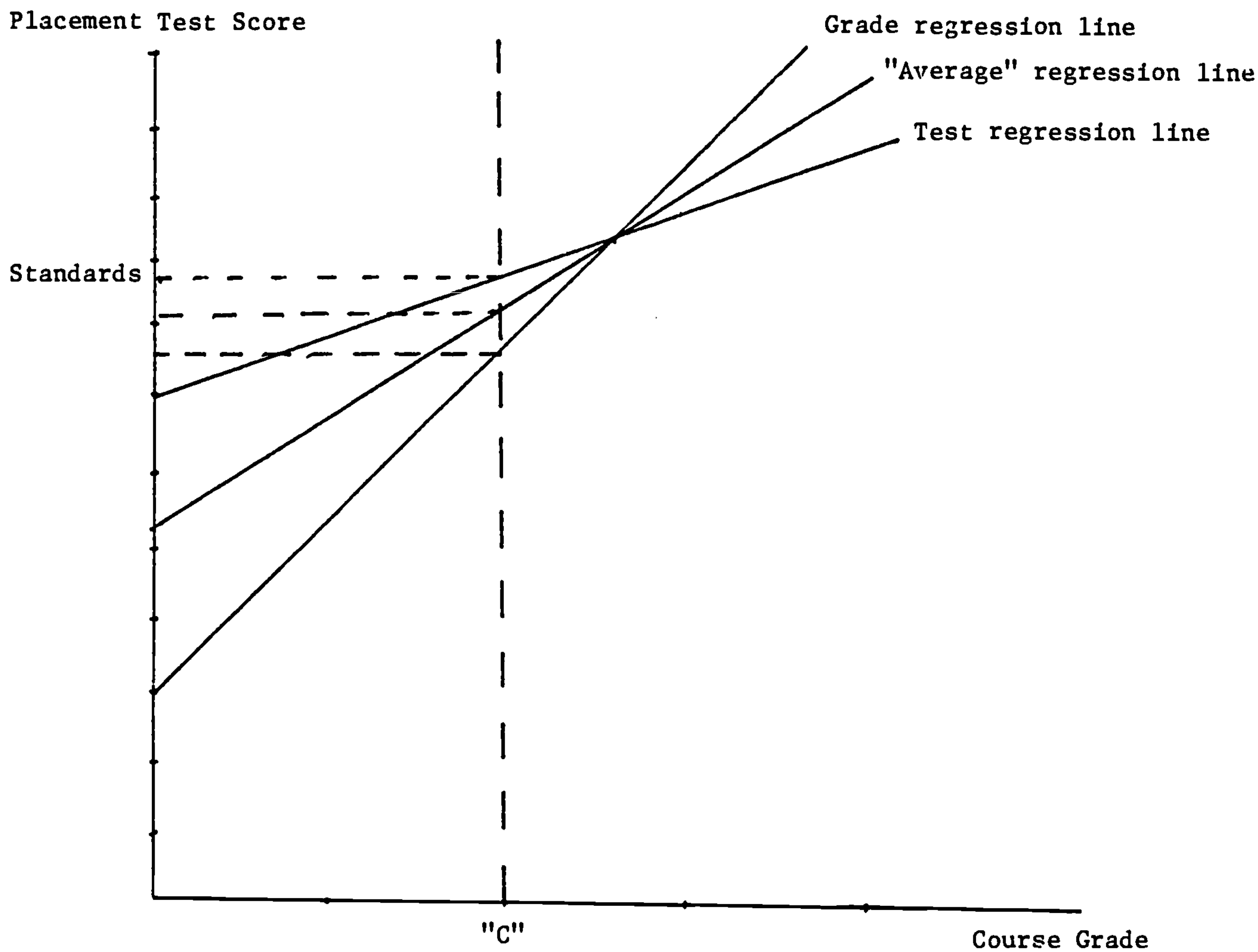
EXAMPLE OF THE LOWEST C METHOD



Note: each * represents a single observation

FIGURE 2

EXAMPLE OF THE REGRESSION PROCEDURES



ERIC Full Text Provided by ERIC

FIGURE 3

RESIDUAL PLOT FOR OUTLIERS

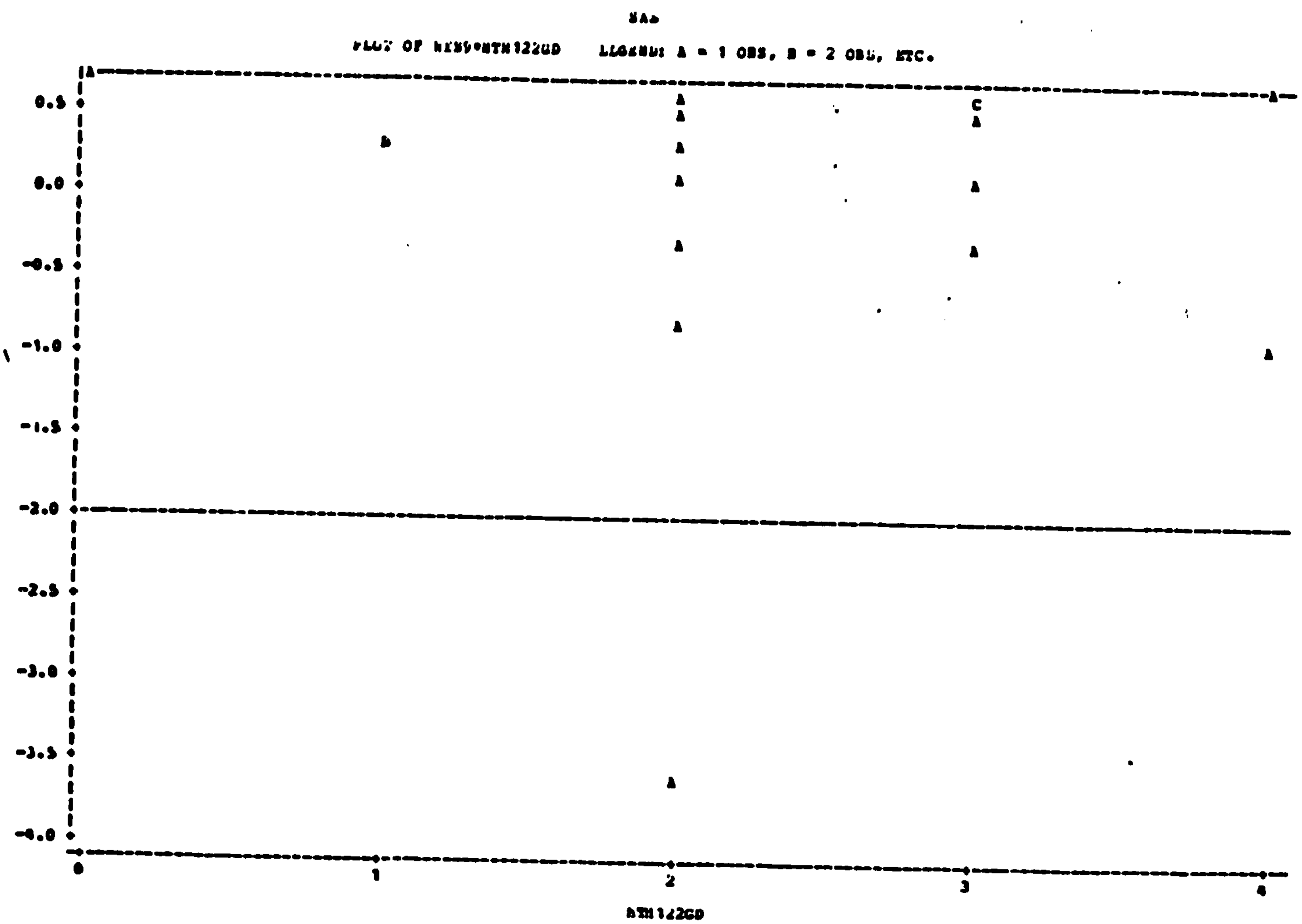


FIGURE 4

CRITERION CLASSIFICATION

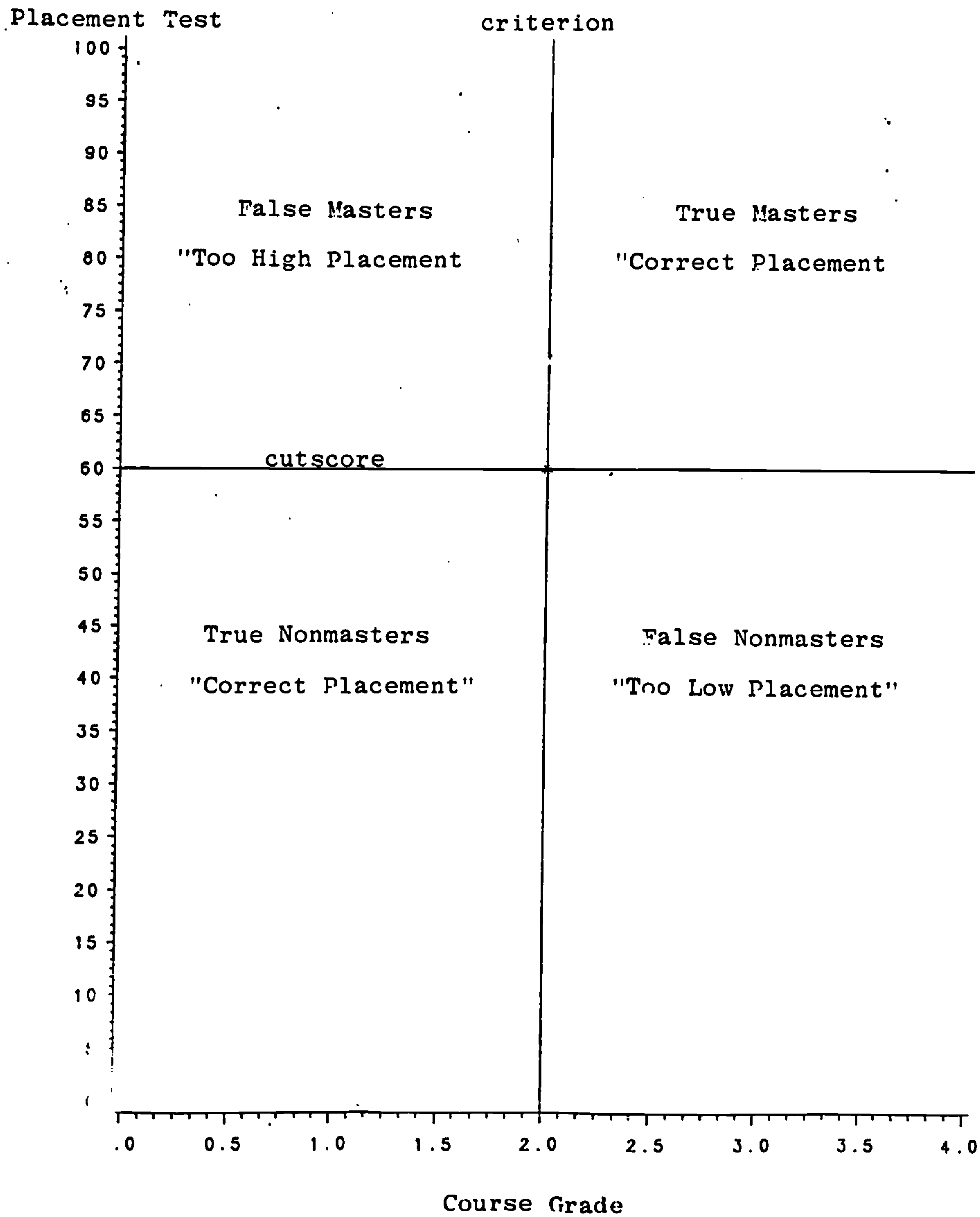


Table 1

Match Of Placement Test By Courses

Test and Course	Number (Withdrawal=missing)	Number (Withdrawal=failing)	Number (total Withdrawal)
English Placement Test			
English 121	336	380	44
Arithmetic Placement Test			
Math 120	61	99	38
Math 121	68	79	11
Algebra Placement Test			
Math 120	67	96	29
Math 121	19	22	3
TOTAL SCORES	551	676	125

Note: The marginal totals (TOTAL SCORES) are not total subjects. One subject could have taken any (or all five classes) and as such, would have contributed more than one score to the total. The reduced data pool originally contained 1062 students.

Table 2

Cutscores Established By Procedure

Lowest C Method		Equal Percentile Method		Regression Of Placement Test On Grade		Regression Of Grade On Placement Test		"Average" Regression	
withdrawal equal missing	withdrawal equal failing	withdrawal equal missing	withdrawal equal failing	withdrawal equal missing	withdrawal equal failing	withdrawal equal missing	withdrawal equal failing	withdrawal equal missing	withdrawal equal failing
01	01	12	23	58	63	N/A	25	50	53
51	51	27	52	49	84	80	N/A	83	92
56	56	43	51	82	77	88	96	80	83
06	06	27	52	88	48	23	83	44	55
09	09	43	51	28	36	27	36	32	36

NOTE: All cutscores are expressed as percentile ranks (PR's).

BEST COPY AVAILABLE

Table 3

High, Correct, And Low Placement For Lowest C And Equal Percentile

		Lowest C Method withdrawal equal missing		Lowest C Method withdrawal equal failing		Equal Percentile Method withdrawal equal missing		Equal Percentile Method withdrawal equal failing	
		N	Percent	N	Percent	N	Percent	N	Percent
Engscore by Eng 121	High	44	13	88	23	41	12	74	20
	Correct	292	87	292	77	287	85	282	74
	Low	0	0	0	0	8	3	24	6
Cutscore		01		01		12		23	
Arthscore by MTH 120	High	18	30	47	48	19	31	44	44
	Correct	41	67	50	50	42	69	52	53
	Low	2	3	2	2	0	0	3	3
Cutscore		51		51		27		52	
Arthscore by MTH 121	High	24	35	32	40	26	38	33	42
	Correct	41	60	45	56	40	59	44	56
	Low	3	5	3	4	2	3	2	2
Cutscore		56		56		43		51	
Algscore by MTH 120	High	16	24	45	47	11	16	13	17
	Correct	51	76	51	53	44	66	36	47
	Low	0	0	0	0	12	18	27	36
Cutscore		06		06		27		52	
Algscore by MTH 121	High	3	16	5	23	0	0	2	9
	Correct	16	84	17	77	8	42	9	41
	Low	0	0	0	0	11	58	11	50
Cutscore		09		09		43		51	

BEST COPY AVAILABLE

Table 4

High, Correct, And Low Placement For Regression (Withdrawal = Missing)

		Regression Of Placement Test On Grade			Regression Of Grade On Placement Test			"Average Regression"	
			N	Percent		N	Percent		N	Percent
Engscore by ENG 121	High		21	6		--	--		31	9
	Correct		195	58		--	--		221	66
	Low		120	36		--	--		84	25
Cutscore		58			N/A			50		
Arthscore by MTH 120	High		18	30		12	20		12	20
	Correct		41	67		40	66		40	66
	Low		2	3		9	14		9	14
Cutscore		49			80			83		
Arthscore by MTH 121	High		10	15		6	9		10	15
	Correct		41	60		43	63		41	60
	Low		17	25		19	28		17	25
Cutscore		82			88			80		
Algscore by MTH 120	High		2	3		11	16		8	12
	Correct		23	34		45	68		36	54
	Low		42	63		11	16		23	34
Cutscore		88			23			44		
Algscore by MTH 121	High		0	0		2	11		0	0
	Correct		12	63		13	68		12	63
	Low		7	37		4	21		7	37
Cutscore		28			27			32		

Table 5

High, Correct, And Low Placement For Regression (Withdrawal = Failing)

		Regression Of Placement Test On Grade			Regression Of Grade On Placement Test			"Average Regression"		
			N	Percent		N	Percent		N	Percent
Engscore by Eng 121	High		38	10		74	19		55	14
	Correct		207	54		282	74		224	59
	Low		135	36		24	6		101	27
Cutscore		63			25			53		
Arthscore by MTH 120	High		27	27		--	--		13	13
	Correct		63	64		--	--		64	65
	Low		9	9		--	--		22	22
Cutscore		84			N/A			92		
Arthscore by MTH 121	High		17	22		2	3		13	16
	Correct		50	63		48	61		49	62
	Low		12	15		29	36		17	22
Cutscore		77			96			83		
Algscore by MTH 120	High		15	16		3	3		13	14
	Correct		58	60		53	55		55	57
	Low		23	24		40	42		28	29
Cutscore		48			83			55		
Algscore by MTH 121	High		2	9		2	9		2	9
	Correct		13	59		13	59		13	59
	Low		7	32		7	32		7	32
Cutscore		36			36			36		

Table 6

Correlations (Withdrawal = Missing)

(STD = Standardized Raw Score)
Integer values represent sample size

W=8	RDG	ENG	ART	ALG	PSY	SOC	MTH 10X	MTH 120	MTH 121	MTH 122	ENG 121	ENG 125	STD READ	STD ENG	STD ART	STD ALG
RDG	887 1.00	.6447 884	.5102 702	.3713 233	.3230 336	.2073 232	.0894 82	.1727 88	.4013 79	-.1127 24	.1266 312	.1460 17	.0551 847	.6453 885	.4313 703	.3986 232
ENG		938 1.00	.4703 701	.4578 237	.3707 353	.2885 243	-.0546 84	.2071 91	.4078 83	-.2244 27	.2052 336	.3059 17	.0728 844	.9814 938	.4899 702	.4797 236
ART			716 1.00	.22995 55	.2728 269	.2274 196	.0027 74	.3062 61	.3507 68	.0080 18	.1096 252	-.0908 13	.0396 673	.4710 702	.9491 716	.3008 54
ALG				347 1.00	.0969 112	.1063 66	.3945 72	.1881 67	.4463 19	-.063 6	.1657 101	-.4260 9	.1579 219	.4557 237	.2447 55	.9829 345
PSY					379 1.00	.6198 133	.6463 42	.6625 52	.6701 40	-.0519 14	.4804 179	.8501 7	.1401 323	.3663 353	.3293 270	.0872 112
SOC						264 1.00	.0747 24	.6311 28	.4556 32	.4264 8	.5653 120	.8165 6	.1607 222	.2811 244	.2338 196	.1132 65
MTH 10X							130 1.00	.6677 20	.6883 4	.000 0	.2993 38	.4267 4	.0402 81	-.0428 84	.0385 74	.4105 71
MTH 120								1.00 118	1.00 2	.8729 5	.4933 58	-.8660 3	.2408 86	.2224 91	.4195 61	.2199 66
MTH 121									1.00 90	0000 0	46615 47	0000 0	-.0284 76	.3992 83	.3413 68	.4697 19
MTH 122										1.00 27	.4182 16	.0000 1	.1371 24	-.2448 27	.0275 18	-.0837 6
ENG 121											1.00 354	.2332 6	16478 302	.20[illegible]5 336	.1728 253	.1919 100
ENG 125												1.00 18	.3268 17	.2939 17	-.068 13	-.4392 9
STD READ													1.00 847	.3993 839	.0728 674	.1865 219
STD ENG														1.00 939	.4899 703	.4812 236
STD ART															1.00 717	.3081 54
STD ALG																1.00 345

ERIC

Table 7

Correlations (Withdrawal = Failing)

(STD = Standardized Raw Score)

Integer values represent sample size

W=0	RDG	ENG	ART	ALG	PSY	SOC	MTH 10X	MTH 120	MTH 121	MTH 122	ENG 121	ENG 125	STD READ	STD ENG	STD ART	STD ALG
RDG	1.000 887	.6447 884	.5102 702	.3713 233	.5398 390	.2469 260	.0439 103	.2183 132	.3832 93	-.0067 37	.1087 350	.2686 23	.0551 847	.6453 885	.4313 703	.3986 232
ENG		1.00 938	.4703 701	.4578 237	.3817 409	.3211 272	-.0738 105	.2391 138	.3982 97	.0252 40	.1688 380	.2869 23	.0728 844	.9813 938	.4899 702	.4797 236
ART			1.00 716	.22995 55	.2544 318	.2657 218	.0569 92	.35996 99	.3368 79	.1060 27	.1449 285	.0552 18	.0396 613	.4710 702	.9491 716	.3008 54
ALG				1.00 347	.1928 125	.1842 72	.4028 99	.2483 96	.1936 22	.3711 10	.2313 118	.3404 11	.1579 219	.4557 237	.2449 55	.9829 354
PSY					1.00 438	.5469 161	.4797 60	.6826 84	.6057 48	.3012 20	.4371 227	.6713 11	.1328 375	.3776 409	.2987 319	.1926 125
SOC						1.00 293	.3426 31	.5653 46	.4149 79	.3203 13	.5045 141	.8165 6	.1339 247	.3101 273	.2991 219	.1848 71
MTH 10X							1.00 171	.5826 333	.2899 6	.000 0	.3559 50	.0000 5	.0555 99	-.0653 105	.0639 92	.4207 98
MTH 120								1.00 180	.7303 4	.8131 8	.4340 91	.4774 7	.1222 129	.2320 138	.4008 99	.2719 95
MTH 121									1.00 104	.00 2	.6124 55	.0000 1	.0462 90	.3846 97	.3285 79	.2312 22
MTH 122										1.00 40	.3425 27	.0000 1	.1300 36	.0066 40	-.0025 27	.3683 10
ENG 121											1.00 403	.6752 8	.1447 339	.1723 380	.20635 286	.2369 117
ENG 125												1.00 24	.2535 22	.2876 23	.0413 18	.31765 11
STD READ													1.00 847	.0746 845	.0728 674	.1865 219
STD ENG														1.00 939	.4899 703	.4812 236
STD ART															1.00 717	.3081 54
STD ALG																1.00 354

BEST COPY AVAILABLE

ERIC