ACT Research Report Series 92-3

The Effects of Data Truncation on Estimated Validity Indices for Course Placement

Jeff Schiel
Julie Noble

October 1992

THE EFFECTS OF DATA TRUNCATION ON ESTIMATED VALIDITY INDICES FOR COURSE PLACEMENT

Jeff Schiel
Julie Noble

ABSTRACT

Traditionally, correlation coefficients have been used to validate course placement decisions based on test scores and high school grades. Because placement systems restrict the range of both the predictor and outcome variables, correlation coefficients based on data from students enrolled in particular courses are understated relative to what they would be if placement had not occurred. Alternative methods have therefore been examined for validating placement systems. One such approach uses validity indices estimated from logistic regression analyses and distributions of predictor variables to determine placement effectiveness.

The ASSET Basic Skills test scores and course grades of entering freshman from four postsecondary institutions were analyzed to determine the impact of prior selection on the accuracy of estimated validity indices. Estimated validity indices based on truncated distributions of test scores and course grades were compared to the same indices based on full distributions. It was found that greater degrees of truncation are associated with a loss of accuracy in estimated validity indices. However, the loss of accuracy in the estimates was small when less than 15% of the data for the full distributions were truncated.

THE EFFECTS OF DATA TRUNCATION ON ESTIMATED VALIDITY INDICES FOR COURSE PLACEMENT

It is common practice for postsecondary institutions to use standardized test scores for placing students into college-level courses. If a student's test score is at or above a specified cutoff, then she or he would be placed into a standard-level course. If instead the student's score is below the cutoff, she or he would be placed into a developmental or lower-level course.

Placement decisions, whether correct or incorrect, may affect individual students in several ways. For example, if a student is incorrectly placed in a standard-level course, she or he may be unable to complete it satisfactorily because the level of the course work exceeds the student's level of knowledge and skills. On the other hand, if a student is placed in a developmental course, then she or he may have to pay additional tuition, simply because of the extra course work that must be undertaken. Further, the student may have to allocate more time toward earning a degree than she or he originally anticipated. If the student is incorrectly placed in the developmental course, then the level of course work may not be sufficiently challenging, and she or he may become discouraged.

Placement decisions may also affect the institution. If many students are identified as needing remediation, for example, it may be necessary to schedule extra sections of a particular developmental course or to hire additional teaching staff. If the students are incorrectly identified as needing remediation, such hiring or scheduling efforts may be superfluous.

Because of the importance of placement decisions, it is essential that they be as accurate as possible. If test scores are used to make these decisions, but are not valid for use in course placement, then placement decisions based on the scores cannot be accurate. Traditionally, correlation coefficients have been used to document the strength of the statistical relationships between test scores and course grades, and thereby serve as a measure of the validity of the test scores. There are, however, some disadvantages associated with using correlation coefficients for this purpose.

At most institutions, students are placed into standard-level courses using test scores and/or other related information. Students scoring above a specified cutoff score are placed into the course, while students scoring below the cutoff are placed into remedial courses. When outcomes (i.e., grades) for the standard course are examined and associated with test scores, correlations between test scores and course grades can only be developed for students placed in the standard-level course. Thus, due to prior placement, the range of the test scores is restricted. Moreover, if the placement test effectively identifies high-risk students, there will be few students in the standard course who earn poor grades; therefore, the range of course grades will also be restricted. The magnitude of correlation coefficients is directly related to the degree of variability in the measures of interest. Thus, correlation coefficients will be smaller than those that would be obtained if all tested students were allowed to enroll in the standard-level course. In addition, as the accuracy of placement increases, the correlation decreases. A low correlation for placement and admissions tests is often perceived as evidence of invalidity, when it could, in fact, be the exact opposite.

Correlational and linear regression results are based on several assumptions. The conditional mean grade is assumed to be a linear function of test scores, grades and test scores are assumed to have the same variance, and the variance of the conditional distribution of grades, given test scores, is assumed to be constant throughout the score range. One or more of these assumptions is usually violated. Further, linear regression can yield predicted grades that are outside the range of grades (i.e., less than 0 or greater than 4, assuming a five-point grade scale).

A more significant limitation of correlations is that they do not provide direct information on the effectiveness of a particular placement rule. For example, if a college is using a particular cutoff score for placement into freshman English, then faculty and administrators may be interested in the proportion of students who were correctly placed (i.e., the proportion who scored at or above the cutoff and, in fact, succeeded in the course, and the proportion who scored below the cutoff and who would have failed the standard course had they enrolled in it). A correlation between performance on the placement test and freshman English grades can provide a measure of the strength of the relationship between these variables, but it cannot provide information about the proportion of students correctly placed.

ACT has developed an alternative methodology for evaluating placement systems (Sawyer, 1989). This method uses estimated validity indices generated from logistic regression models and distributions of predictor variables to determine the accuracy of placement decisions. Logistic regression allows for curvilinear relationships and it models directly a student's probability of success in the standard-level course.

Just as in estimating correlation coefficients, the available data are subject to prior selection. For example, when evaluating the relationship between test scores and course grades for a standard course, the data pertain only to those students who enrolled in and completed the standard course, and not to all students who could have taken the course (i.e., the test score range is restricted). With extrapolation, logistic regression allows one to estimate easily and directly the probability of success (e.g., a grade of C or higher; a grade of B or higher) in the standard course, given a particular cutoff score, for all tested students (including those scoring below the cutoff as well as those scoring above the cutoff). One can, for example, estimate the following four proportions for any cutoff score:

1. The proportion of students who scored below the cutoff and who would have failed the standard course had they enrolled in it (*true negative*).
2. The proportion of students who scored below the cutoff but who would have succeeded in the standard course (*false negative*).
3. The proportion of students who scored above the cutoff and actually succeeded in the standard course (*true positive*).
4. The proportion of students who scored above the cutoff but actually failed the standard course (*false positive*).

Placement validation using this methodology relies, in part, on evaluating the proportion of students correctly placed, given the cutoff score used for placement. This proportion of correct decisions, or "accuracy rate," is defined as the sum of the proportions of true positives and true negatives. Alternative cutoff scores can also be examined by evaluating the proportion of students that would be correctly placed, given particular test score values.

An illustration of a logistic regression function is provided in Figure 1. The estimated probability of success in a standard-level course, given a placement test score, is shown for one institution. The placement test score is displayed on the horizontal axis and the probability of earning a grade of C or higher is displayed on the vertical axis. As shown in the figure, the estimated probability of success increases as the placement test score increases.

Figure 2 illustrates the relationships between the cutoff score used for placement and the estimated accuracy rate, the estimated success rate, and the proportion selected in the standard-level course for this same institution. (The estimated success rate is defined as the proportion of true positives divided by the sum of the proportions of true and false positives.) The proportion selected decreases as the placement test score increases. Conversely, the estimated success rate increases as the placement test score increases. The estimated accuracy rate also increases as the placement test score increases, but achieves a maximum value around a score of 40 and then begins to decrease. This shows that with respect to accuracy rate, the optimal placement test cutoff score is about 40.

Estimated validity indices are useful for evaluating placement systems. ACT is developing a service that will, through the use of estimated validity indices, provide information on the effectiveness of placement systems of individual colleges and universities. The service, for example, might use estimated accuracy rates to help an institution identify the optimal cutoff score for a particular course.

Because validity indices are estimates, it is important to examine them to ensure that they are accurate. It is important to know, for instance, how an estimated accuracy rate based on a truncated distribution of test scores and course grades (i.e., one in which placement has occurred) compares to the same statistic based instead on a full distribution (i.e., one in which there has been no prior placement). If there is little difference between the two estimated accuracy rates, then this would suggest that these indices can be used effectively to evaluate placement practices for courses in which placement has already occurred. The purpose of this study is to investigate the accuracy of estimated validity indices based on truncated data distributions.

There are several techniques that could be used to investigate the effects of truncation on estimated validity indices. One could, for example, use an analytical method, but the mathematics required would be extremely complex. Another method would be to simulate the occurrence of truncation, using computer-generated data. This type of simulation study is currently being conducted by ACT, and the results should be available in the fall of 1992.

The present study uses a different method to examine the effects of truncation. The occurrence of truncation is simulated, but the data used are actual data, gathered from students at postsecondary institutions.

Data

The ASSET system was designed to assist in educational advising, course placement, and retention planning for students entering two-year postsecondary institutions. The ASSET Basic Skills tests measure students' basic skills and knowledge in writing, reading, and mathematics. The Advanced Mathematics tests measure more advanced mathematical skills and knowledge in elementary, intermediate, and college algebra. Scores for the ASSET tests are reported on a scale ranging from 23 to 55.

In fall, 1988, the ASSET Basic Skills tests were administered to entering freshmen from 23 postsecondary institutions. These institutions were randomly selected from the population of all ASSET user institutions. The sample was stratified by geographical region, with the probability of selection proportionate to the size of the institution. Therefore, the sample represented ASSET user institutions from all six regions across the nation (east, southeast, midwest, southwest, mountain/plains, west) and those ranging in size from 1,000 to more than 25,000 students. Most of the institutions were public institutions and offered two-year degree programs.

The Basic Skills tests were administered to over 15,000 students. Every third student in the sample received the same test (i.e., either Writing Skills, Reading Skills, or Numerical Skills); therefore, the sample size for each test was about 5,000. Institutions provided fall (1988) semester grades for tested students who were enrolled in four specific standard-level courses: accounting, history, psychology, and biology. Across institutions, the median numbers of students enrolled in these courses were 22, 72, 55, and 40, respectively.

The data used in this study came from 4 of the 23 institutions, for reasons explained in the following section, and pertained to courses in accounting, history, and psychology. Consequently, the participating institutions may not be representative of all two-year postsecondary institutions, or of ASSET user institutions nationwide. The results of this study therefore may not be generalizable to all two-year institutions and courses.

Method

Sawyer (1989) used logistic regression to determine the accuracy of ACT Assessment scores and high school course grades for college course placement. In this study, the conditional probability of success, given test scores, was estimated using a logistic regression function:

$$P\ [\text{Success} \mid X = x] = (1 + e^{-(a + bx)})^{-1},$$

where x is a particular value of the test score X, and where a and b are the model parameters. These parameters were estimated using the SAS (1990) LOGISTIC procedure.

Estimated validity indices are a function of the conditional probabilities estimated from a logistic model and the distribution of the predictor variable(s) in the relevant population. For example, the proportion of true positives can be estimated as:

$$P\ [\text{Success} \mid X \geq x_o] = \sum_{X \geq x_o} P\ [\text{Success} \mid X = x]\ f(x)$$

for a particular cutoff score x_o, where P [Success | X = x] is the estimated conditional probability and $f(x)$ is the distribution of the predictor variable(s) (e.g., ASSET test scores for students enrolled in accounting at a particular institution).

Because prior selection had not occurred in the standard-level courses, the full distributions of students' test scores and grades were available to estimate the probability of success. Course success (defined as a grade of C or higher) was predicted from the relevant ASSET test score, by institution. With one exception (discussed below), only models with statistically significant ($p < .05$) regression coefficients were retained for further analysis. The estimated probabilities yielded by the logistic regression models were used in combination with distributions of predictor variables to calculate, for each institution, estimated accuracy rates and success rates.

These procedures were repeated using truncated distributions of students' test scores and grades instead of full distributions. At a truncation score of 37, for example, only the records of students scoring at or above 37 were retained and used in the analyses. The truncation scores varied from 31 to 47, and encompassed a broad range of ASSET cutoff scores.

Of the 23 participating institutions, 1 had statistically significant regression coefficients across a wide range of truncation scores for its accounting course, 2 had statistically significant regression coefficients across a wide range of truncation scores for history, and 1 had statistically significant regression coefficients across a wide range of truncation scores for psychology. No institutions were identified as meeting these criteria for biology courses.

Accuracy rates and success rates based on the truncated distributions (denoted AR_{tr} and SR_{tr}, respectively) were estimated for the full range of ASSET scores from each data set. For example, students enrolled in accounting at one institution had ASSET Numerical Skills scores that ranged from 34 to 55. No students at this institution received ASSET scores of 35 or 50. In this case, AR_{tr} and SR_{tr} were estimated for the full range of ASSET scores, excluding 35 and 50 (see Table 1).

The estimated AR_{tr} and SR_{tr} model the situation an institution would encounter if its actual cutoff score were equal to a particular truncation score, and the institution wanted to investigate alternative cutoff scores. For instance, consider an institution that is presently using a cutoff score of 43 on the ASSET Numerical Skills test for placement into accounting. The institution has data for only those students who achieved a Numerical Skills score greater than or equal to 43. Accuracy rates and success rates associated with potential cutoff scores below 43 may be extrapolated from the data, however, and can be examined to determine whether a cutoff below 43 would be advantageous (e.g., whether it would likely result in a larger proportion of correct placement decisions, compared to the present cutoff).

The estimated accuracy rates and success rates based on the full distributions of students' scores and grades (denoted AR_f and SR_f, respectively) were compared to those based on the truncated distributions (AR_{tr}, SR_{tr}). Differences (AR_δ) between the two types of accuracy rates were computed in the following manner:

$$AR_\delta = AR_{tr} - AR_f.$$

A similar calculation was performed for the success rates. Mean differences were calculated, and means of the absolute values of the differences also were calculated. These latter statistics were calculated by determining the absolute value of each AR_δ or SR_δ and then computing a mean of the absolute values. The mean of the absolute values of the AR_δ, for example, may be expressed as

$$1/n \sum_{i=1}^{n} | AR_{\delta_i} |.$$

When interpreting the results of this study, it is important to remember that validity indices based on the full distributions are themselves estimates. These indices therefore are subject to error, particularly for institutions with small samples.

Results

Accounting

The effects of truncation for students enrolled in accounting at one institution are displayed in Figure 3. The estimated conditional probabilities of earning a C or higher grade in accounting, given the ASSET Numerical Skills score, are shown for the full distribution and for four truncated distributions of students' Numerical Skills scores and accounting grades. A fifth truncated distribution was also examined, but it yielded conditional probabilities so similar to those of the full distribution that it was not included in Figure 3. Differences (AR_δ and SR_δ) for this truncated distribution are reported, however, in Table 1 (described below).

The thick, solid line in Figure 3 represents the estimated conditional probabilities based on the full distribution of test scores and grades. This distribution was then truncated at ASSET Numerical Skills scores of 38, 41, 42, and 44. The resulting conditional probabilities are shown by the thin, dashed lines. Note that ASSET scores for this institution ranged from 34 to 55, and that the estimated conditional probabilities are plotted, for each truncation score, across this range of scores.

The graphs in Figure 3 indicate that the accuracy of the conditional probabilities decreased as the degree of truncation increased (i.e., as the cutoff score value increased). A truncation score of 38, for instance, yielded conditional probabilities that were very similar to those of the full distribution. In contrast, the conditional probabilities at other truncation scores (e.g., 42 and 44) were dissimilar to those of the full distribution.

Figure 4 illustrates the effects of truncation on the estimated accuracy rates and success rates for accounting. The thick, solid line in Figure 4 represents the estimated accuracy rate based on the full distribution of test scores and grades. The thick, dashed line represents the estimated success rate based on this same distribution. The accuracy rates and success rates based on the truncated distributions are shown by the thin, solid lines and thin, dashed lines, respectively.

Because the accuracy rates and success rates are based on the estimated conditional probabilities (Figure 3), we would expect them also to be affected by truncation. This was indeed the case: The graphs in Figure 4 indicate that as the degree of truncation increased, the precision of the estimates of the accuracy rate and success rate decreased. At a truncation score of 38, for example, AR_{tr} was similar to AR_f, but at a truncation score of 44, the differences between AR_{tr} and AR_f increased considerably. Moreover, the differences were larger near the minimum and maximum ASSET Numerical Skills scores, relative to scores near the center of the distribution. For example, at a Numerical Skills score of 41, the absolute values of the differences between AR_f and each AR_{tr} were fairly small, as indicated by the proximity of the five lines, ranging from .01 to .02. At a score of 55, on the other hand, the absolute values of the differences were larger, ranging from .01 to .20. These findings were also true for SR_f and each SR_{tr}. At a Numerical Skills score of 45, for example, absolute values of the differences between SR_f and each SR_{tr} were smaller (.00 to .01) than they were at a score of 55 (.00 to .09).

In Figure 4, the slope of each curve representing the AR_{tr} increases until a maximum AR_{tr} is achieved, then begins to decrease. Provided that the slope of the AR_{tr} curves is not constantly increasing, the estimated maximum value of the AR_{tr} corresponds to the optimal cutoff score for accounting, given a particular truncation score. For example, at a truncation score of 44, the estimated maximum AR_{tr} corresponds to an ASSET Numerical Skills score of about 46. For the full distribution, on the other hand, the estimated optimal cutoff score is about 41. Therefore, truncation was associated with overestimation of the optimal cutoff score. Moreover, the maximum AR_{tr} overestimated the maximum AR_f.

Differences between AR_f and each AR_{tr}, and between SR_f and each SR_{tr} are provided for accounting in Table 1. Differences for the same truncation scores as those in Figure 4 are reported with the addition of differences for a truncation score of 37. Note that some ASSET scores (e.g., 50) are not listed in the first column of Table 1. This occurs because no students at this particular institution received these scores.

Table 1 also contains the estimated accuracy rates and success rates for the full distribution of students' Numerical Skills scores and accounting grades. At a Numerical Skills score of 42, for example, AR_f and SR_f were .66 and .70, respectively. When these proportions were compared to those based on a distribution truncated at a score of 41, the differences (AR_δ, SR_δ) were .03 and -.02, respectively. At a truncation score of 44, in comparison, AR_δ and SR_δ were .04 and -.11, indicating that the estimates were less precise at a larger degree of truncation. Note that the signs (+, -) of the AR_δ and SR_δ indicate whether the AR_{tr} and SR_{tr} over- or underestimated the AR_f and SR_f. A positive value corresponds to overestimation of the AR_f or SR_f; a negative value corresponds to underestimation.

The average AR_δ and SR_δ across ASSET Numerical Skills scores are given at the bottom of Table 1. Typically, the AR_{tr} overestimated the AR_f for each truncation score, and the extent of overestimation increased as the truncation score increased. At a truncation score of 37, for example, the average AR_δ was smaller (.00) than at a truncation score of 44 (.08). Similar results were found for the estimated success rates.

The average of the absolute values of the AR_δ and SR_δ also are shown at the bottom of Table 1, in the row labelled "Mean $|\delta|$." The means of the absolute values of the AR_δ and SR_δ ranged from .00 to .16 and from .00 to .11, respectively, for accounting. In addition, they increased as the truncation score increased.

The means of the $|AR_\delta|$ and $|SR_\delta|$ are helpful in determining the accuracy of the AR_{tr} and the SR_{tr}, without regard to over- or underestimation. They will be discussed further in a section describing the accuracy of estimates.

Table 2 contains cumulative relative frequencies (CRFs) of ASSET Numerical Skills scores. For each truncation score used for accounting, corresponding CRFs are reported, along with a corresponding sample size. For the full distribution of Numerical Skills scores, for example, 61% of the students received a score of 45 or lower. The sample consisted of 49 students. When the distribution was truncated at a score of 41, 49% of the students received a score of 45 or lower and the sample size decreased to 37. At the largest truncation score (44), the sample size decreased to 26.

<u>History</u>

<u>Institution A.</u> The effects of truncation for history are illustrated for one institution (Institution A) in Figures 5 and 6. The distribution of history grades and ASSET Reading Skills scores was truncated; statistically significant regression coefficients were found when truncation scores of 31, 34, 36, and 43 were used. As occurred for Accounting, the estimated conditional probabilities (Figure 5) and the estimates of the accuracy rate and success rate (Figure 6) decreased in accuracy as the truncation score increased. In addition, the differences between AR_f, SR_f, and each corresponding AR_{tr} and SR_{tr} were larger near the minimum and maximum ASSET scores, compared to ASSET scores near the center of the distribution (e.g., between about 35-43). The maximum AR_{tr} overestimated the maximum AR_f at extreme degrees of truncation (e.g., 36 and 43). Furthermore, the estimated optimal

cutoff scores themselves exceeded the estimated optimal cutoff score associated with the full distribution.

Accuracy rate and success rate differences for history are reported for Institution A in Table 3. Results are reported for several truncation scores that, because of the similarity of their results to those of other truncation scores, were not included in Figure 6 (33, 35, 37, 38). All average AR_{δ} were positive, suggesting that the AR_{tr} typically overestimated the AR_{f}. The mean of the absolute values of the AR_{δ} increased as the truncation score increased, with exceptions occurring at truncation scores of 37 and 38. The mean of the absolute values of the SR_{δ} also increased as the truncation score increased, with one exception occurring at a score of 38. As was found for accounting, these results indicate that accuracy rates and success rates generally were estimated with less accuracy as truncation increased.

CRFs and sample sizes are reported for history in Table 4. Sample sizes corresponding to the truncation scores ranged from 104 (full distribution) to 44 (truncation score = 43).

Institution B. The effects of truncation for history are illustrated for another institution (Institution B) in Figures 7 and 8. Statistically significant regression coefficients were found when truncation scores of 34, 35, 36, and 37 were used. The lines in Figure 7 are close together, suggesting that there was little difference between conditional probabilities based on the full distribution and those based on the truncated distributions. This is probably due, in part, to the relatively small range of the truncation scores.

The estimates of the accuracy rate and success rate (Figure 8) decreased only slightly in accuracy as the truncation score increased. The differences between AR_{f}, SR_{f}, and each corresponding AR_{tr} and SR_{tr} were largest above a Reading Skills score of 43. Note that the maximum value of the AR_{f} was associated with the maximum Reading Skills score (51). In this case, no optimal Reading Skills cutoff score can be identified. The maximum AR_{tr} overestimated the maximum AR_{f} for most truncation scores.

Because of their proximity, the individual graphs in Figure 8 are not labelled according to truncation score. The same truncation scores that are represented in Figure 7 are also represented in this figure, however.

Accuracy rate and success rate differences for history are reported in Table 5 for Institution B. Results are reported for an additional truncation score (33) that was not illustrated in Figures 7 and 8. As occurred for Institution A, the AR_{tr} typically overestimated the AR_f, but to a smaller degree: The mean of the absolute values of the AR_δ and SR_δ did not exceed .01 for any truncation score. This suggests that these estimates were more precise, compared to those for Institution A.

Table 6 contains CRFs and sample sizes for history (Institution B). The sample sizes for this institution were smaller, in general, than those of Institution A, ranging from 55 to 62. The number of student records varied little across truncation scores. For example, at a truncation score of 33, 61 student records were included. The full distribution, in comparison, contained 62 student records.

Psychology

The effects of truncation on the estimated conditional probabilities is shown in Figure 9 for psychology. Figure 10 shows the effects of truncation on accuracy rates and success rates for this course. At truncation scores of 32, 33, and 35, each AR_{tr} and SR_{tr} differed only slightly from the AR_f and SR_f. At a truncation score of 40, however, the differences in the statistics were greater, particularly for the accuracy rate. In addition, the maximum AR_{tr} at this truncation score overestimated the maximum AR_f, and the estimated optimal cutoff score itself exceeded the estimated optimal cutoff score associated with the full distribution. Note that at a truncation score of 33, slight underestimates of AR_f and SR_f were obtained across most Reading Skills scores, whereas at a truncation score of 35, slight overestimates were obtained.

Table 7 contains accuracy rate and success rate differences for psychology. Results for a truncation score of 31 are also reported. While increases in the means of the absolute values of the AR_δ and SR_δ clearly corresponded to increases in the truncation scores for accounting and history, this trend was less evident for psychology. For example, the mean $|AR_\delta|$ at truncation scores of 31, 32, 33, and 35 were identical (.01). Had there been more truncation scores for which statistically significant regression coefficients were identified, then perhaps trends in the results would be more discernable. Regardless, it is evident that at the most extreme degree of truncation (40) the AR_{tr} and SR_{tr} differed the most from the AR_f and SR_f.

Table 8 contains CRFs and sample sizes for psychology. Sample sizes for this course ranged from 83 (full distribution) to 46 (truncation score = 40).

Accuracy of Estimates

The preceding results indicate that estimated accuracy rates and success rates based on truncated distributions of test scores and grades differ from those based on full distributions. In some cases, particularly for the lowest truncation scores, the loss of accuracy was small and these estimates therefore could be considered acceptable. In other cases, the loss of accuracy was large, suggesting that these estimates would not be acceptable. In Figure 4, for example, a truncation score of 38 for the ASSET Numerical Skills test yielded estimated accuracy rates and success rates for accounting that were similar to those of the full distribution (e.g., the average differences were .01 and .00, respectively), while a truncation score of 44 yielded dissimilar estimates.

One method of determining whether the AR_{tr} and SR_{tr} are sufficiently similar to the AR_f and SR_f is to choose a "threshold" for the mean of the absolute values of the AR_δ and SR_δ. For example, if the absolute values of either the AR_δ or SR_δ differ, on average, by more than .05, then the estimates could be considered unacceptably imprecise. A threshold of .05 seems reasonable; an accuracy rate of .70, for example, could be meaningfully different from an accuracy rate of .76 when an institution is interested in making the largest possible proportion of correct placement decisions.

Accounting. The mean of the absolute values of the AR_δ or SR_δ for accounting did not exceed .05 until a truncation score of 42 was used (Table 1). This indicates that the loss in accuracy of the AR_{tr} was unacceptable at truncation scores greater than or equal to 42. The graphs in Figure 4 confirm this conclusion: The lines representing the AR_{tr} at truncation scores of 38 and 41 are fairly close to the line representing the AR_f. The other AR_{tr}, however, are considerably distant from the AR_f.

The CRFs in Table 2 can assist in determining the minimum proportion of the full distribution needed for accurate estimation of validity indices for accounting. For this particular institution, the estimates of AR_{tr} were noticeably inaccurate when a cutoff score of 42 was imposed (33% of the full distribution was not included). This implies that to achieve accurate estimates, at least two-thirds of the full distribution must be included.

History. It is evident from Table 3 (Institution A) that the mean of the absolute values of the AR_δ or SR_δ first exceeded .05 at a truncation score of 35. Therefore, the loss in accuracy of the estimates for this institution was unacceptable at truncation scores greater than or equal to 35. Table 4 shows that 15% of the students in the full distribution were not included when a truncation score of 35 was used.

Across all truncation scores, the means of the absolute values of the AR_δ and SR_δ did not exceed .05 for Institution B (Table 5). In fact, they did not exceed .01. Thus, the accuracy of the estimates for this institution was acceptable across all truncation scores. This institution, however, had a smaller range of truncation scores for which statistically significant regression coefficients were identified, relative to that of Institution A. Moreover, the decrease in the number of student records at each truncation score was smaller, compared to that of Institution A. For example, the maximum percentage of student records in the full distribution that were not included when the data for Institution B were truncated was 11% (truncation score = 37). This was considerably smaller than the maximum percentage for Institution A (58% at a truncation score of 43). Had larger truncation scores been used and/or had greater decreases in sample size occurred, then perhaps the findings for Institution B would more closely resemble those of Institution A.

Psychology. A loss in accuracy did not become very noticeable until a truncation score of 40 was used. At this truncation score, the means of the absolute values of the AR_δ and SR_δ were .18 and .15, respectively (Table 7). The CRFs in Table 8 indicate that 45% of the students in the full distribution for psychology were not included at this truncation score.

The estimated accuracy rates and success rates were imprecise for accounting, history, and psychology at different degrees of truncation. The CRFs at which the accuracy of the estimates was unacceptable varied across these three courses, ranging from 15 to 45. It appears, therefore, that the loss in accuracy of the estimates may be related to such factors as sample size, institution, and course.

Statistical Significance of Regression Coefficients

All logistic regression models had statistically significant ($p < .05$) regression coefficients, with the exception of the model based on the full distribution of students' history grades and ASSET Reading Skills scores for Institution A. The coefficient associated with Reading Skills score for this model had a p-value of .069.

When the full distribution of history grades and Reading Skills scores was truncated at a score of 31, the number of student records included in the analysis decreased from 104 to 101, but the resulting logistic regression model had a statistically significant regression coefficient associated with Reading Skills score. In fact, truncating the distribution at a score of 29 resulted in a loss of only two student records, and the resulting model still had statistically significant regression coefficients. The inclusion of two particular student records, therefore, prevented the model based on the full distribution from meeting the criterion of statistical significance.

Further examination of the records of these two students revealed that their Reading Skills performance was low; they each earned a score of 28. Only 2% of students nationwide earn Reading Skills scores of 28 or below. Contrary to what we might expect based on their Reading Skills performance, the two students both received passing grades in history (one student received a B and the other received a C). However, an outlier analysis for the full distribution of history grades and ASSET scores did not identify these particular observations, or any others, as statistically significant ($p < .01$) outliers. It therefore seemed reasonable to include all 104 observations when developing the logistic regression model based on the full distribution of test scores and history grades.

These findings suggest that the statistical significance of regression coefficients in logistic regression models may be determined, in some instances, by a very small proportion of student records. In the case of the model developed for the full distribution of history grades and ASSET scores, there seemed to be little reason for not accepting it as a useful model even though it was not statistically significant at the .05 level. The estimated accuracy rates and success rates based on this model were nearly identical to those based on an alternative model. For example, means of the absolute values of the AR_δ and SR_δ for a truncation score of 29 were both small (.02 and .01, respectively), with a loss of only two student records.

Conclusions

The findings of this study suggest that when distributions of grades and test scores are truncated, as occurs when students are placed into a course on the basis of a cutoff score, the estimated accuracy rates and success rates differ from those obtained when the full data distribution is used. In general, the greater the degree of truncation (i.e., course selectivity), the less accurate are the estimated accuracy rates and success rates. Estimated maximum accuracy rates are typically overestimated, and the extent of overestimation increases as the degree of truncation increases. In addition, the estimated optimal cutoff scores themselves tend to be overestimated when truncation is extreme. The loss in accuracy of estimated validity indices due to truncation implies that these statistics should be accompanied by suitable estimates of variability, such as confidence intervals. ACT is presently planning research to develop such estimates.

The estimated accuracy rates and success rates were acceptably accurate when less than 15% of the full distribution of students' test scores and course grades was truncated. Greater degrees of truncation often resulted in unacceptably imprecise estimates. This finding has implications for using estimated validity indices to evaluate placement systems. For example, consider a placement test cutoff score that results in placing 48% of an institution's entering freshmen into a lower-level course. Complete data are available, in this case, for only those students who enrolled in and completed the standard-level course (representing 52% of the original sample). The distribution of these data may, unfortunately, be truncated to the extent that estimates of validity indices will not be sufficiently accurate.

This study identified only a small number of institutions with statistically significant logistic regression models across a wide range of truncation scores. Consequently, the results should be interpreted cautiously and confirmed through future research using a larger number of institutions and courses.

References

SAS Institute, Inc. (1990). SAS/STAT User's Guide, Version 6 (4th ed., Vol. 2). Cary, NC: SAS Institute Inc.

Sawyer, R. (1989). Validating the use of ACT Assessment scores and high school grades for remedial course placement in college. (ACT Research Report No. 89-4). Iowa City, Iowa: The American College Testing Program.

TABLE 1

Effects of Truncation, Across ASSET Numerical Skills Scores, on Estimated Validity Indices for Accounting

ASSET cutoff score	Full distribution		Truncation score									
			37		38		41		42		44	
	AR	SR	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ
34	.61	.61	.00	.00	-.01	-.01	-.06	-.06	-.14	-.14	-.19	-.19
36	.63	.63	.00	-.01	-.01	-.01	-.05	-.06	-.12	-.14	-.17	-.19
37	.64	.63	.00	.00	-.01	-.01	-.04	-.05	-.11	-.13	-.16	-.18
38	.65	.65	.00	-.01	-.01	-.01	-.03	-.05	-.08	-.13	-.13	-.18
39	.66	.66	.00	-.01	-.01	-.01	-.02	-.05	-.06	-.12	-.10	-.17
40	.66	.66	.00	.00	.00	-.01	-.01	-.04	-.05	-.11	-.08	-.16
41	.66	.68	.00	.00	.01	.00	.01	-.03	.01	-.09	-.02	-.14
42	.66	.70	.00	.00	.01	.00	.03	-.02	.05	-.06	.04	-.11
43	.65	.72	.00	.00	.01	.00	.05	-.01	.09	-.04	.10	-.07
44	.64	.74	.00	.00	.01	.00	.05	-.01	.12	-.02	.14	-.05
45	.62	.76	.00	.00	.02	.00	.06	.00	.14	.01	.18	-.01
46	.60	.78	.00	.00	.02	.00	.07	.01	.16	.02	.21	.02
47	.58	.79	.00	.00	.01	.01	.07	.02	.16	.04	.22	.05
48	.54	.82	.01	.00	.02	.00	.08	.02	.17	.06	.23	.07
49	.52	.83	.01	.00	.02	.01	.07	.03	.17	.07	.23	.09
51	.49	.86	.00	-.01	.01	.00	.07	.03	.16	.07	.22	.09
52	.46	.87	.01	.00	.02	.01	.07	.03	.16	.08	.21	.10
53	.43	.88	.01	.00	.02	.01	.07	.04	.15	.08	.21	.10
54	.42	.89	.00	.00	.01	.01	.06	.04	.14	.08	.20	.09
55	.40	.90	.01	.00	.02	.01	.06	.03	.15	.07	.20	.09
Mean			.00	-.00	.01	.00	.03	-.01	.06	-.02	.08	-.04
Mean $\lvert\delta\rvert$			.00	.00	.01	.01	.05	.03	.12	.08	.16	.11

TABLE 2

Cumulative Relative Frequencies of ASSET Numerical Skills Scores for Accounting

ASSET cutoff score	Full distribution	Truncation score				
		37	38	41	42	44
34	4					
36	6					
37	10	4				
38	14	9	5			
39	16	11	7			
40	25	20	16			
41	33	28	25	11		
42	41	37	34	22	12	
43	47	44	41	30	21	
44	55	52	50	41	33	15
45	61	59	57	49	42	27
46	67	65	64	57	52	39
47	76	74	73	68	64	54
48	80	78	77	73	70	62
49	86	85	84	81	79	73
51	90	89	89	87	85	81
52	94	94	93	92	91	89
53	96	96	96	95	94	92
54	98	98	98	97	97	96
55	100	100	100	100	100	100
N	49	46	44	37	33	26

TABLE 3

Effects of Truncation, Across ASSET Reading Skills Scores, on Estimated Validity Indices for History (Institution A)

ASSET cutoff score	Full distribution		Truncation score															
			31		33		34		35		36		37		38		43	
	AR	SR	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ
28	.60	.60	-.01	-.01	-.02	-.02	-.03	-.03	-.07	-.07	-.10	-.10	-.09	-.09	-.07	-.07	-.26	-.26
29	.60	.60	-.01	-.01	-.01	-.01	-.02	-.02	-.06	-.07	-.09	-.09	-.08	-.08	-.05	-.06	-.24	-.26
31	.60	.60	.00	-.01	-.01	-.01	-.01	-.02	-.05	-.06	-.08	-.09	-.07	-.08	-.05	-.06	-.23	-.25
33	.60	.61	.00	-.01	.00	-.01	-.01	-.02	-.04	-.07	-.06	-.09	-.05	-.08	-.03	-.06	-.22	-.26
34	.61	.61	-.01	-.01	-.01	-.01	-.01	-.02	-.04	-.06	-.05	-.08	-.05	-.07	-.03	-.05	-.20	-.25
35	.61	.62	.00	.00	.00	.00	.00	-.01	.00	-.04	-.01	-.06	-.01	-.05	.00	-.03	-.12	-.23
36	.61	.63	.01	.00	.01	.00	.01	-.01	.02	-.03	.02	-.05	.02	-.04	.02	-.03	-.08	-.21
37	.61	.63	.01	.00	.01	.00	.01	.00	.02	-.03	.03	-.04	.03	-.04	.03	-.02	-.06	-.20
38	.60	.64	.02	.00	.02	.00	.02	.00	.05	-.02	.06	-.03	.06	-.03	.05	-.01	.00	-.19
39	.60	.65	.01	.00	.01	.00	.02	.00	.06	-.01	.08	-.02	.07	-.01	.06	.00	.05	-.16
40	.59	.65	.02	.01	.02	.01	.03	.01	.07	.01	.10	.00	.09	.01	.07	.01	.10	-.13
41	.58	.66	.02	.01	.02	.01	.03	.01	.09	.02	.11	.02	.10	.02	.08	.02	.15	-.10
43	.56	.68	.02	.01	.02	.02	.04	.02	.10	.04	.14	.05	.12	.05	.10	.04	.24	-.02
44	.53	.70	.02	.01	.03	.02	.04	.02	.11	.06	.15	.07	.14	.07	.11	.06	.29	.04
45	.53	.70	.02	.02	.02	.02	.04	.03	.11	.06	.15	.08	.13	.08	.10	.07	.30	.06

(continued on next page)

TABLE 3 (continued)

ASSET cutoff score	Full distribution		Truncation score															
			31		33		34		35		36		37		38		43	
	AR	SR	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ
46	.52	.70	.02	.02	.03	.02	.04	.03	.11	.07	.15	.09	.13	.08	.10	.07	.30	.08
48	.49	.72	.02	.02	.02	.02	.04	.03	.10	.08	.14	.10	.13	.09	.10	.08	.31	.12
49	.45	.74	.01	.02	.02	.03	.03	.04	.09	.09	.12	.12	.11	.11	.08	.10	.29	.16
51	.44	.75	.01	.02	.01	.03	.03	.04	.08	.10	.11	.13	.10	.12	.07	.10	.28	.18
53	.42	.76	.01	.03	.01	.04	.02	.05	.08	.11	.11	.14	.09	.13	.07	.12	.27	.20
Mean			.01	.01	.01	.01	.02	.01	.04	.01	.05	.01	.05	.01	.04	.01	.06	-.08
Mean|δ|			.01	.01	.02	.01	.02	.02	.07	.06	.09	.07	.08	.07	.06	.05	.20	.17

TABLE 4

Cumulative Relative Frequencies of ASSET Reading Skills Scores for History (Institution A)

ASSET cutoff score	Full distribution	Truncation score							
		31	33	34	35	36	37	38	43
28	2								
29	3								
31	5	2							
33	8	5	3						
34	15	13	11	8					
35	20	18	16	14	6				
36	22	20	18	16	8	2			
37	28	26	24	22	15	10	7		
38	35	33	31	29	23	18	16	9	
39	39	38	36	34	28	24	22	16	
40	45	44	42	41	35	31	30	24	
41	58	56	56	54	50	47	46	41	
43	67	66	66	65	61	59	58	55	23
44	69	68	68	67	64	61	61	57	27
45	71	70	70	69	66	64	63	60	32
46	80	79	79	78	76	75	74	72	52
48	90	90	90	90	89	88	88	87	77
49	93	93	93	93	92	92	91	91	84
51	97	97	97	97	97	96	96	96	93
53	100	100	100	100	100	100	100	100	100
N	104	101	99	96	88	83	81	75	44

TABLE 5

Effects of Truncation, Across ASSET Reading Skills Scores, on Estimated Validity Indices for History (Institution B)

ASSET cutoff score	Full distribution		Truncation score									
			33		34		35		36		37	
	AR	SR	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ
32	.35	.35	.01	.01	-.01	-.01	-.01	-.01	.00	.00	.00	.00
33	.37	.36	.00	.00	-.01	-.01	-.01	-.01	-.01	-.01	.00	.00
34	.39	.37	.00	.00	-.01	-.01	-.01	-.01	.00	-.01	.00	-.01
35	.40	.37	.01	.00	.00	-.01	.00	-.01	.00	.00	.00	.00
36	.43	.38	.00	.00	-.01	-.01	-.01	-.01	.00	.00	.00	.00
37	.44	.38	.00	.00	-.01	-.01	.00	.00	.00	.00	.00	.00
38	.47	.39	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00
39	.49	.40	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
40	.53	.42	.00	.00	.01	.00	.01	.00	.01	.00	.01	.00
41	.55	.43	-.01	.00	.01	.00	.01	.00	.00	.00	.00	.00
43	.59	.46	.00	.00	.02	.00	.02	.00	.02	.00	.01	.00
44	.63	.49	.00	.00	.03	.01	.02	.01	.02	.01	.01	.00
45	.66	.52	-.01	-.01	.02	.02	.02	.01	.01	.01	.01	.00
46	.66	.52	.00	.00	.03	.02	.02	.02	.01	.01	.01	.01
48	.67	.57	.00	.00	.03	.03	.02	.03	.02	.02	.01	.01
49	.67	.62	.00	-.01	.02	.04	.02	.03	.01	.02	.00	.01
51	.67	.63	-.01	-.01	.02	.04	.01	.04	.01	.02	.00	.02
Mean			-.00	-.00	.01	.01	.01	.01	.01	.00	.00	.00
Mean $\lvert\delta\rvert$			.00	.00	.01	.01	.01	.01	.01	.01	.00	.00

TABLE 6

Cumulative Relative Frequencies of ASSET Reading Skills Scores for History (Institution B)

ASSET cutoff score	Full distribution	Truncation score				
		33	34	35	36	37
32	2					
33	5	3				
34	7	5	2			
35	10	8	5	3		
36	11	10	7	5	2	
37	16	15	12	10	7	6
38	19	18	15	14	11	9
39	27	26	24	22	20	18
40	31	30	27	26	23	22
41	42	41	39	38	36	35
43	55	54	53	52	50	49
44	65	64	63	62	61	60
45	66	66	64	64	63	62
46	81	80	80	79	79	78
48	90	90	90	90	89	89
49	92	92	92	91	91	91
51	100	100	100	100	100	100
N	62	61	59	58	56	55

TABLE 7

Effects of Truncation, Across ASSET Reading Skills Scores, on Estimated Validity Indices for Psychology

ASSET cutoff score	Full distribution		Truncation score									
			31		32		33		35		40	
	AR	SR	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ	AR_δ	SR_δ
28	.77	.77	.00	.00	.01	.01	.02	.02	-.01	-.01	-.23	-.23
30	.78	.78	-.01	-.01	.00	.00	.01	.01	-.01	-.01	-.23	-.23
31	.78	.78	.00	.00	.00	.01	.01	.02	-.01	-.01	-.22	-.23
32	.79	.80	.00	.00	.00	.01	.00	.01	-.01	-.01	-.18	-.22
33	.79	.81	.00	-.01	.00	.00	.00	.01	-.01	-.01	-.17	-.22
34	.79	.82	.00	-.01	.00	.00	.00	.00	.00	-.01	-.14	-.22
35	.79	.83	.00	.00	-.01	.01	-.01	.01	-.01	.00	-.09	-.19
36	.78	.85	.00	.00	-.01	.00	-.02	.00	.00	-.01	-.04	-.17
37	.77	.86	.00	.00	-.01	.00	-.01	.00	.00	-.01	-.01	-.16
38	.75	.86	.01	.00	.00	.00	-.01	.00	.01	.00	.03	-.14
39	.71	.88	.01	.00	-.01	.00	-.02	.00	.01	.00	.12	-.09
40	.67	.90	.01	.00	-.01	.00	-.02	-.01	.01	-.01	.18	-.05
41	.64	.91	.01	.00	-.01	-.01	-.02	-.01	.01	-.01	.20	-.03
43	.59	.92	.01	.00	-.01	.00	-.02	-.01	.01	.00	.23	.00
44	.47	.94	.01	.00	-.01	-.01	-.02	-.01	.01	.00	.25	.02
45	.43	.94	.01	.01	.00	.00	-.02	-.01	.01	.00	.25	.03
46	.41	.95	.01	.00	.00	-.01	-.02	-.01	.01	-.01	.25	.03
48	.31	.96	.00	.00	-.01	.00	-.02	-.01	.01	.00	.23	.03
49	.26	.97	.01	.00	.00	-.01	-.01	-.01	.02	.00	.24	.02
51	.24	.98	.01	.00	.00	-.01	-.01	-.01	.01	.00	.23	.02
Mean			.00	-.00	-.00	-.00	-.01	-.00	.00	-.01	.05	-.10
Mean$\lvert\delta\rvert$			.01	.00	.01	.00	.01	.01	.01	.01	.18	.15

TABLE 8

Cumulative Relative Frequencies of ASSET Reading Skills Scores for Psychology

ASSET cutoff score	Full distribution	Truncation score				
		31	32	33	35	40
28	1					
30	2					
31	7	5				
32	8	6	1			
33	11	9	4	3		
34	17	15	10	9		
35	22	20	16	15	6	
36	24	22	18	17	9	
37	28	26	22	21	13	
38	37	36	33	32	25	
39	45	43	40	40	33	
40	49	48	46	45	39	9
41	57	56	53	53	48	22
43	72	72	70	70	67	50
44	77	77	75	75	73	59
45	80	79	78	78	75	63
46	92	91	91	91	90	85
48	96	96	96	96	96	94
49	99	99	99	99	99	98
51	100	100	100	100	100	100
N	83	81	77	76	69	46

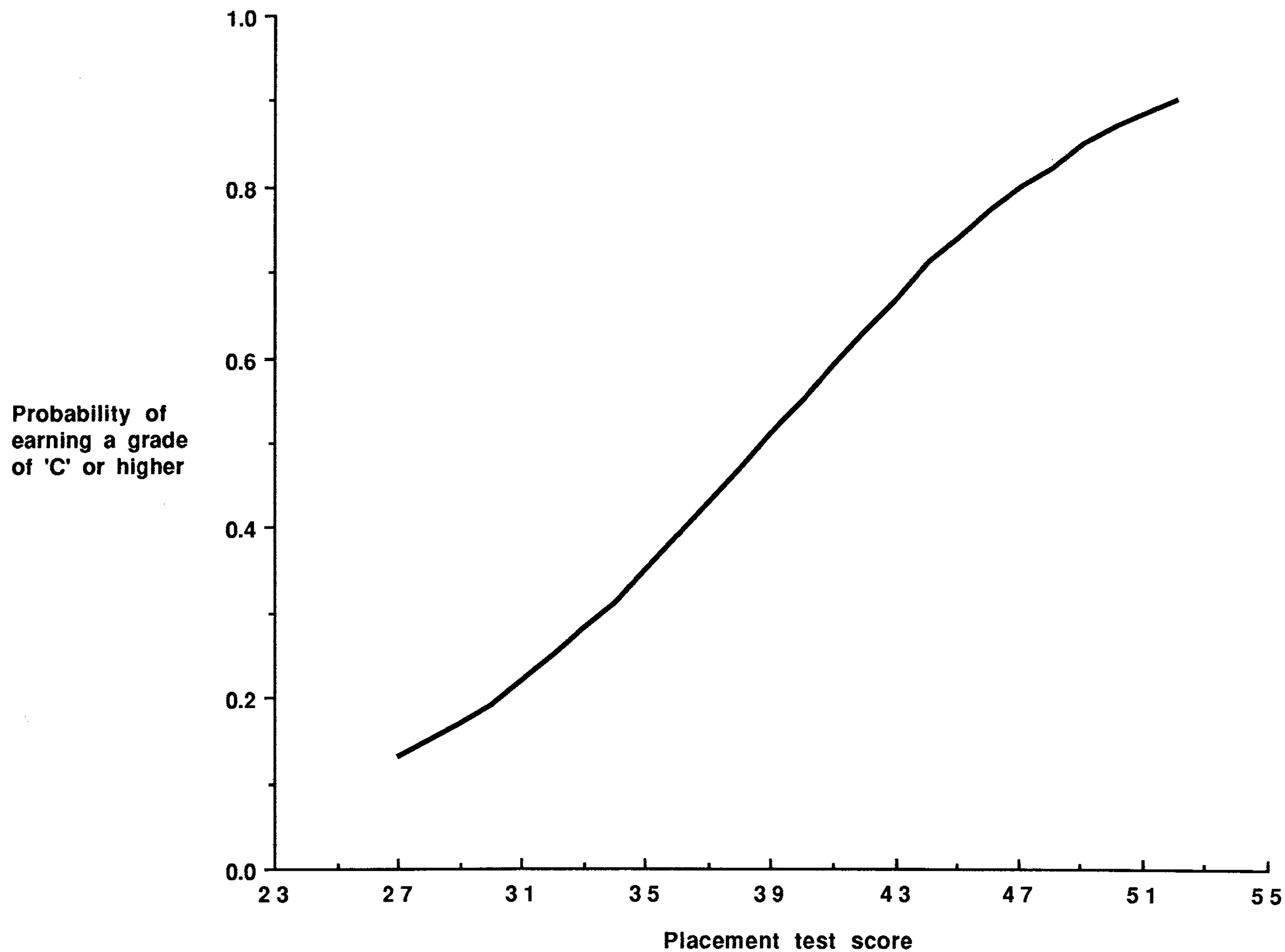

FIGURE 1. Estimated Probability of Success in Standard Course

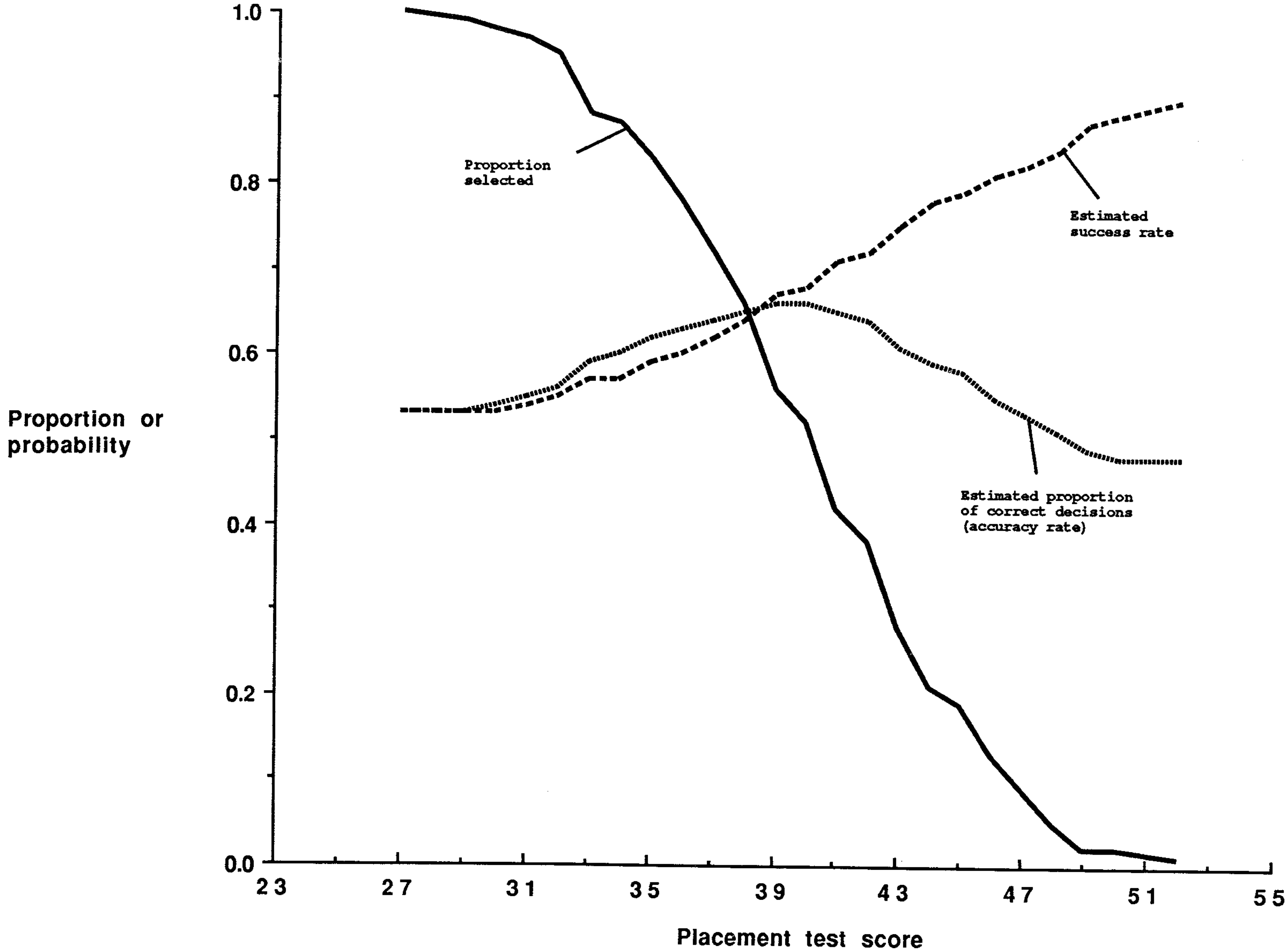

FIGURE 2. Estimated Validity Indices for Standard Course (Criterion for Success = 'C' or Higher)

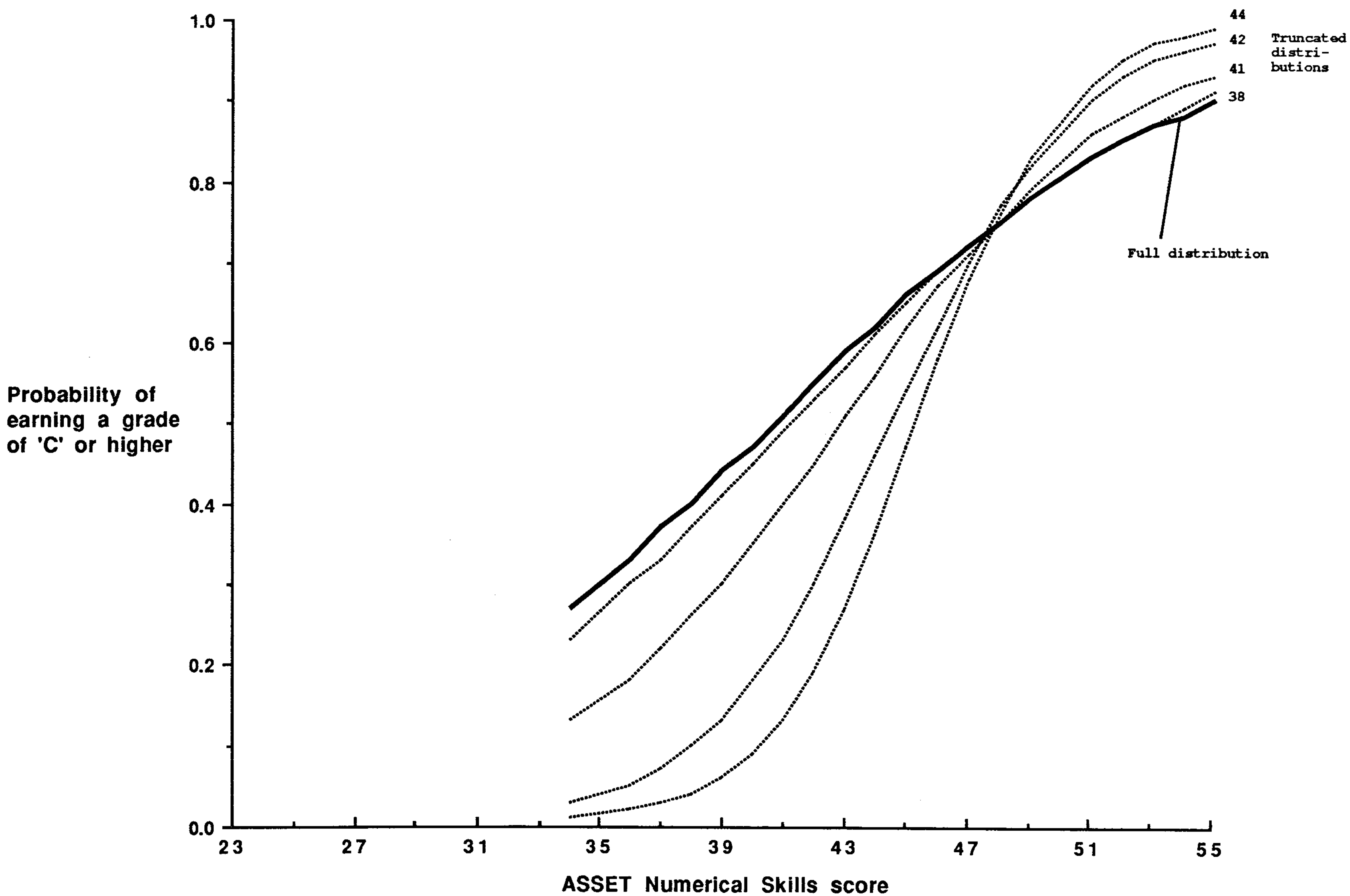

FIGURE 3. Effects of Truncation on Conditional Probabilities for Accounting (Based on Data for One Institution)

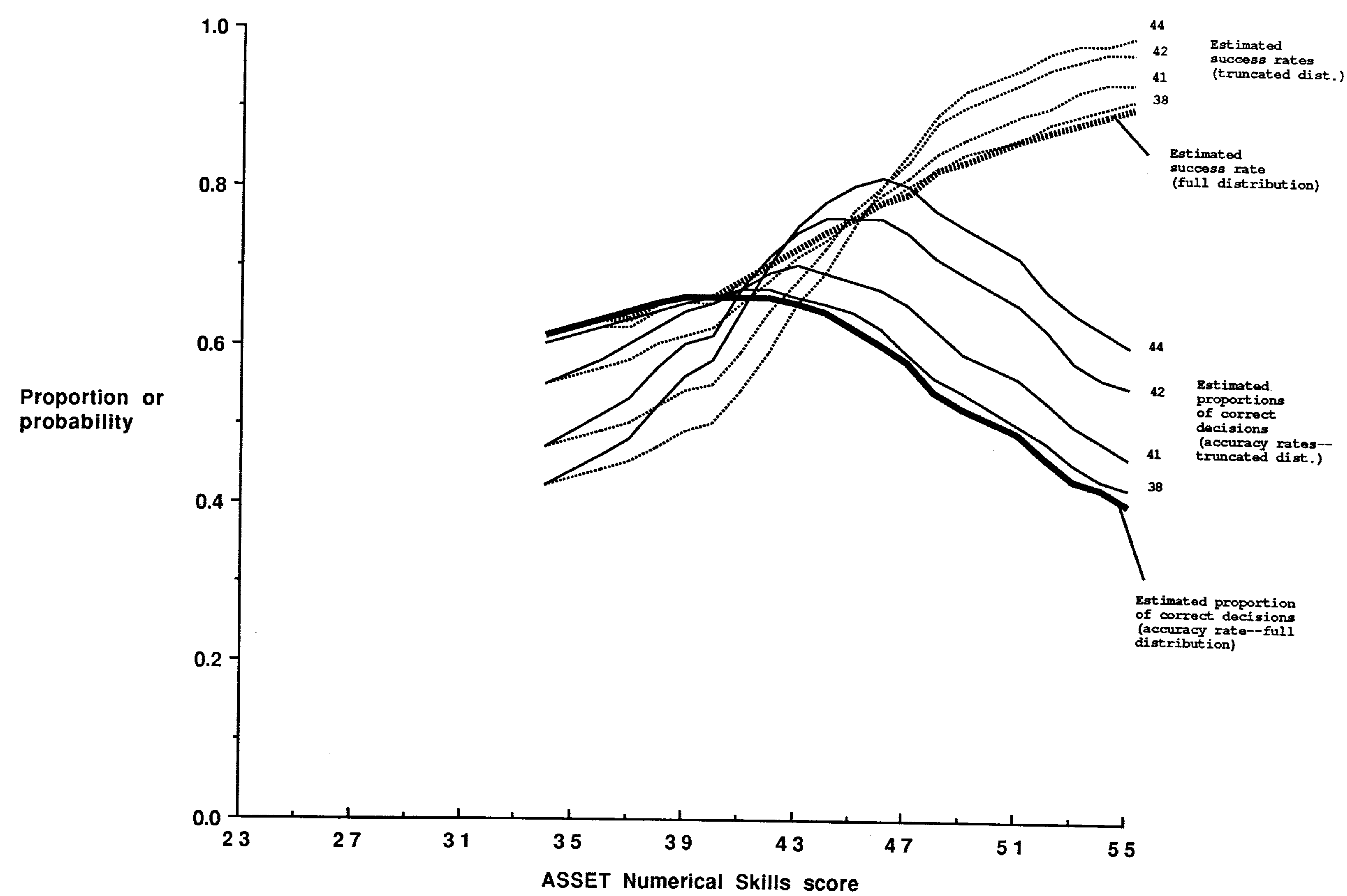

FIGURE 4. Effects of Truncation on Estimated Validity Indices for Accounting (Based on Data for One Institution)

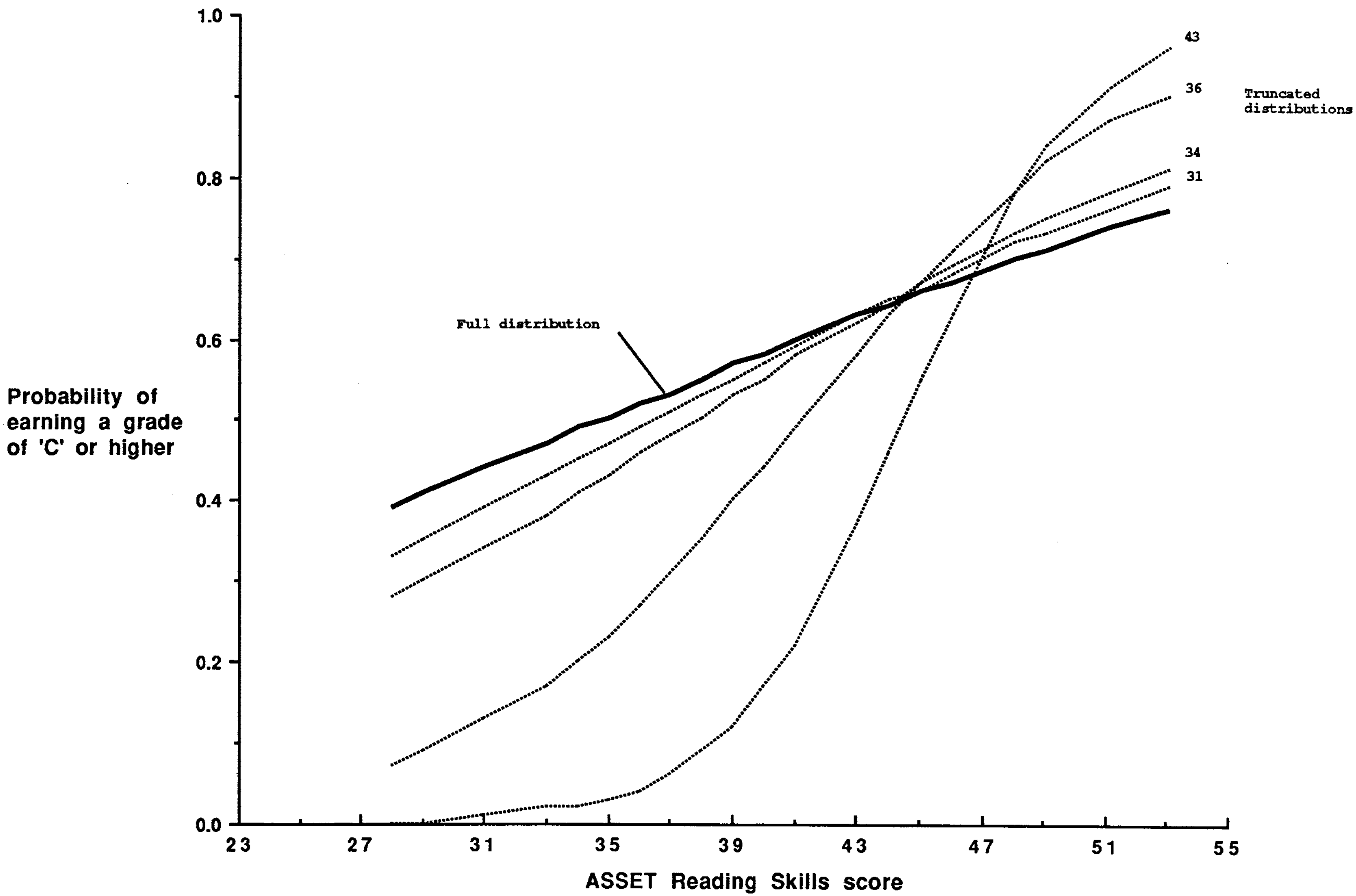

FIGURE 5. Effects of Truncation on Conditional Probabilities for History (Institution A)

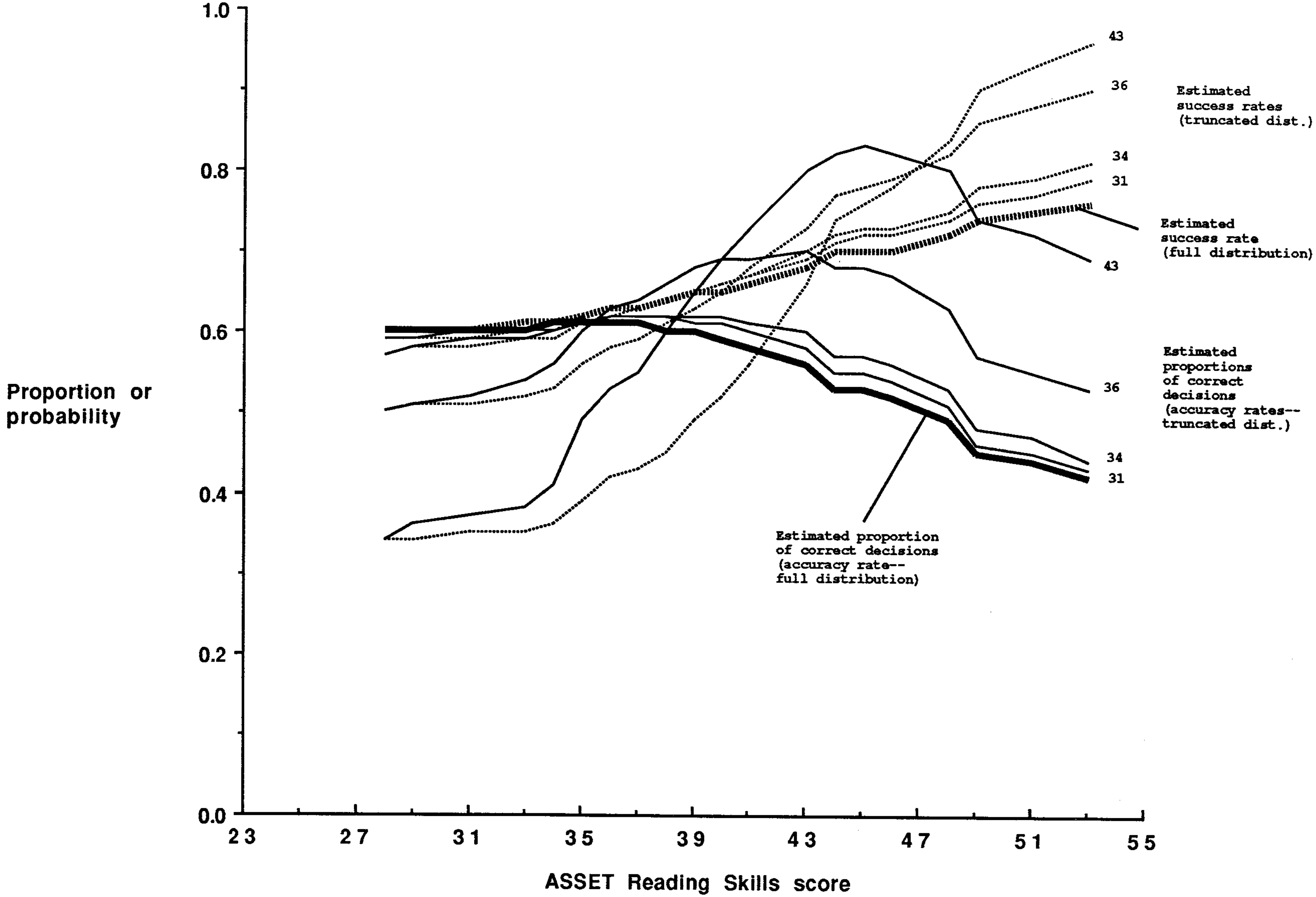

FIGURE 6. Effects of Truncation on Estimated Validity Indices for History (Institution A)

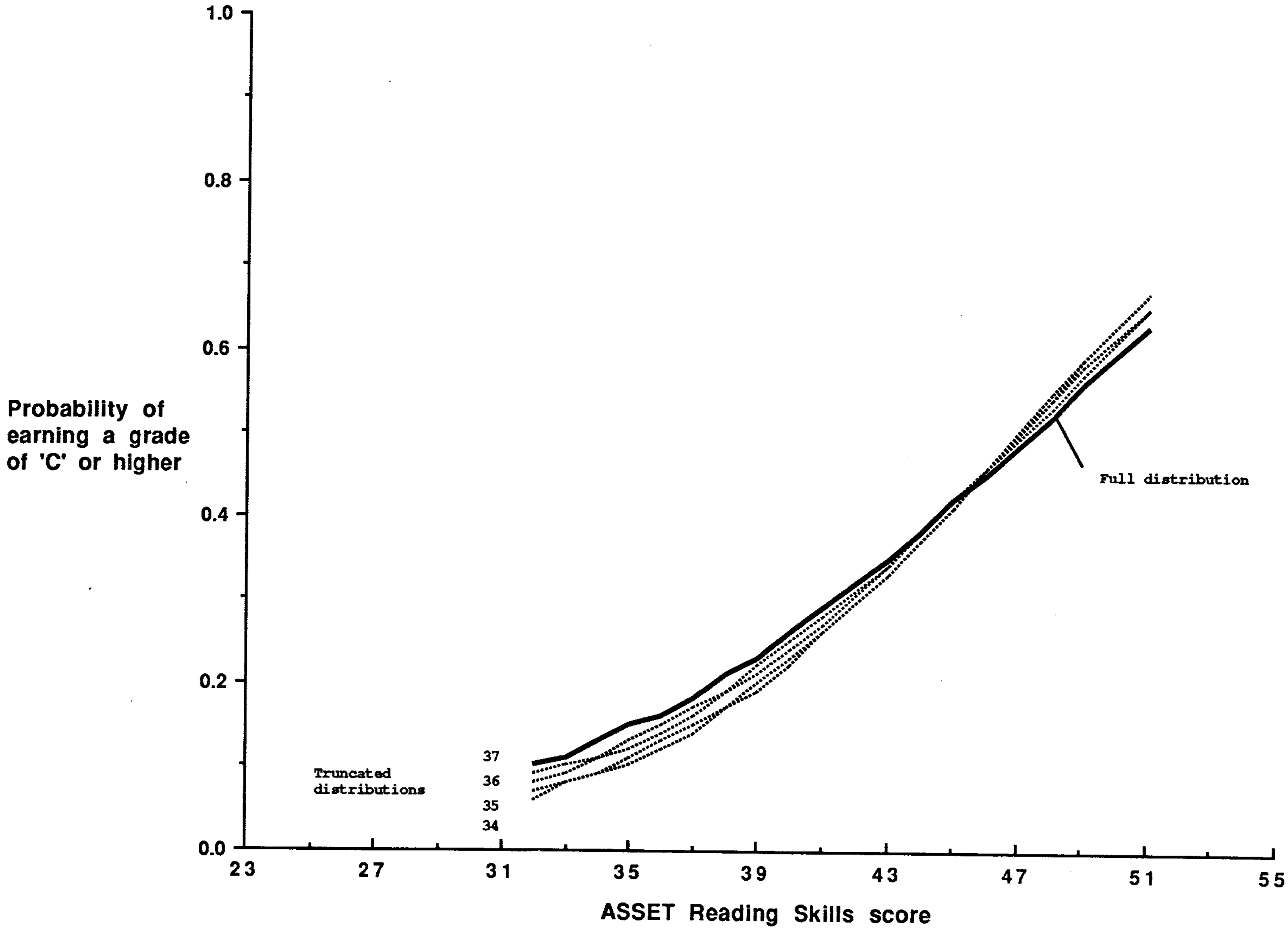

FIGURE 7. Effects of Truncation on Conditional Probabilities for History (Institution B)

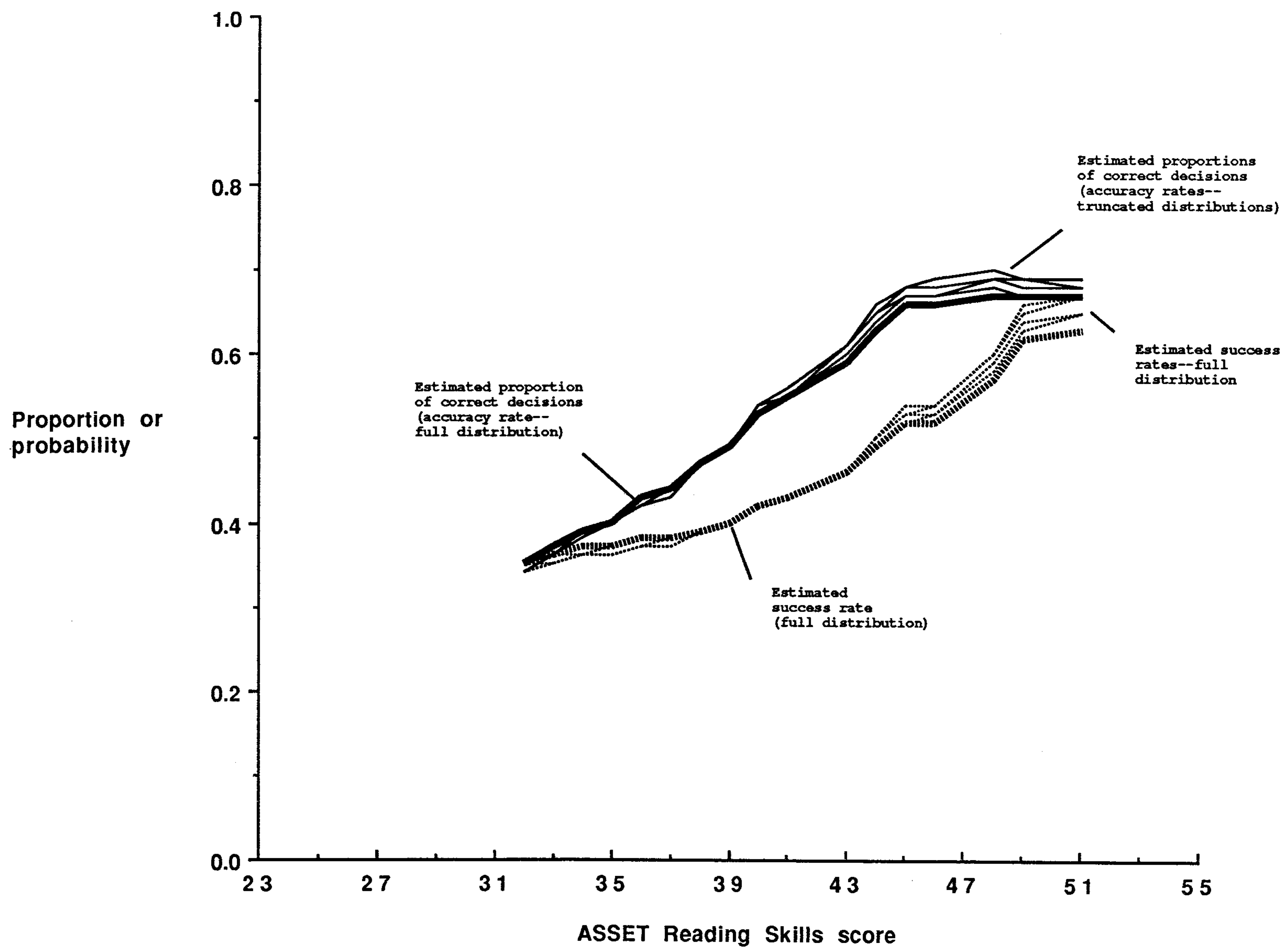

FIGURE 8. Effects of Truncation on Estimated Validity Indices for History (Institution B)

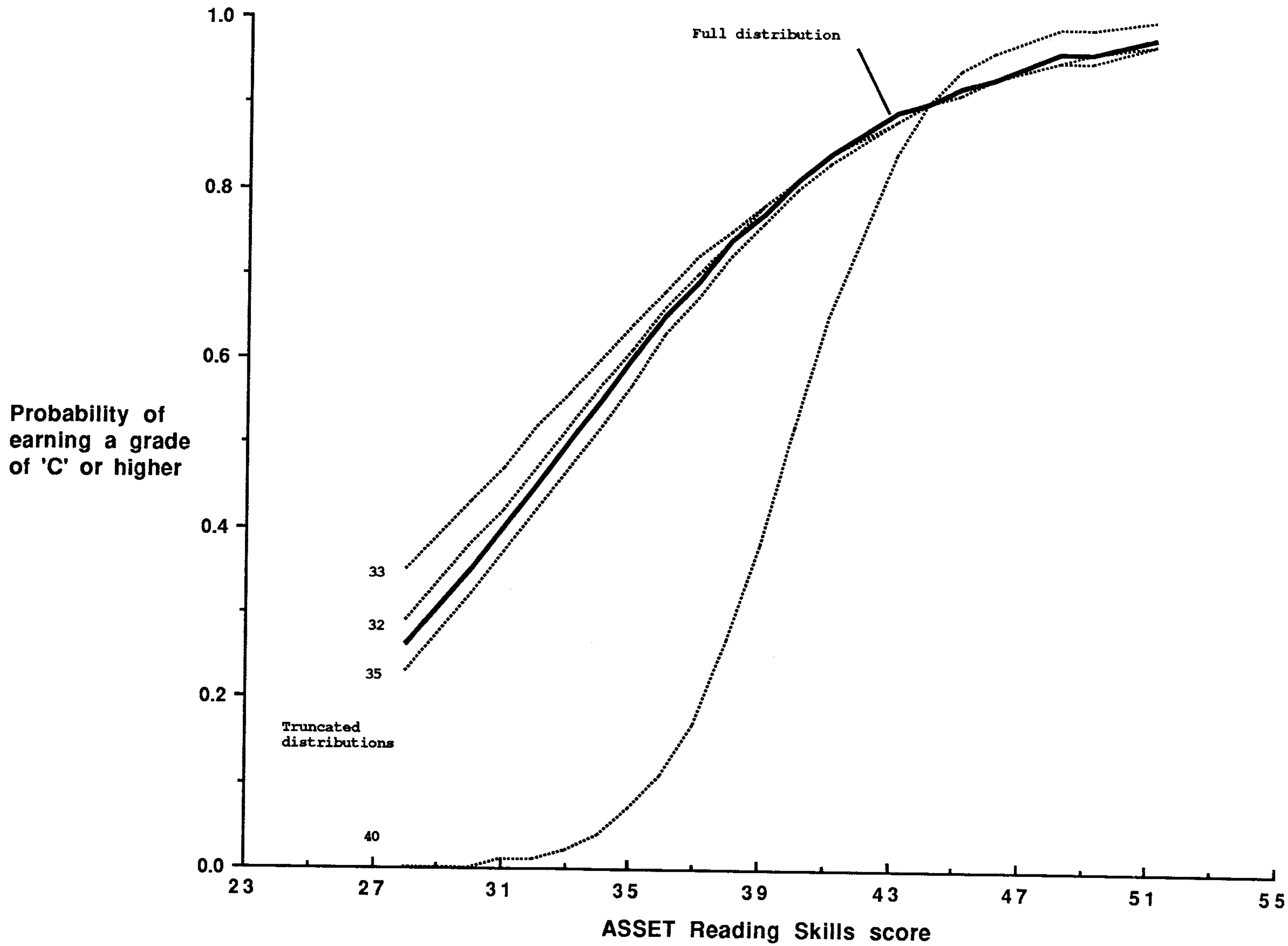

FIGURE 9. Effects of Truncation on Conditional Probabilities for Psychology (Based on Data for One Institution)

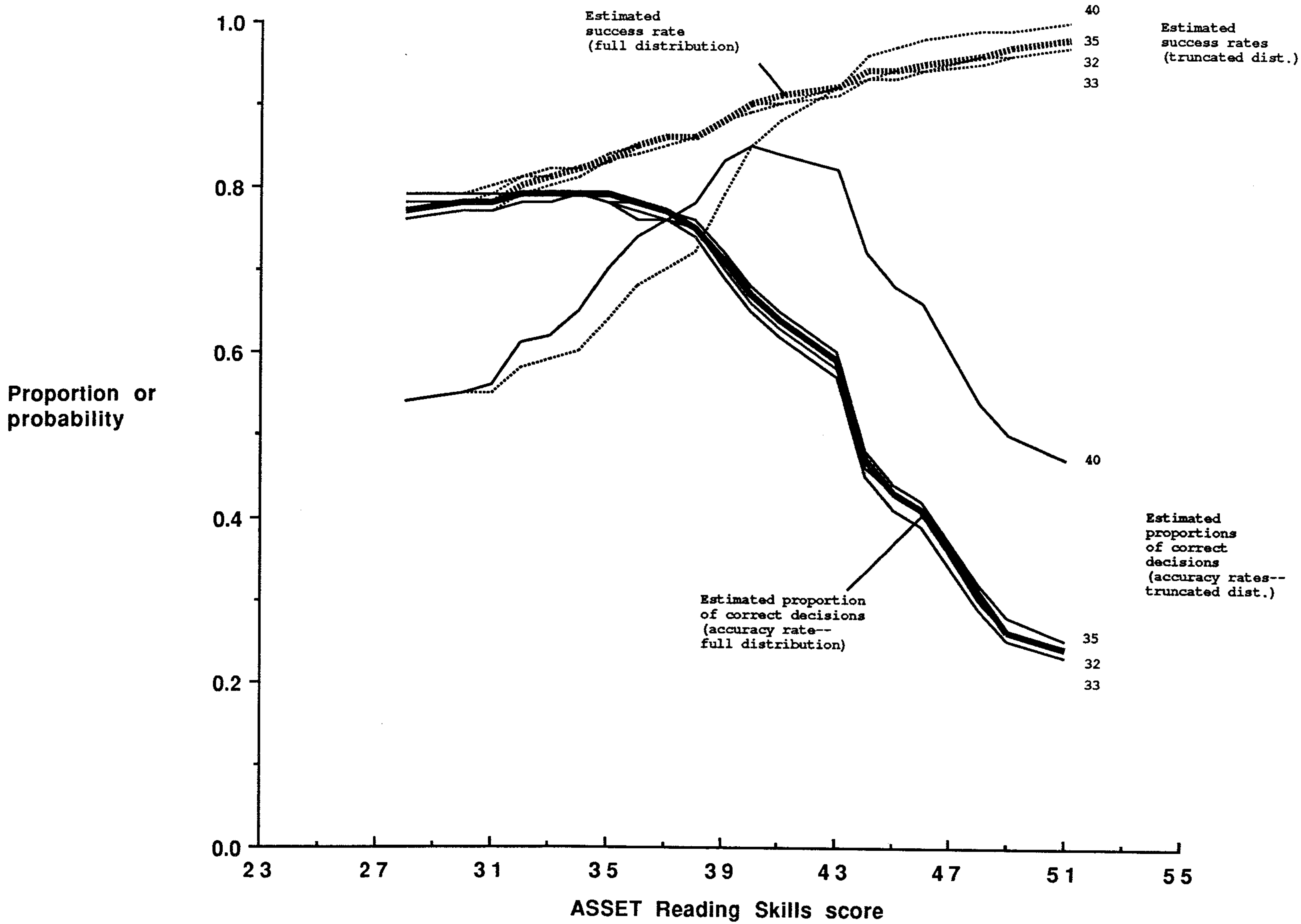

FIGURE 10. Effects of Truncation on Estimated Validity Indices for Psychology (Based on Data for One Institution)