DOCUMENT RESUME

ED 455 248 TM 033 049

AUTHOR	Bassiri, Dina; Schulz, E. Matthew
TITLE	Constructing a Universal Scale of High School Course Difficulty.
PUB DATE	2001-04-11
NOTE	22p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001). Contains small print.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Entrance Examinations; Grade Point Average; *Grades (Scholastic); Grading; *High School Students; High Schools; Item Response Theory; *Prediction; *Rating Scales; Scaling
IDENTIFIERS	*ACT Assessment; *Rasch Model

ABSTRACT

This study examined the usefulness of applying the Rasch rating scale model (D. Andrich, 1978) to high school grades. ACT Assessment test scores (English, Mathematics, Reading, and Science Reasoning) were used as "common items" to adjust for different grading standards in individual high school courses both within and across schools. This scaling approach yielded an ACT Assessment-adjusted high school grade point average (AA-HSGPA) that was comparable across schools, cohorts, and among students within the same school and cohort who took different courses. The AA-HSGPA was constructed for all ACT-tested students in 50 selected high schools. First-year college grades at a large midwestern public university were available for approximately 3,500 of these students. AA-HSGPA was a better predictor of first-year college grade point average (CGPA) than the regular high school grade point average (HSGPA). As expected, the regression of CGPA on HSGPA differed for high schools grouped by difficulty (easy or hard), but the regressions of CGPA on AA-HSGPA and the ACT Composite score did not. The best model for predicting CGPA included both the ACT Composite score and AA-HSGPA. (Contains 4 tables, 5 figures, and 22 references.) (Author/SLD)

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

ED 455 248

Constructing a Universal Scale of High School Course Difficulty

Dina Bassiri
E. Mathew Schulz
ACT, Inc.

Paper to be presented at the

annual meeting of the
American Educational Research Association
Seattle, Washington
April 11, 2001

TM033049

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 AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Bassiri

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Abstract

This study examined the usefulness of applying the Rasch rating scale model (Andrich, 1978) to high school grade data. ACT Assessment test scores (English, Mathematics, Reading, and Science Reasoning) were used as "common items" to adjust for different grading standards in individual high school courses both within and across schools. This scaling approach yielded an ACT Assessment-adjusted high school gpa (AA-HSGPA) that was comparable across schools, cohorts, and among students within the same school and cohort who take different courses. The AA-HSGPA was constructed for all ACT-tested students in 50 selected high schools. First-year college grades at a large midwestern public university were available for approximately 3,500 of these students. AA-HSGPA was a better predictor of first-year college grade point average (CGPA) than the regular high school grade point average (HSGPA). As expected, the regression of CGPA on HSGPA differed for high schools grouped by difficulty (easy or hard), but the regressions of CGPA on AA-HSGPA and the ACT Composite Score (ACTC) did not. The best model for predicting CGPA included both the ACT composite score and AA-HSGPA.

BEST COPY AVAILABLE

Constructing a Universal Scale of High School Course Difficulty

Theoretical Perspective

As a measure of academic achievement, the grade point average (GPA) is limited by the extraneous effects of schools and courses. The problem of school effects was first to be recognized. Linn (1966) noted that variation in grading practices among high schools made the high school gpa (HSGPA) a suboptimal predictor of achievement in college. The problem of course effects has received attention more recently. Students can earn a higher grade point average (gpa) simply by taking easier courses. These problems fuels grade inflation at the high school level (Ziomek and Svec, 1995), and creates incentives for students to avoid courses in difficult subjects such as mathematics and science (Johnson, 1997).

In an effort to improve the prediction of college GPA (CGPA) by HSGPA, Lindquist (1963) regressed CGPA on HSGPA separately for each combination of high school and college. This approach, called "central prediction systems"(Linn, 1966) was an attempt to control for school effects. The approach was disappointing, however. One or more of the variables needed for the regressions were missing for too many students (Lindquist, 1963). Sample size requirements were often not met for particular high school and college pairs. Linn (1966) noted that the accuracy of central prediction systems often was not so much better than that of simply using a standardized test, such as the ACT Assessment, in conjunction with HSGPA, to predict college achievement as to justify the cost and expense of these systems.

In an attempt to construct more reliable, predictable measures of college achievement, scaling models have been used to control for the effects of college courses (Young, 1990; Johnson, 1997). These models have been applied only to college grade data, possibly because course taking patterns differ more among college students than among high school students. Scale-adjusted CGPA measures, like CGPA, represent student achievement on a single (unidimensional) scale and are constructed entirely from course grade data.

Scaling procedures have shown promise with college grade data. Compared to CGPA, scale-adjusted CGPA is more predictable by pre-admissions variables such as HSGPA and standardized tests (Young, 1990; Caulkins, Larkey, & Wei, 1996; Johnson, 1997). Scale-adjusted CGPA is also better able to predict which of two students received the higher grade if both students took the same college course (Johnson, 1997).

The question to be addressed by the present study is whether a scaling approach might be effective in controlling for both school and course effects on HSGPA. With regard to controlling for course effects *within* a high school, the motivation and procedures are similar to those in studies with college grade data. For example, within any given high school, some college-bound students take core college-preparatory courses, while others don't. These differences reduce the ability of HSGPA to predict college achievement. A

scaling-adjusted HSGPA can be constructed from data within a high school because courses taken by the same students are "common items", and students who take two different courses are "common persons" in scaling terminology.

In order to control for school effects in a scaling procedure, however, one needs one or more common items between the schools. Linn (1966) suggested that an external variable, such as a standardized test, could be used in this fashion. In this study, we will use ACT Assessment test scores as "common items." The ACT Assessment tests (English, Mathematics, Reading, and Science Reasoning) are curriculum-based achievement tests (ACT, 1997). ACT test scores have strong relationships to high school course loads and to self-reported grades in high school (Noble & McNabb, 1989). Scores on these tests are comparable across high schools and over time (ACT, 1997). The ACT Assessment is taken by over one million high school juniors and seniors in thousands of high schools (ACT, 2000). With these characteristics, ACT Assessment tests are ideal "common items" for constructing a universal scale of high school course difficulty. The scale-adjusted HSGPA in our study is called "AA-HSGPA" because it is adjusted for school effects by the ACT-Assessment tests.

A potential advantage of a scaling approach over a central prediction system concerns the sample size and data requirements for estimating prediction equations. An AA-HSGPA can be constructed for any student in a high school in which a sufficient number of students take the ACT Assessment. With some scaling models, as few as thirty students may be needed to estimate the difficulty of a course on the universal scale. These data can be collected over time, including more than one year, as long as the grading practices for the courses or school remain approximately the same. The AA-HSGPA can be used to predict achievement at *any* college. Within a college, AA-HSGPA can be used to predict achievement of students from different high schools. There do not have to be sufficient numbers of students from a given high school in order to estimate school-specific regression equations.

Source of Data

Sample A) This sample included data from all of the high schools represented in the four freshman cohorts from a public university (referred to as Midwest University). There were 13,740 students from over 1, 300 high schools. These students were in the high school graduating class of 1996, 1997, 1998, or 1999, had self-reported high school grades and ACT Assessment scores, and were attending a high school that had one or more graduates who attended the Midwest University between 1997 and 2000 inclusively.

ACT Assessment scores include four subject-area scores—English, Mathematics, Reading, and Science Reasoning—on a 1 to 36 point scale, plus a Composite score. The Composite score is the rounded average of the four subject area scores and has an average of 21 among ACT-tested students nationally.

ERIC Full Text Provided by ERIC

Self-reported high school grades came from the Course Grade Information Section (CGIS) of the Student Profile component of the ACT Assessment. This component is completed voluntarily by students when they register for the ACT Assessment. Students are asked to report grades in standard college preparatory courses such as English I to English IV, Algebra, Geometry, etc. Course grades are reported on A to F scale. Previous research shows that self-reported grades are highly comparable to actual grades received (Sawyer, Laing & Houston,1988). For all analyses high school course grades were numerically coded as A=4, B=3, C=2, D=1, and F=0.

Sample B) Fifty schools were selected that were representative of the distribution of 'difficulty' of 1300 plus high schools and had ten or more graduates who attended the Midwest University between1997and 2000 inclusively. The selection of these fifty schools was done in the following manner. The mean ACT composite score (ACTC) and mean high school GPA (HSGPA), based on self-reported grades, was computed for each of the 1300-plus schools (for 13,740 students). Mean HSGPA was regressed on mean ACT composite score, with the number of students within school as the weight for each observation. Fifty schools were selected so the distribution of their weighted residuals matched that of the entire sample.

This sample included all ACT-tested students who were from the fifty selected high schools regardless of which college they attended or even whether they attended college. There were 31,791 students in this sample. The records were obtained from the ACT high school profile data set that ACT Inc. keeps of all ACT-tested students by high school and year of high school graduation.

Sample C) This sample included all Sample B students who earned a grade at the Midwest University in the two semesters of the year following their high school graduation i.e. had a college GPA. There were 3,329 of these students. The number of students per school (50 schools) ranged from 5 to 294.

Computation of College GPA (CGPA): A two-semester college GPA was computed for all ACT-tested students who earned one or more grades in each semester of the year following their high school graduation. For example, the college GPA of 1996 high school graduates was based on grades earned during fall semester of 1996 and spring semester of 1997. ACT-tested status and graduation year were indicated by whether a student was present in the ACT high school profile data set for the given graduation year.

The number of students for whom a college GPA was computed, by cohort, was 785 for 1996, 848 for 1997, 921 for 1998, and 775 for 1999. The number of grades per student ranged from 2 to 19. Over seventy five percent of these students had 8, 9, or 10 grades. The college GPA was the unweighted average of the grades. Grades included pluses (+) and minuses (–). These were numerically coded as A=4, B=3, C=2, D=1, F=0 with 0.33 added or subtracted for a plus or minus. (There were no F- grades.) The mean GPA, by cohort, was 2.74 for 1996, 2.78 for1997, 2.74 for1998, and 2.80 for 1999

Samples D (Students from Easy Schools) and E (Students from Hard Schools): Three or more schools were selected from each end of the ordered weighted residuals to represent easy and hard schools with regard to grading policy of the fifty high schools. These residuals resulted from regressing mean HSGPA on mean ACTC score as was explained previously. Note that negative residuals correspond to schools with strict grading policy (hard schools), whereas positive residuals correspond to schools with lenient grading policy (easy schools). Easy and hard schools each contained a total of 179 and 178 students, respectively who attended Midwest University between 1996 to 1999.

Methods

1) Scaling Analyses

All scaling analyses were performed with the Bigsteps computer program, Version 2.27 (Wright & Linacre, 1990). A combination of the Rating Scale Model (RSM) (Andrich, 1978) and the Partial Credit Model (PCM) (Masters, 1982) was in most analyses. For J+1 distinct grades numerically coded, say, F=0, D=1, ..., A=4, (J+1 = 5) the RSM is:

$$\ln\left(\frac{P_{nij}}{P_{nij-1}}\right) = \beta_n - \delta_i - \tau_j, \qquad j=1,2,\ldots,J;\ i=1,\ldots,I;\ n=1,\ldots,N \tag{1}$$

where

- ln means to take the natural log,
- P_{nij} is the probability that student n earns grade j in course i,
- P_{nij-1} is the probability that student n earns grade j-1 in course i,
- β_n is the AA-HSGPA of student n,
- δ_i is the difficulty of course i,
- τ_j is the "step threshold"--the *relative* difficulty of earning grade j or higher rather than grade j-1; $\tau_0 \equiv 0$.

In the RSM, the items share a common set of step thresholds (τ_j). In the PCM, each item has its own set of step thresholds. This is accomplished notationally with the replacement of "τ_{ij}" for "τ_i" in Equation 1.

The RSM was used for all high school courses. One reason for this is that there were insufficient numbers of students per grade within each course within a school. For example, in some courses, no "F" or "D" was given. This situation presents problems in estimating course difficulty.if courses have their own set of step threshold parameters. The RSM was the best performing IRT model for scaling college grade data (Lei, Bassiri, & Schulz, 2001). Other IRT models used in that study included the PCM, the graded response model (Samejima, 1969), and the generalized partial credit model (Muraki, 1982).

The PCM was used for the ACT tests. For purposes of scaling the ACT tests as common items with high school courses, the ACT scale scores were recoded as follows: [0,16]=0, [17,19]=1, [20,22]=2, [23,25]=3, [26,36]=4. This coding created as many scores as there were grades in high school courses (five), and yielded approximately equal numbers of students per score across the four tests. The PCM was used for the ACT tests because there were sufficient numbers of students per new score for each ACT test and because the distribution of the new scores differed from across tests in ways that suggested they should have separate step thresholds.

In order to construct AA-HSGPA measures, we perform three basic series of Bigsteps analyses. Each series of analyses used Sample B--all ACT-tested students from 50 selected high schools. All analyses used the first 23 (high school) courses in the CGIS (Table 1).

All-data Scaling Analysis (RSA1): The purpose of this analysis was to bring the ACT tests onto a scale that represents high school course grades. This was done by jointly calibrating (estimating the parameters of) ACT tests and high school courses. We did not consider it necessary at this point for the high school course parameters to be school-specific, so only twenty-seven items were defined for the analysis: 23 high school courses and 4 ACT tests.

School-specific Scaling Analysis (RSA2): School-specific analyses were performed in order to obtain school-specific parameters for the 23 courses on a common scale--the universal scale of high school course difficulty. This was done by anchoring (fixing) the ACT test parameters at the estimates obtained for them in RSA1 and including the tests as items (for a total of 27 items) in each school-specific analysis. The ACT tests were thus treated as common items across schools.

Estimating AA-HSGPA (RSA3): These analyses were performed so that AA-HSGPA measures would be based exclusively on high school grades (calibrated to a universal scale of difficulty). Analyses were performed separately by school. Each analysis involved just twenty three items--the high school courses. ACT tests were not included. Course parameters were anchored at the estimates obtained in RSA2.

2) Regression Analyses

Regression analyses were performed separately on data from samples C, D (easy) and E (hard) and on D and E combined as follows:

Sample C: Five regression analyses (using proc reg in SAS) were performed on Sample C data.

1) CGPA = ACTC
2) CGPA = HSGPA
3) CGPA = ACTC, HSGPA
4) CGPA = AA-HSGPA
5) CGPA = ACTC, AA-HSGPA

Samples D and E Separately: Three regression analyses (using proc reg in SAS) were performed on each sample.

1) CGPA = ACTC
2) CGPA = HSGPA
3) CGPA = AA-HSGPA

Samples D and E combined: Nine regression analyses (using proc glm in SAS) were performed on the combined samples where high school type (HS-type), easy or hard, was dummy coded as [hard =1] and [easy=0]. For the ACTC the regression formulas were

1) Common Intercept & Slope (Model I)
CGPA = ACTC
2) Separate Intercept & Common Slope (Model II)
CGPA = ACTC, HS-type
3) Separate Intercept & Slope (Model III)
CGPA = ACTC, HS-type, ACTC * HS-type

The same models as above were applied to the other two predictors (HSGPA and AA-HSGPA) replacing ACTC in the equations. Note also that each predictor was centered around its respective mean.

Results

Summary Statistics: Table 2 shows descriptive statistics for ACTC, HSGPA, AA-HSGPA and CGPA for samples C, D, and E. Notice that easy and hard schools are very similar in terms of HSGPA (3.57 vs 3.59), but different with respect to ACTC, AA-HSGPA and CGPA (2.69 vs 2.96). Also note that AA-HSGPA is not on the same scale as HSGPA i.e., not on 0 to 4 scale.

All-data Rating Scale Analysis (RSA1): The average difficulty of the 23 high school courses over all fifty high schools is represented in Table 1 and plotted in Figure 1. Numbers on overall curve correspond to the high school courses sequence number as shown in Table 1. Notice that speech (5), psychology (23) and Beginning Calculus (10) are ranked as easiest, moderate and hardest overall, respectively.

School-specific Rating Scale Analysis (RSA2): Also, in Figure 1, the difficulty of the high school courses within two schools (denoted School A and School B) are plotted against overall average course difficulty. These schools were selected because they had the most extreme average course difficulty as measured by the courses' school-specific parameter estimates in RSA2. School A's average course difficulty was 0.79 and school B's was -2.30. Overall average course difficulty was -0.43. Unexpectedly, psychology with an overall course difficulty of -0.43 had very similar course difficulty in School A (-1.41) and School B (-1.45).

Step Difficulties for 46 high schools are depicted in Figure 2. Recall that the 23 high school courses were constrained to have a common set of step parameters within a school, but the step parameters were allowed to vary across schools. Four high schools were excluded from this plot since no F grades were assigned to students in any course within these schools and, as a result, step 1 difficulty equaled zero. As shown in Figure 2, step difficulties get progressively larger as we move from step 1 to step 4. This indicates that across schools it is harder to replace grade B with A than it is to replace grade F with D or grade D with C or grade C with B. Also, step 1 and 2 difficulty have the most and the least dispersion, respectively, across the 46 high schools.

The across-school difficulty distribution of three selected courses is shown in Figure 3. The courses are the easiest (5), moderate (23) and hardest (10) overall according to Table 1. Course difficulty of the moderate overall (psychology) is more variable across 46 schools than that of the hardest overall (beginning calculus) followed by the easiest overall (speech) course.

Sample C Regression Analysis: The proportion of variance (R^2) explained by each of the five regression models was significant at the .01 level (see Table 3 for R^2 and p values). However, AA-HSGPA was a better predictor compared to unadjusted one (HSGPA) with or without ACTC in the model (0.294 vs 0.264 and 0.286 vs 0.236, respectively).

Samples D and E Combined Regression Analysis: Table 4 reports R^2 and F-ratio for the difference between the three models I, II, and III regarding intercept and slope of regression lines. Once again the proportion of variance accounted for by each of the three predictors was significant at the .01 level. There was an increase in predictability measured by R^2 when the predictor variable was AA-HSGPA rather than raw high school gpa (HSGPA). This is apparent from Table 4 under R^2, when we compare third and second row.

As expected there was a significant difference between hard and easy schools regarding unadjusted high school gpa (HSGPA): This inference is based on the significant F-ratio for difference between model II and I (14.07) and between model III and I (7.29). These results indicate that there is a significant main effect regarding high school type (easy or hard). However, in the absence of significant F-ratio for difference between model III and II we can conclude that interaction effect is insignificant. Thus separate intercept and

common slope is the appropriate model to be considered for predicting CGPA from HSGPA from high schools of different grading difficulty.

Regarding the other two predictors ACTC and AA-HSGPA, there were no significant differences between easy and hard schools. Neither main effects nor interaction effects were significant. This is apparent from Table 4 as none of the F-ratios for difference were significant for the two predictors.

Samples D and E Separate Regression Analysis: Schematically the above inferences may be drawn from Figures 4, 5, and 6. As there is only significant high school type effect when predicting CGPA from HSGPA (see Figure 5) and no significant difference between easy and hard schools when ACTC or AA-HSGPA are considered (see Figures 4 and 6, respectively). Expectedly this predicts that a student from a hard school will have a higher CGPA than a student from an easy school with an identical HSGPA. Note that the three points on each regression line corresponds to the mean, 5% and 95% quantile points of the respective predictor variable.

Conclusion

The results of this study reveal the problems with using the high school grade point average alone to predict achievement in college. As shown in Figure 5, HSGPA has a different relationship to college achievement, depending on the grading standards of the high school. This is the basic phenomenon that led early researchers to consider using high school and college-specific regressions in "central prediction systems" to improve information for college admissions decisions (Linn, 1966; Lindquist, 1963).

Our results also show and confirm the notion that a standardized test such as the ACT Assessment is a common yardstick for comparing academic achievement of students across schools. That the ACT Assessment is a common yardstick as shown by Figure 4 and the nonsignificance of separate intercept and slopes for easy versus hard schools in Table 4.

The unique contribution of this study, however, is that it presents an additional way to use the ACT Assessment as a common yardstick. The traditional use of the ACT Assessment is to consider a student's test scores, along with other factors, such as HSGPA or AA-HSGPA, to predict their achievement in college and to make college admissions decisions. This use concerns differences among students. Another way is to control for school effects when the AA-HSGPA is constructed. This use concerns differences among schools. There is no overlap or redundancy between these two uses.

The continued, traditional use of ACT Assessment scores is consistent with the results of this study. Although grade-based, high school achievement measures, such as HSGPA and class rank (which is based on HSGPA) typically slightly outperform standardized tests such as the SAT or ACT in predicting college GPA (Willingham, Lewis, Morgan, & Ramist, 1990; Noble 1991; present study), such results have *not* been interpreted to suggest that college admissions decisions should be based on high school grades alone. One reason standardized tests continue to be used is that they *improve* the prediction of college GPA. In the present study, for example, the two highest R^2 values for predicting CGPA were $R^2 \doteq .286$ when ACTC was combined with HSGPA, and $R^2 = .294$ when ACTC was combined with AA-HSGPA.

Another reason ACT Assessment tests should continue to be used traditionally is related to the nature of course grades. Course grades are based on many factors such as attendance, social conformity, willingness to please authority figures, and class participation. These factors do not represent cognitive academic achievement. Indeed, one possible explanation for the fact that grade-based measures from high school and college correlate slightly better with each other than with a standardized test is that they have more noncognitive factors in common. Their wide-spread use as measures of academic achievement is sometimes attributed to the convenience of obtaining them rather than to their validity as measures of academic achievement (Johnson, 1997).

Finally, it seems to be common wisdom that no single variable can be relied upon for making admissions decisions or for representing academic achievement. Besides being comparable across schools, standardized tests are used in combination with grade data because they represent unique variables. As objective measures of cognitive academic achievement, ACT Assessment scores are less influenced by the same, possibly extraneous, factors that influence course grades. Therefore, even if high school grades could somehow be made comparable across schools, as with the AA-HSGPA, they would not represent the same variable as an ACT Assessment score. The consequence of this point is illustrated in this study by the fact that the ACTC improved the prediction of CGPA over that of AA-HSGPA alone.

The present study showed that scaling approaches to high school grade data could be useful. However, no adjustment was made on college grade data. Linn (1966) suggested using regression regression-adjusted gpa at both the high school and college levels in "central prediction systems." A related question for future research, then, is how much improvement in predicting achievements in college from high school grade data can be gained when adjustment is made at both levels. We predict that the best model will consist of AA-HSGPA and ACT Assessment scores predicting a scale-adjusted CPGA.

References

ACT, Inc., (1997). ACT Assessment Technical Manual. Iowa City, IA: Author.

ACT, Inc., (2000). The High School Profile Report. Iowa City, IA: Author.

Andrich, D. (1978). A rating formulation for ordered response categories. Psychometrika, 43, 561-573.

Caulkins, J. P., Larkey, P. D., and Wei, J. (1996) Adjusting GPA to Reflect Course Difficulty. The Heinz School of Public Policy and Management, Carnegie Mellon University.

Johnson, V. E. (1997). An alternative to traditional GPA for evaluating student performance. Statistical Science, 12, 251-278.

Lei, P., Bassiri, D., & Schulz, E. M. (April, 2001). Alternatives to the grade point average as measures of academic achievement in college. Paper presented at the American Educational Research Association annual meeting. Seattle, WA.

Linn, R. L. (1966). Grade adjustments for prediction of academic performance: A review. Journal of Educational Measurement, 3, 313-329.

Lindquist, E. F. (1963). An evaluation of a technique for scaling high school grades to improve prediction of college success. Educational and Psychological Measurement, 23, 623-646.

Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

Maxey, J. (2001). Selected Trends on ACT-Tested Students. Iowa City, IA: ACT, Inc.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement, 16(2), 159-176.

Noble, J., & McNabb, T. (1989). Differential coursework and grades in high school: Implications for performance on the ACT Assessment. ACT Research Report 89-5. Iowa City, IA: ACT, Inc.

Noble, J. P. (1991). Predicting college grades from ACT Assessment scores and high school course work and grade information. ACT Research Report 91-3. Iowa City, IA: ACT, Inc.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17.

Sawyer, R., Laing, J., & Houston, M. (1988). Accuracy of self-reported high school courses and grades of college-bound students (ACT Research Rep. No. 88-1). Iowa City, Iowa: ACT, Inc.

Stricker, L. J., Rock, D. A., Burton, N. W., Muraki, E., & Jirele, T. J. (1994). Adjusting college grade point average criteria for variations in grading standards: A comparisons of methods. *Journal of Applied Psychology*, *79*(2), 178-183.

Willingham, W. W., Lewis, C., Morgan, R., & Ramist, L. (1990). Predicting college grades: an analysis of institutional trends over two decades. Princeton, NJ: Educational Testing Service.

Wright, B. D., & Linacre, J. M. (1990). A User's Guide to BIGSTEPS: Rasch-Model Computer Program Version 2.0. Chicago, IL: MESA Press.

Wright, B.D., & Stone, M. H. (1979). Best Test Design. Chicago, IL: MESA Press.

Young, J. W. (1990). Adjusting the cumulative GPA using item response theory. Journal of Educational Measurement, 27, 175-186.

Young, J. W. (1993). Grade adjustment methods. *Review of Educational Research*, *63*(2), 151-165.

Ziomek, R. L., and Svec, J. C. (1995). High school grades and achievement: Evidence of grade inflation. (ACT Research Rep. No. 95-3). Iowa City, Iowa.

Table 1

List of First Twenty-Three Courses in the CGIS and Percent Course Taken and Course Difficulty from 50 Schools

Course Number	Course Title	% Taken	Rating Scale Model Difficulty *
1	English 9	94	-0.54
2	English 10	94	-0.53
3	English 11	85	-0.48
4	English 12	14	-0.83
5	Speech	51	-0.93
6	Algebra I	91	-0.16
7	Algebra II	78	0.32
8	Geometry	89	0.17
9	Trigonometry	29	0.46
10	Beginning Calculus	3	0.75
11	Other Advanced Math	18	0.39
12	Computer Science	8	-0.90
13	General Science	72	-0.55
14	Biology	91	-0.15
15	Chemistry	73	0.14
16	Physics	21	0.24
17	U.S. History	90	-0.54
18	World History	69	-0.44
19	Other History	23	-0.43
20	American Government	29	-0.55
21	Economics	18	-0.55
22	Geography	27	-0.81
23	Psychology	16	-0.47

* Based on all-data Rating Scale analysis.

Table 2

Descriptive Statistics for Models Predicting CGPA using ACTC, HSGPA and AA-HSGPA as Predictors by Different Samples of High Schools

Variable	Sample			
	D Students from Easy Schools	E Students from Hard Schools	D & E Combined	C
ACTC	22.29 (3.49)	25.84 (3.63)	24.06 (3.56)	24.61 (3.79)
HSGPA	3.57 (0.38)	3.59 (0.37)	3.58 (0.37)	3.53 (0.39)
AA-HSGPA	2.68 (1.57)	3.46 (1.49)	3.07 (1.53)	3.12 (1.46)
CGPA	2.69 (0.72)	2.96 (0.70)	2.82 (0.71)	2.76 (0.71)

Note. SDs are given in parentheses. Also, SDs for combined schools are based on square root of pooled within variances of easy and hard schools.

Table 3

R^2 Values for Models Predicting College GPA (CGPA) N=3,329 from 50 Schools

Predictor	R^2	P
ACTC	0.125	0.0001
HSGPA	0.236	0.0001
ACTC, HSGPA	0.264	0.0001
AA-HSGPA	0.286	0.0001
ACTC, AA-HSGPA	0.294	0.0001

Note: The R^2 values in this table are somewhat lower than, or not strictly comparable to, those that are typically obtained or might be obtained. Typically, the R^2 between CGPA and a standardized test battery is based on multiple predictors (e.g., all four ACT tests), and the CGPA is the credit-weighted average of course grades. Likewise, in the prediction research services at ACT, Inc., high school grades are represented as multiple predictors consisting of subject area gpa's in English, Mathematics, Social Studies, and Natural Sciences. In one recent ACT report (Maxey, 2001) the median correlations between the freshman CGPA and predictive indices based on either 1) ACT tests, 2) self-reported high school grades, and 3) their combination were, respectively, .423, .484, and .530. These translate to R^2 values of .18, .23, and .28, respectively. These values were were based on 291 colleges and the 1997-98 college freshman cohort.

Table 4

Easy vs Hard Schools

Predictor	R^2			F-Ratio for Difference		
	Common Intercept & Slope (I)	Separate Intercept & Common Slope (II)	Separate Intercept & Slope (III)	II vs I	III vs I	III vs II
ACTC	0.21*	0.21*	0.21*	0.20	0.13	0.05
HSGPA	0.22*	0.25*	0.25*	14.07*	7.29*	0.55
AA-HSGPA	0.34*	0.34*	0.35*	0.98	1.24	1.50

* Significant at .01 level

Figure 1

Course Difficulty Overall and Within two Schools with Extreme Average Course Difficulty

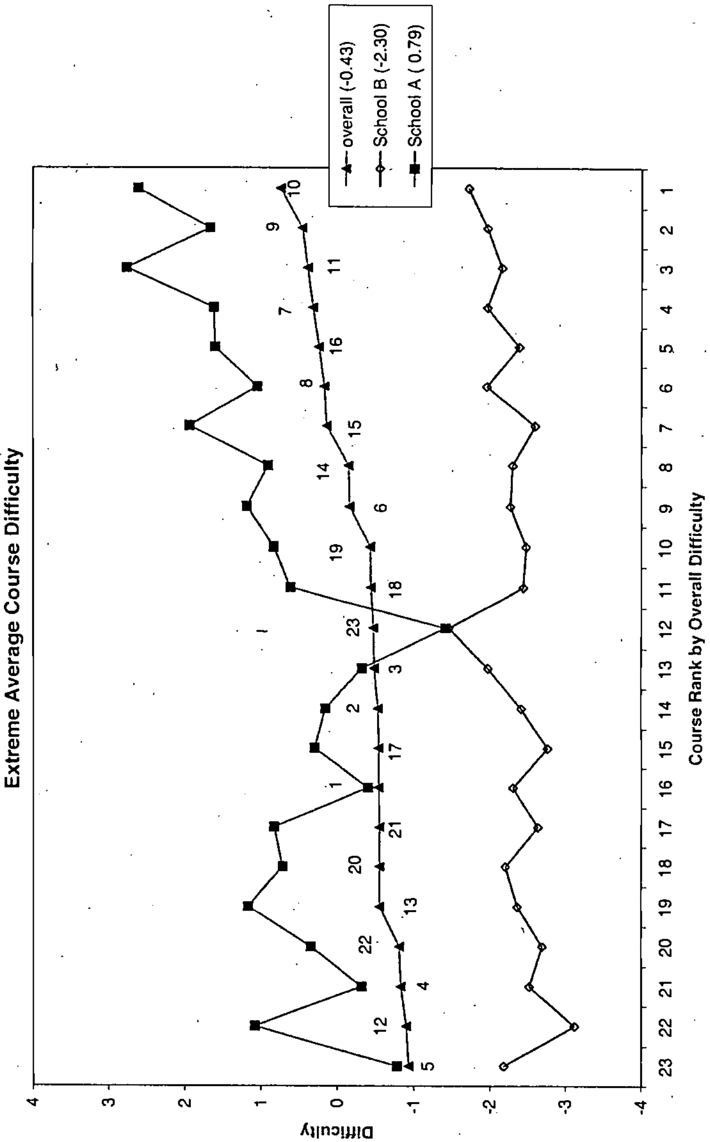


Figure 2

Plot of Step Difficulties (τ_j) for 46 Schools

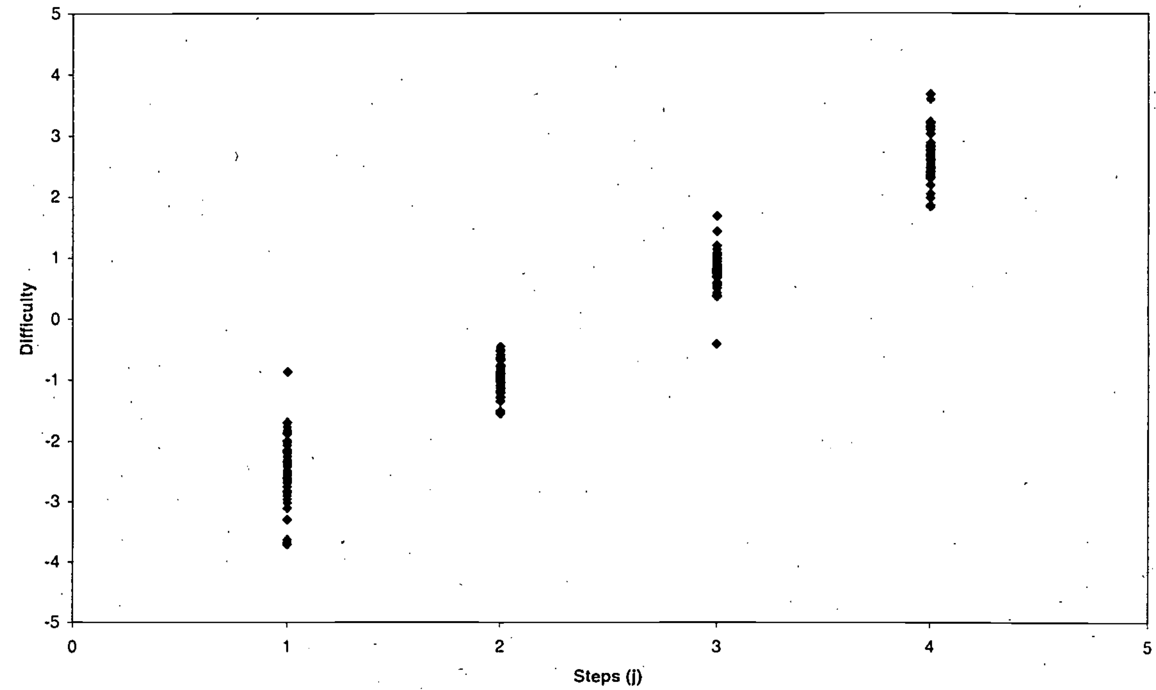


Figure 3

Plot of Course Difficulties for 46 Schools

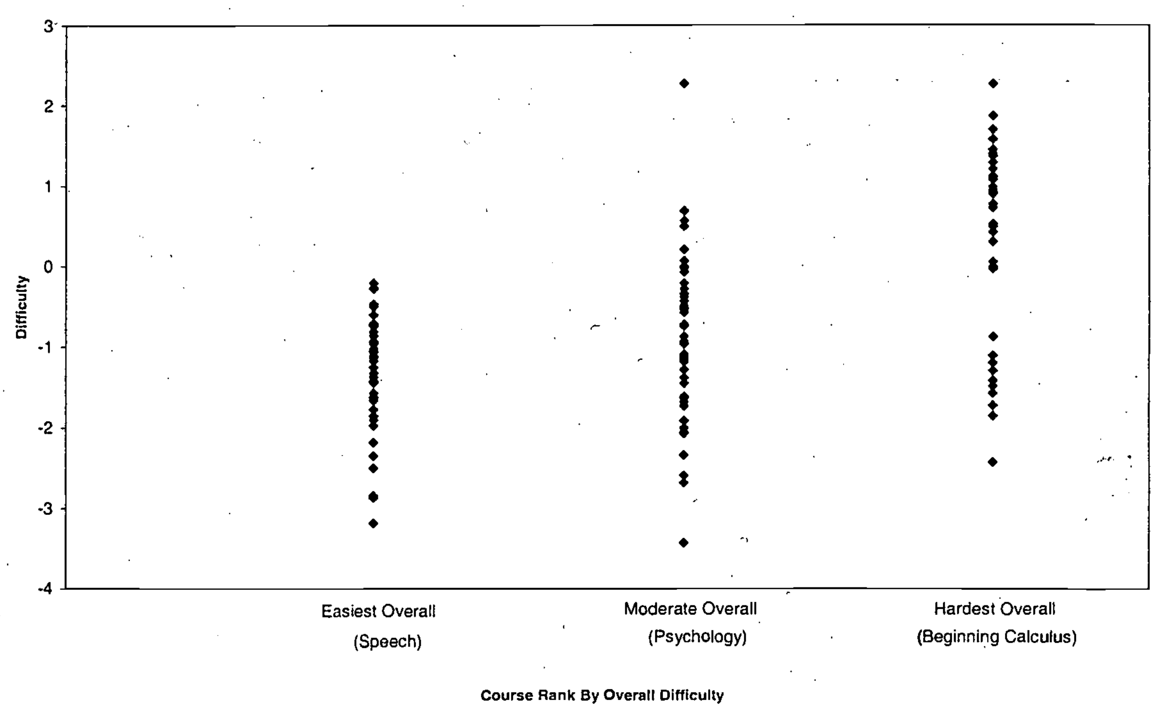
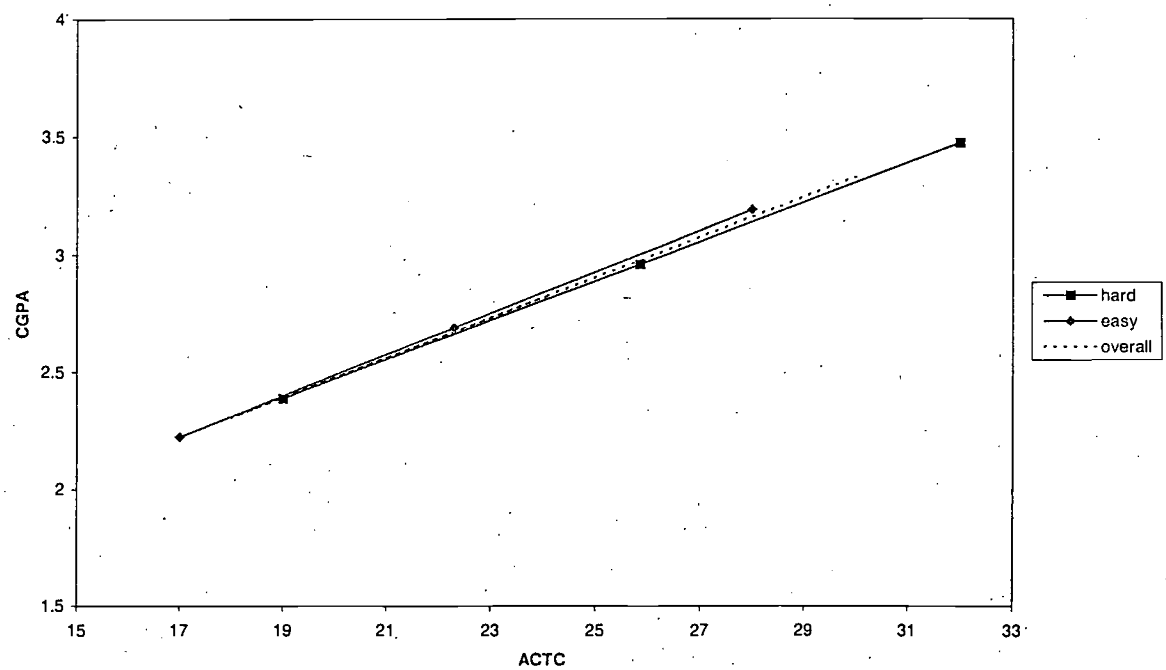


Figure 4

Plot of Regression Line of CGPA on ACTC for Hard and Easy Schools



21

Figure 5

Plot of Regression of CGPA on HSGPA for Hard and Easy Schools

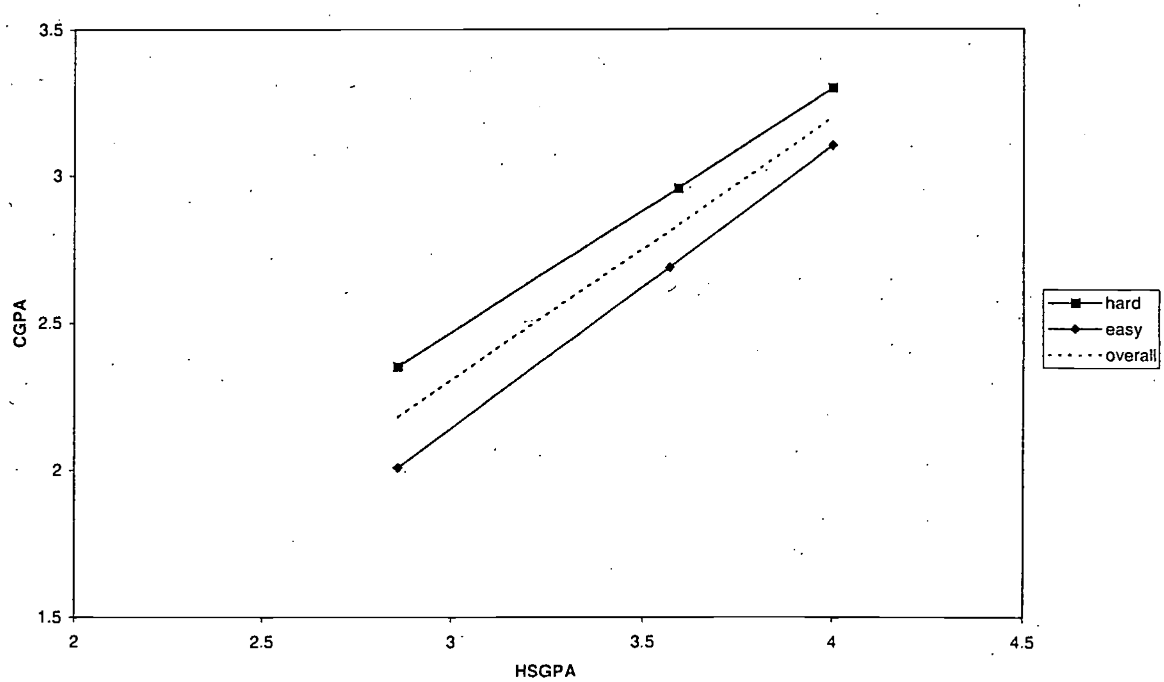
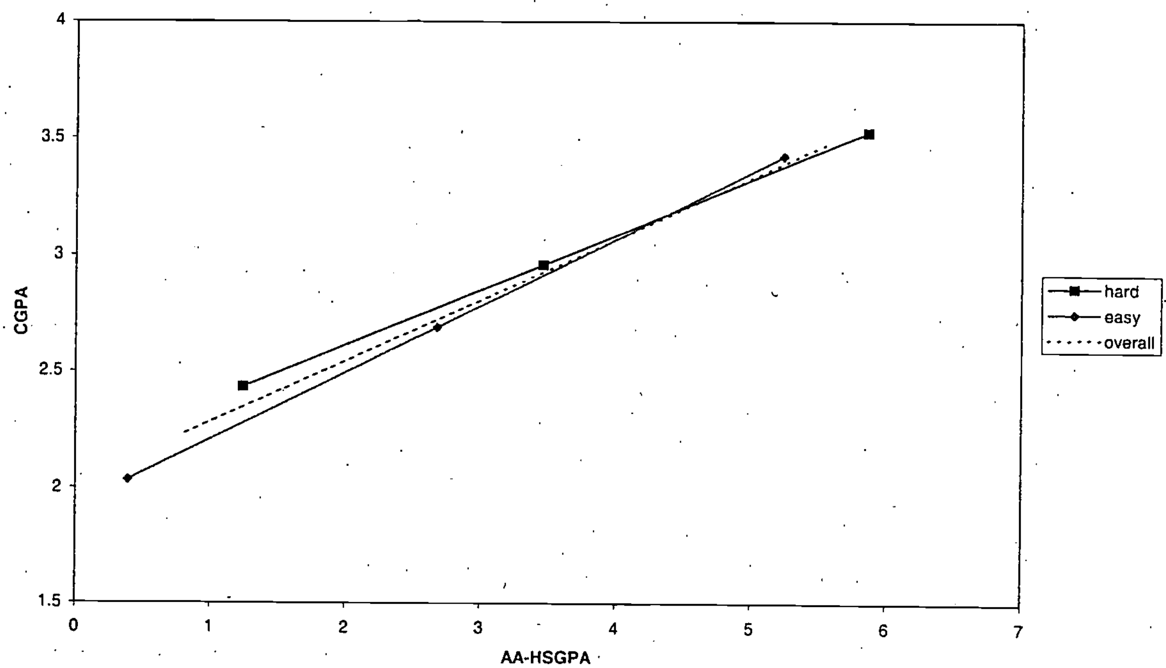


Figure 6

Plot of Regression Line of CGPA on AA-HSGPA for Hard and easy Schools



23

ERIC

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM033049

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Constructing a Universal Scale of High School Course Difficulty

Author(s): Bassiri, Dina; Schulz, E. Matthew

Corporate Source: ACT, Inc.

Publication Date: 4/11/2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Dina Bassiri	Printed Name/Position/Title: Research Associate
Organization/Address: ACT, Inc., PO Box 168, 2201 N Dodge, Iowa City, IA 52243	Telephone: 319-337-1235 FAX: 319-339-3020
	E-Mail Address: bassiri@act.org Date: 5/2/01

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)

ERIC Full Text Provided by ERIC