DOCUMENT RESUME

ED 473 070 HE 035 661

AUTHOR	Noble, Julie; Sawyer, Richard
TITLE	Accuracy of High School Grades and College Admissions Test Scores for Predicting Different Levels of Academic Achievement in College. AIR 2002 Forum Paper.
PUB DATE	2002-06-03
NOTE	24p.; Paper presented at the Annual Meeting of the Association for Institutional Research (42nd, Toronto, Ontario, Canada, June 2-5, 2002).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; *College Entrance Examinations; *College Students; *Grade Point Average; *Grade Prediction; Higher Education; Scores
IDENTIFIERS	*ACT Assessment

ABSTRACT

This study compared the cross-validated accuracy of American College Test (ACT) Composite score and high school grade point average (GPA) (HSAV) for predicting different levels of first-year college GPA. Logistic regression models were developed by institution, for predicting GPA levels of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or higher. These models were then applied to data from the next year (cross-validation over time) for the same institutions. Data were taken from 2 years of ACT Prediction Research history files. The 1996-1997 files contained HSAX, AXT Composite scores, and college grades for 219,435 first-year students from 301 postsecondary institutions, and the 1996-1997 data consisted of the same information for 782,957 students. The resulting base-year and cross-validation year statistics were summarized across institutions. HSAV predictions of first-year GPA levels of 2.00, 2.50, and 3.00 were somewhat more accurate than those based on ACT Composite score. HSAV was an ineffective predictor for higher GPA levels, however. ACT Composite score predictions, in contrast, were effective at all first-year GPA levels. (Contains 4 tables and 21 references.) (Author/SLD)

Accuracy of High School Grades and College Admissions Test Scores for Predicting Different Levels of Academic Achievement in College

Julie Noble and Richard Sawyer

ACT, Inc.

Association for Institutional Research Annual Forum

Toronto, Canada

June 3, 2002

Accuracy of High School Grades and College Admissions Test Scores for Predicting Different Levels of Academic Achievement in College

Abstract

This study compared the crossvalidated accuracy of ACT Composite score and high school GPA (HSAV) for predicting different levels of first-year college GPA. Logistic regression models were developed, by institution, for predicting GPA levels of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or higher. These models were then applied to data from the next year (crossvalidation over time) for the same institutions. The resulting base-year and cross-validation year statistics were summarized across institutions.

HSAV predictions of first-year GPA levels of 2.00, 2.50, and 3.00 were somewhat more accurate than those based on ACT Composite score. HSAV was an ineffective predictor for higher GPA levels, however. ACT Composite score predictions, in contrast, were effective at all first-year GPA levels.

Accuracy of High School Grades and College Admissions Test Scores for Predicting Different Levels of Academic Achievement in College

College admissions officials typically use both high school GPA and scores on college entrance tests (such as the ACT Assessment) to predict (directly or indirectly) an applicant's probability of academic success in the first year of college (Breland, H., Maxey, J., Gernand, R., Cumming, T. and Trapani, C., in press). Academic success is typically measured by first-year college GPA.

Although high school GPA and college GPA both measure educational achievement, they also include other personal characteristics such as effort, attendance, conformity, and motivation (Goldman & Widawski, 1976; Stiggins, et al., 1989). In contrast, ACT scores primarily measure educational achievement in college-preparatory courses (ACT, 1997c). One might therefore expect high school GPA and ACT scores to be related to first-year GPA in different ways: High school GPA likely relates to both the cognitive and the noncognitive components of college GPA. ACT scores, on the other hand, likely relate only to the cognitive components of college GPA.

Goldman's and others' research in the 1970s found that the presence of non-achievement components in college grades was related to average student ability (e.g., Goldman & Hewitt, 1975; Goldman, Schmidt, Hewitt, & Fisher, 1974; Goldman & Widawski, 1976). They found that *high* college grades were more likely to reflect cognitive achievement and less likely to reflect noncognitive factors. One might therefore expect that

a. predictions of moderate levels of college GPA (e.g., 2.00 or higher) based on high school GPA would be more accurate than predictions based on ACT scores, but

b. predictions of high levels of college GPA (e.g., 3.0 or higher) based on ACT scores would be more accurate than predictions based on high school GPA.

The research comparing predictions of first-year GPA from ACT scores or high school average has largely been correlational in nature (e.g., ACT, 1998d). Recent summary statistics for institutions participating in ACT's Research Services showed that, across 129 colleges, the median multiple correlation (across institutions) between first-year GPA and the four ACT scores (in English, mathematics, reading, and science reasoning) was .43 (Maxey, in press). The corresponding median correlation between first-year GPA and four high school average grades was .48. The eight-predictor ACT score/high school grade average multiple correlation with first-year GPA was .53. These results were not corrected for range restriction caused by selection on ACT scores and/or high school grades.

Other research has examined prediction accuracy at first-year GPA levels of 2.00 or higher or at 3.00 or higher (ACT, 1997c). This study showed that high school GPA was a slightly better predictor of first-year college GPAs of 2.00 or higher than were ACT scores. For predicting whether students earn a 3.00 or higher college GPA, however, ACT Composite score and high school average had the same accuracy: The typical percentage of accurate predictions was 79% using either predictor, and the typical percentage of correct classifications using a joint ACT Composite/high school average model was 80%.

The purpose of this study is to compare the accuracy of ACT Composite score and high school GPA for predicting successive levels of first-year college GPA. The levels considered were first-year GPAs of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or higher.

In practice, prediction equations based on data from one freshman class are applied to the test scores and/or high school averages of *future* applicants. Because students may differ over time in their test scores, high school GPA, or college grades, predictive validity statistics developed from one year's data may misstate the strength of the relationship associated with actual use of such predictions. Crossvalidation allows us to study the accuracy of using prediction equations developed from one freshman class to forecast the success of a subsequent class. This procedure models the actual use of prediction equations by institutions, and it avoids the tendency of estimates of prediction accuracy based on a single year's data to be overly optimistic. A second purpose of this study, therefore, was to determine the crossvalidated accuracy of prediction equations and cutoffs based on ACT Composite score, high school average, and both variables used jointly.

Data for the Study

The data for the study were taken from two years of ACT's Prediction Research history files (ACT, 1997b; 1998c). The 1996-97 file consisted of the high school GPAs (HSAV), ACT Composite scores, and college grades of 219,435 first-year students from 301 postsecondary institutions (minimum sample size = 50). High school GPAs were based on students' self-reports of their grades in 30 college-preparatory high school courses provided at the time they took the ACT Assessment.

In addition, 728,957 nonenrolled students were identified from the 1996-97 ACT Class Profile history (ACT, 1997a), a database consisting of enrollment information and ACT Assessment records of both enrolled and nonenrolled students from over 900 institutions. Nonenrolled students had requested that their ACT scores be sent to at least one of the 301 institutions, but they did not enroll at that institution. These students, plus those who actually enrolled in an institution and completed their first year, comprised the base-year "applicant pool" for that institution. In fact, some of the nonenrolled students did not apply for admission to the institutions to which they sent their scores, but it was not possible to distinguish the nonapplicants from actual applicants. The analyses in this paper are based on data from all score senders; they are considered to be "applicants." All students had valid ACT Composite scores and HSAVs; enrolled students also had valid first-year GPAs.

The 1997-98 Prediction Research file consisted of the same elements described above for 214,924 first-year students from 294 colleges (minimum sample size = 50). In addition, 749,002 nonenrolled students were identified from the 1997-98 ACT Class Profile history (ACT, 1998a). These files comprised the crossvalidation-year data set.

Operationally, first-year GPA predictions based on one-years' data are applied to data for students two years later; for example, equations based on 1996-97 data for an institution would be used for applicants in fall 1998-99. However, to maximize sample sizes and the number of institutions for this study, the crossvalidation analysis was based on 1996-97 and 1997-98 data, rather than 1996-97 and 1998-99 data. In addition, Sawyer and Maxey (1980) found than crossvalidation statistics using a one-year period between base and crossvalidation years were very similar to those using a two-year period.

The analyses were based on data from institutions that had participated in ACT's Prediction Research Service in both 1996-97 and 1997-98. The base-year and crossvalidation-year files therefore consisted of data from 216 institutions: The base-year file consisted of 164,436 enrolled students and 528,082 nonenrolled students and the crossvalidation-year file consisted of 166,126 enrolled students and 539,241 nonenrolled students.

Method

Base-year and crossvalidation-year mean ACT Composite scores and mean HSAV values were computed by institution. Means were calculated for enrolled students, as well as for students in the entire applicant pool. Mean first-year GPAs were calculated by institution for students who completed the first year of college. Distributions of the means of these variables were then summarized across institutions for both the base year and the crossvalidation year using minimum, median, and maximum values.

Base-Year Logistic Regression Analysis

Logistic regression is a method for estimating the statistical relationship between a dichotomous criterion variable (e.g., first-year GPA of 2.0 or higher) and students' standing on one or more predictor variables (e.g., ACT Composite score or high school average). It follows the same general principles as linear regression, but fits a non-linear model with a predicted outcome bounded by 0 and 1. Sawyer (1996) used logistic regression and statistical decision theory to devise a method for validation that frames validity evidence in terms of probable outcomes, given the admissions measures and outcome criteria used.

Validity statistics are generated from logistic regression and frequency distributions of scores on the admissions measure(s) to determine the effectiveness of the admissions criteria. The validity statistics emphasize specific cutoff values on the admissions measure and on the outcomes (i.e., first-year GPA). For example, the validity statistics address whether or not students with a specific test score would be successful in college (e.g., first-year GPA of 2.0 or higher).

Admissions decisions are usually made based on multiple measures. ACT does not advocate making admissions decisions solely on the basis of a single measure; the use of one or two predictors and single cutoffs in this paper is a mathematical simplification. However, the methods used here, such as those used with the joint ACT and HSAV model, could be generalized to multiple measures.

We estimated three validity statistics for each predictor (or predictor combination) and cutoff score:

a. the maximum percentage of correct classifications (accuracy rate (AR)),

b. the percentage of successful students among those who would be expected to be successful (success rate (SR)), and

c. the increase in the percentage of correct classifications over expecting all applicants to be successful (increase in accuracy rate (ΔAR)).

Correct classifications include students scoring above the optimal cutoff who were successful and students scoring below the optimal cutoff who would have not been successful.

The percentage of all enrolled students with GPAs at or above each criterion level was also calculated for each institution. This percentage corresponds to the estimated accuracy rate resulting from expecting all applicants to be successful, regardless of test score or high school average. This value is referred to here as the "baseline" accuracy rate. The arithmetic difference between the maximum accuracy rate and the baseline accuracy rate represents the increase in accuracy rate (ΔAR) that results from using test scores or high school average.

Logistic regression models were constructed based on (a) ACT Composite score, (b) high school average (HSAV), and (c) ACT Composite and HSAV used jointly for predicting first-year success. The success criteria included first-year GPAs of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or higher. The logistic regression weights from each model were applied to the ACT Composite scores and/or HSAV values of all students at each institution with valid predictor data (i.e., the applicant pool), resulting in estimated probabilities of success for each student and model.

For each institution and success criterion, optimal base-year cutoffs were identified for the three types of predictor models (a)-(c). Optimal cutoffs are those that maximize the estimated percentages of correct classifications. It can be shown that optimal cutoffs also correspond to a .50 probability of success for a given model. For the two-predictor model, combinations of ACT Composite and HSAV cutoffs corresponding to a probability of success of .50 were identified. ARs, SRs, and ΔARs were then estimated for each predictor (or predictor combination) and optimal cutoff. All statistics were calculated through the conditional probabilities of each outcome for individual students in the applicant pool, as estimated by the regression models (Sawyer, 1996). For comparison purposes, median baseline accuracy rates (the percentages of all enrolled students with GPAs at or above each criterion level) were also reported. Distributions of these statistics were summarized across institutions using minimum, median, and maximum values.

As noted previously, a probability of .50 corresponds to the maximum accuracy rate. Probability distributions that cross .50 will yield accuracy rate distributions that increase to a maximum and then decrease. If the probability distribution for an institution does not cross .50, the maximum accuracy rate and optimal cutoff score are generally not interpretable. Therefore, in reporting maximum accuracy rates for all base-year predictor models and criterion levels, the probability distribution for each institution was required to cross .50.

Crossvalidation Analyses

The accuracy of predictions based on the base-year ACT Composite and high school average (HSAV) logistic regression models was assessed using the crossvalidation-year data. The logistic regression weights from each base-year model were applied to the ACT Composite scores and HSAV values of all applicants to each institution, resulting in estimated probabilities of success for each student and model. The *base-year* optimal cutoffs for each institution were then applied to the corresponding crossvalidation-year probability distributions, and crossvalidated ARs, SRs, and ΔARs were calculated (see p. 5 for descriptions of these statistics). For the two-predictor model, the logistic regression coefficients developed from the base year were applied to the crossvalidation-year applicant pool data to estimate probabilities of success. These probability "scores" were then used to calculate ARs, SRs, and ΔARs using a cutoff value of .50. Distributions of all crossvalidated statistics were summarized across institutions using minimum, median, and maximum values.

Results

Of the 216 institutions for which data were available for both 1996-97 and 1997-98, logistic regression models could be developed for all institutions and for all criterion values, with the exception of one institution for the 3.75 criterion. (For this institution, all students with a GPA of 3.75 or higher had HSAVs of 4.00.) Different criterion levels resulted in different numbers of institutions for which the fitted probability curves crossed .50, however:

- For all criterion levels except 2.00, successive increases in the criterion levels resulted in decreases in the numbers of institutions for which the fitted probability curves crossed .50, particularly for the HSAV models at the 3.25, 3.50, and 3.75 criterion levels.
- For the 2.00 criterion level, fewer institutions had fitted probability curves based on ACT Composite scores that crossed .50.

For the 2.50 to 3.50 criterion levels the final sample consisted of 84 institutions with 58,482 enrolled and 186,029 nonenrolled students for which all models and criterion levels could be evaluated. For the 2.00 criterion level, there were 58 institutions and 39,925 enrolled and 166,583 nonenrolled students for which all models could be developed. For the 3.75 criterion level, there were only 15 institutions for which an HSAV model could be developed. In comparison, ACT Composite and joint models could be developed for all 84 institutions. HSAV results are therefore not reported for the 3.75 criterion level. Results for the HSAV, ACT Composite, and the joint model for the 2.00 success criterion can be compared with each other, but they can not be compared to the results for other criterion levels. The substantial decline in the numbers of institutions is discussed later in this section.

Descriptive Statistics

The distributions of descriptive statistics across the 84 institutions are summarized in Table 1. For both enrolled students and the applicant pool, median, minimum, and maximum numbers of students, mean ACT Composite scores, HSAV values, and first-year GPA (enrolled students only) are reported for the base year and crossvalidation year.

Table 1

Distributions, Across Institutions, of Means and Standard Deviations of ACT Composite Scores, High School Grade Averages, and First-Year GPAs, by Year and Applicant/Enrollment Status

Enrollment status	Predictor variable	N		Mean		SD	
		Med	Min/Max	Med	Min/Max	Med	Min/Max
Base year							
Applicant pool	ACT Composite	1,183	219/19,675	20.6	17.5/26.0	3.97	3.35/4.81
	HSAV			3.10	2.76/3.65	.59	.37/.71
Enrolled students	ACT Composite	388	50/3,319	21.4	17.9/26.0	3.79	3.02/4.69
	HSAV			3.18	2.70/3.65	.57	.37/.76
	First-year GPA			2.63	2.30/3.13	.90	.55/1.28
Crossvalidation year							
Applicant pool	ACT Composite	1,268	227/21,386	20.9	17.7/25.4	4.07	3.27/4.81
	HSAV			3.14	2.73/3.59	.59	.44/.68
Enrolled students	ACT Composite	371	58/4,190	21.6	17.4/27.0	3.84	3.13/4.84
	HSAV			3.22	2.65/3.72	.57	.32/.71
	First-year GPA			2.69	2.05/3.13	.88	.48/1.48

As expected, for both years the institutions' mean ACT Composite scores and HSAVs were typically higher among enrolled students than among the students in the entire applicant pool. The corresponding standard deviations were smaller for enrolled students. The distributions of crossvalidation year mean ACT Composite score, HSAV, and first-year GPA showed slightly higher values than those for the base year.

Mean ACT Composite scores for enrolled students from both years were typically lower than those for first-year college students nationally (mean ACT Composite = 21.7; ACT, 1998e). Mean HSAVs were similar to those of first-year college students nationally (mean HSAV = 3.23; ACT, 1998e).

Base-Year Logistic Regression Results

Figures 1 and 2 show the median probabilities corresponding to all six criterion levels for the ACT Composite (Figure 1) and HSAV (Figure 2) models. The probabilities were summarized across the 216 institutions (215 institutions for the 3.75 criterion level) for which all three models could be developed.

As shown in Figure 1, the median probability distributions for all criterion levels ranged from near zero for an ACT Composite score of 1, to between .83 and .98 for an ACT Composite score of 36. A student with an ACT Composite score of 21 (the approximate median mean ACT Composite score across the 84 institutions) would typically have a .81 probability of a 2.00 first-year GPA. The corresponding probabilities for the other criterion levels would be .62 (2.50), .36 (3.00), .20 (3.25), .11 (3.50), and .04 (3.75), respectively.

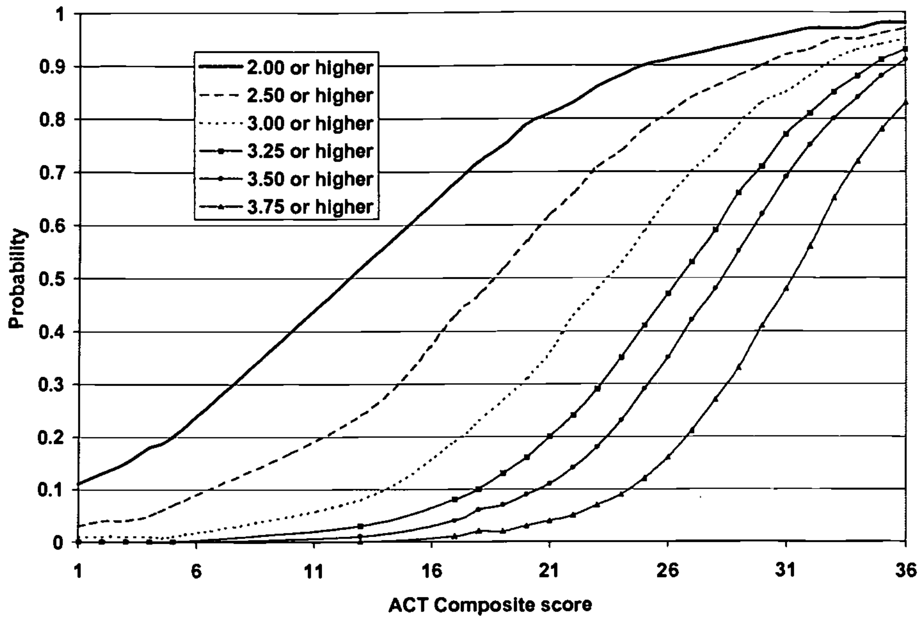

Figure 1. Median Probabilities of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or Higher First-Year GPA, Based on ACT Composite Score

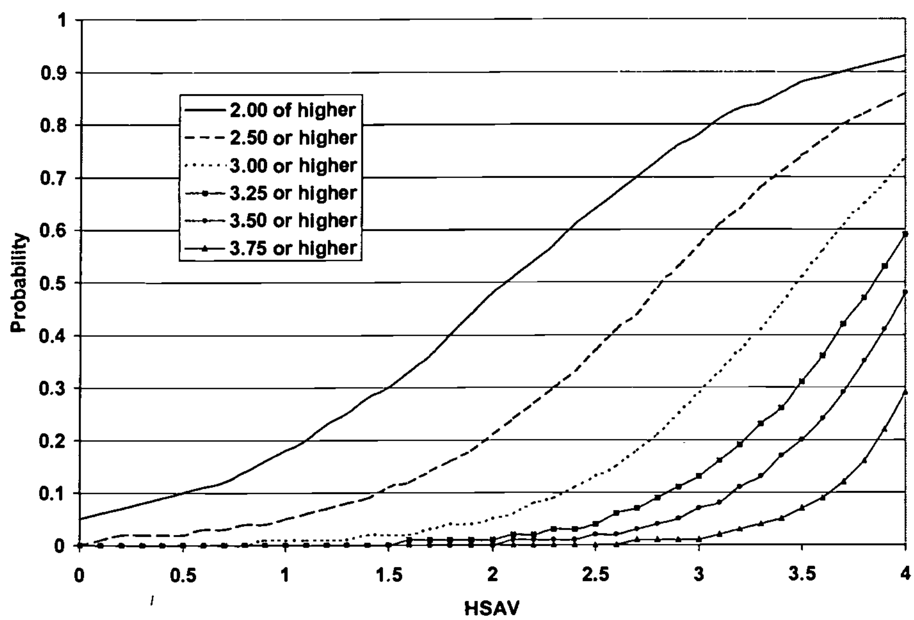

Figure 2. Median Probabilities of 2.00, 2.50, 3.00, 3.25, 3.50, and 3.75 or Higher First-Year GPA, Based on High School Average (HSAV)

As shown in Figure 2, the median probability distributions for HSAV ranged from near 0 (HSAV = 0), to between .29 and .93 (HSAV = 4.00). A student with an HSAV of 2.00 would typically have a .48 probability of a 2.00 or higher first-year GPA, and a .21 probability of a 2.5

or higher first-year GPA. The corresponding median probabilities for the other criterion levels would be .05 or lower. In comparison, a student with an HSAV of 3.2 (the approximate median mean HSAV across the 84 institutions) would typically have a .83 probability of a 2.00 or higher first-year GPA. The corresponding median probabilities for the other criterion levels would be .64 (2.50), .37(3.00), .19 (3.25), .11 (3.50), and .03 (3.75), respectively.

Note that for the criterion levels of 3.50 and 3.75, an HSAV of 4.00 corresponded to a median probability of success of less than .50. Moreover, for the criterion levels of 2.50 and 3.00, there was little difference in the median probabilities for HSAV values of less than 2.00. Similarly, for HSAV values of 2.50 for the 3.00, 3.25, 3.50, and 3.75 criterion levels, there was little difference in the corresponding median probabilities. Any substantive differentiation among students' probabilities across all criterion levels therefore appeared to occur between HSAV values of 3.0 and 4.0.

Loss of Institutions Due to Lack of Model Fit

A probability curve for a predictor variable is required to cross .50 to calculate a maximum accuracy rate. For example, the typical probability of a 3.25 or higher first-year GPA for a student with a 4.00 HSAV was less than .50. We refer to this type of model as a "nonviable model" for an institution. Models for institutions with probability curves crossing .50 are referred to as "viable models." Table 2 includes the numbers of viable and nonviable models, by predictor. Note, for the 3.75 criterion level at one institution, all students with GPAs of 3.75 or higher had HSAVs of 4.00.

Further examination of the institutions with nonviable models for the 3.50 criterion level revealed that all institutions with nonviable ACT Composite models also had nonviable HSAV models. For example, of the 118 institutions with nonviable HSAV models, 16 also had nonviable ACT Composite models (leaving 102 institutions with nonviable models using HSAV alone). All of the 98 institutions with viable HSAV models also had viable ACT Composite models. In comparison, for the 3.25 criterion level, of the 44 institutions with nonviable HSAV models, seven also had nonviable ACT Composite models. Of the 172 institutions with viable HSAV models, only one had a nonviable ACT Composite model.

Table 2

Number of Institutions Where Probability Distributions Did and Did Not Cross .50

		Number		
Criterion	Predictor variable	Viable models	Nonviable models	Total possible models
2.0 or higher	ACT Comp	151	65	216
	HSAV	207	9	216
	ACT Comp & HSAV	210	6	216
2.5 or higher	ACT Comp	214	2	216
	HSAV	216	0	216
	ACT Comp & HSAV	216	0	216
3.0 or higher	ACT Comp	211	5	216
	HSAV	209	7	216
	ACT Comp & HSAV	214	2	216
3.25 or higher	ACT Comp	208	8	216
	HSAV	172	44	216
	ACT Comp & HSAV	210	6	216
3.5 or higher	ACT Comp	200	16	216
	HSAV	98	118	216
	ACT Comp & HSAV	203	13	216
3.75 or higher	ACT Comp	168	47	215
	HSAV	15	200	215
	ACT Comp & HSAV	175	40	215

As can be seen in Table 2, there was a large number of nonviable models for HSAV, especially for the 3.25, 3.50, and 3.75 criterion levels. For these criterion levels, 20%, 55%, and 93%, respectively, of the HSAV models were nonviable, compared to 4%, 7%, and 22%, respectively, of the ACT Composite models. In contrast, for the 2.00 criterion level, 30% of the

ACT Composite models were nonviable, compared to 4% of the HSAV models. For all of the nonviable ACT models for this criterion level, all fitted probabilities of success exceeded .50.

The institutions with nonviable models were investigated further to attempt to explain the large discrepancy between the ACT Composite and HSAV results. The criterion level of 3.50 was targeted, because of the relatively large numbers of nonviable models for HSAV. The data file was matched to ACT's Institutional Data Questionnaire history (ACT, 1998b) to obtain descriptive information about these institutions. Of the 216 institutions possible, the 102 with nonviable models represented all regions of the country, both two-year and four-year institutions (29% two-year, 69% four-year, and 3% unable to determine), and all types of admissions policy (2% highly selective, 16% selective, 27% traditional, 13% liberal, 41% open, and 2% unable to determine). When compared with the institutions with viable models, very few differences were found in their ACT Composite scores, high school averages, or first-year GPA distributions.

Validity Statistics

Table 3 shows median baseline accuracy rates, optimal cutoff scores, estimated accuracy rates (ARs), estimated increases in accuracy rates (ΔARs), and estimated success rates (SRs) for the 84 institutions for which validity statistics could be calculated.

As one would expect, median optimal ACT Composite and HSAV cutoff scores increased across criterion levels from 2.00 to 3.75. For example, the median optimal ACT Composite score for a GPA level of 2.50 or higher was 18; the corresponding optimal cutoff scores for the other criterion levels were 22, 25, 27, and 30, respectively. Note, however, that statistics could not be calculated for the HSAV model for the 3.75 criterion level, due to the substantial numbers of institutions where the probability of a 3.75 or higher GPA for students with a 4.00 HSAV was less than .50. (The probability distributions did not cross .50 for 69 of the 84 institutions.)

Table 3

Medians, Across 84 Institutions, of Base-Year Logistic Regression Statistics

Criterion level	Baseline accuracy rate	Predictor variable	Optimal cutoff	Accuracy rate (AR)	Increase in accuracy rate (ΔAR)	Success rate (SR)
2.00 or higher GPA(1)	75	ACT Comp	14	.76	.00	.77
		HSAV	2.21	.79	.02	.80
		ACT Comp & HSAV(3)	--	.79	.02	.81
2.50 or higher GPA	61	ACT Comp	18	.69	.07	.70
		HSAV	2.78	.71	.09	.73
		ACT Comp & HSAV(3)	--	.74	.11	.75
3.00 or higher GPA	39	ACT Comp	22	.71	.31	.65
		HSAV	3.39	.73	.32	.67
		ACT Comp & HSAV(3)	--	.76	.36	.70
3.25 or higher GPA	25	ACT Comp	25	.79	.54	.63
		HSAV	3.73	.79	.52	.60
		ACT Comp & HSAV(3)	--	.81	.57	.67
3.50 or higher GPA	17	ACT Comp	27	.84	.67	.61
		HSAV	3.91	.83	.65	.55
		ACT Comp & HSAV(3)	--	.86	.69	.64
3.75 or higher GPA(2)	7	ACT Comp	30	.93	.85	.57
		HSAV	--	--	--	--
		ACT Comp & HSAV(3)	--	.93	.86	.59

Notes:
(1) All 2.00 models are based on 58 institutions for which the base-year fitted probability distributions based on ACT Composite scores crossed .50.
(2) HSAV prediction statistics could be calculated for only 15 of the 84 institutions
(3) A range of optimal combinations of ACT Composite score and HSAV correspond to a probability of .50 for the joint model.

Correspondingly, median baseline accuracy rates (median percentages of students with GPAs at or above each criterion level) decreased across all criterion levels. A relatively high percentage of students (median = 75%) had GPAs of 2.00 or higher.

For criterion levels of 2.00, 2.50, and 3.00, the median ARs and ΔARs indicated that the HSAV models were somewhat more accurate than the ACT Composite models. However, for criterion levels of 3.25 and 3.50, the median ARs for the ACT Composite equaled and then exceeded those for HSAV. (The corresponding median ΔARs were both somewhat higher for the ACT Composite models.) For the 3.75 model, the median ARs for the ACT Composite model and the joint model were identical, reflecting the small contribution of HSAV to the joint model. For all criterion levels except the 2.00 level, the median ARs and ΔARs for the ACT Composite & HSAV joint model exceeded those for the single-predictor ACT Composite and HSAV models.

Median SRs showed a similar result: For criterion levels of 2.00 and 2.50, median SRs for the HSAV model were higher than those for the ACT Composite. For all other criterion levels, the median SRs for the ACT Composite were higher than those for HSAV. For all criterion levels, median SRs for the joint model exceeded those for the HSAV and ACT Composite models.

Crossvalidation-Year Logistic Regression Results

The crossvalidated logistic regression results are shown in Table 4. Median baseline accuracy rates and crossvalidated estimated accuracy rates (ARs), increases in accuracy rates (ΔARs), and success rates (SRs) are shown for each criterion level. The median optimal cutoffs for the ACT Composite, HSAV, and joint models were the same as those for the base-year. For all criterion levels except 2.50 or higher, median baseline accuracy rates were slightly higher for the crossvalidation year than for the base year.

Table 4

Medians, Across 84 Institutions, of Crossvalidated Logistic Regression Statistics

Criterion level	Baseline accuracy rate	Predictor variable	Accuracy rate (AR)	Increase in accuracy rate (ΔAR)	Success rate (SR)
2.00 or higher GPA[1]	78	ACT Comp	.78	.00	.80
		HSAV	.79	.01	.83
		ACT Comp & HSAV	.80	.02	.83
2.50 or higher GPA	61	ACT Comp	.69	.06	.70
		HSAV	.72	.08	.74
		ACT Comp & HSAV	.73	.10	.76
3.00 or higher GPA	40	ACT Comp	.70	.29	.63
		HSAV	.73	.31	.66
		ACT Comp & HSAV	.75	.35	.70
3.25 or higher GPA	26	ACT Comp	.77	.50	.61
		HSAV	.78	.50	.61
		ACT Comp & HSAV	.81	.53	.68
3.50 or higher GPA	19	ACT Comp	.83	.65	.58
		HSAV	.82	.64	.54
		ACT Comp & HSAV	.86	.67	.67
3.75 or higher GPA[2]	8	ACT Comp	.92	.85	.53
		HSAV	--	--	--
		ACT Comp & HSAV	.93	.85	.71

Notes:
(1) All 2.00 models are based on 58 institutions for which the base-year fitted probability distributions based on ACT Composite scores crossed .50.
(2) HSAV prediction statistics could be calculated for only 15 of the 84 institutions.

The crossvalidated ARs and ΔARs were very similar to those for the base year: Differences between base-year and crossvalidated median AR did not exceed .02 for all three models. Differences between base-year and crossvalidated median ΔARs for all three models were .04 or less across all criterion levels.

The base-year results showed somewhat greater prediction accuracy for the HSAV model for criterion levels of 2.50 and 3.00, similar prediction accuracy at a criterion level of 3.25, and somewhat greater prediction accuracy for the ACT Composite at a criterion level of 3.50. The crossvalidated results differed slightly: The crossvalidated median ARs for the 3.25 level slightly favored the HSAV model (.77 vs. .78), although the median ΔARs were the same (.50) for the two models. Crossvalidated ARs and ΔARs continued to favor the ACT Composite for the 3.50 criterion level. For the 3.75 level, the crossvalidated median ΔARs for the ACT Composite and joint models were the same and the corresponding median ARs differed by .01 (.92 and .93, respectively), again illustrating the weak contribution of HSAV to the joint model for this success criterion.

Crossvalidated median SRs were slightly higher than were those for the base year for lower criterion levels (2.00 and 2.50). For higher criterion levels, median SRs were lower than were those for the base year, with the exception of the 3.75 criterion level. For this level, the median SR for the joint model was considerably higher for the crossvalidation year than for the base year. This result might be attributable to the higher baseline accuracy rates for the crossvalidation year sample.

Conclusions

An important finding of this study is the apparent inability of HSAV to predict high levels of academic achievement during the first year of college. For 20%, 55%, and 93% of the institutions, a 4.00 high school average corresponded to a probability of a 3.25, 3.50, and 3.75 or higher GPA of less than .50. Moreover, in some cases HSAV values of less than 3.00 provided little discrimination in terms of student's chances of achieving different first-year GPAs. This evidence suggests that noncognitive factors contribute significantly to high school grades lower than a B.

Consistent with prior research (ACT, 1997c), HSAV predictions of first-year GPAs of 2.50 and 3.00 were somewhat more accurate than those based on ACT Composite score. However, predictions based on ACT Composite score and HSAV jointly were more accurate than those based on HSAV or ACT Composite score alone. For higher first-year GPA criterion levels, ACT Composite score predictions were as accurate or more accurate than those associated with HSAV. These findings appear to support the conclusions of earlier research that:

1. College grades reflect achievement and noncognitive factors (Goldman, et al., 1974; 1975; 1976; Beck, 1999; Pothoven, 1993; Lambert, 1993; McSpirit, Kopacz, Jones, & Chapman, 2000; U.S. Department of Education, 1994), and
2. These noncognitive factors are less pervasive at higher achievement levels (Goldman et al., 1974; 1975; 1976).

Postsecondary institutions seek high achievement for their students, and want to admit students who have a good chance of being successful in college. Some people might criticize the use of standardized admissions tests, stating that such tests are appropriate only for "elite" students, and should not be used for the typical college-bound student. These results suggest that ACT Composite scores provide greater differentiation across levels of achievement than do high school averages in terms of students' probable success during their first year in college.

References

ACT (1997a). *ACT Class Profile Service*. Iowa City, IA: ACT, Inc.

ACT (1998a). *ACT Class Profile Service*. Iowa City, IA: ACT, Inc.

ACT (1998b). *ACT Institutional Data Questionnaire*. Iowa City, IA: ACT, Inc.

ACT (1997b). *ACT Prediction Research Service*. Iowa City, IA: ACT, Inc.

ACT (1998c). *ACT Prediction Research Service*. Iowa City, IA: ACT, Inc.

ACT (1998d). *Research Services Summary Tables*. Iowa City, IA: ACT, Inc.

ACT (1997c). *ACT technical manual*. Iowa City, IA: ACT, Inc.

ACT (1998e). *College student profiles*. Iowa City, IA: ACT, Inc.

Beck, B. (1999). *Trends in undergraduate grades*. A paper presented at the annual fall conference of the Association for Institutional Research in the Upper Midwest in ST. Paul, Minnesota.

Breland, H., Maxey, J., Gernand, R., Cumming, T. and Trapani, C. (In press) *Trends in college admissions: A report of a survey of undergraduate admissions policies, practices, and procedures.*

Goldman, R. D. & Hewitt, B. N. (1975). Adaptation-level as an explanation for differential standards in college grading. *Journal of Educational Measurement*, *12*(3), 149-161.

Goldman, R. D., Schmidt, D. E., Hewitt, B. N., & Fisher, R. (1974). Grading practices in different fields. *American Education Research Journal*, *11*(4), 343-357.

Goldman, R. D. & Widawski, M. H. (1976). A within-subjects technique for comparing college grading standards: Implications in the validity of the evaluation of college achievement. *Educational and Psychological Measurement*, *36*, 381-390.

Pothoven, C. (1993). *C's aren't what they used to be*. The Daily Iowan, pp. 1A, 3A.

Lambert, C. (1993). Desperately seeking summa. *Harvard Magazine*, 36-40.

McSpirit, S, Kopacz, P, Jones, K, & Chapman, A. (2000). Faculty opinion on grade inflation: Contradictions about its cause. *College & University Journal*, *75*(3), 19-25.

Maxey, E. J. (in press). *Selected trends in ACT-Tested Students*. Iowa City, IA: ACT, Inc.

Sawyer, R. L. (1996). Decision theory models for validating course placement tests. *Journal of Educational Measurement, 33,* 271-290.

Sawyer, R. L., & Maxey, E. J. (1980). The validity of college grade prediction equations over time. *Journal of Educational Measurement*, *16*(4), 279-284.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, *8*(2), 5-14.

U. S. Department of Education, Office of Educational Research and Improvement (1994). *What do student grades mean? Differences across schools* (Office of Research Rep. OR 94-3401). Washington, DC: Office of Research.

NOTICE

Reproduction Basis

- [x] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

- [ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").