Research Article

Individual Differences in Course Choice Result in Underestimation of the Validity of College Admissions Systems

Christopher M. Berry[1] and Paul R. Sackett[2]

[1]Wayne State University and [2]University of Minnesota

ABSTRACT—*We demonstrate that the validity of SAT scores and high school grade point averages (GPAs) as predictors of academic performance has been underestimated because of previous studies' reliance on flawed performance indicators (i.e., college GPA) that are contaminated by the effects of individual differences in course choice. We controlled for this contamination by predicting individual course grades, instead of GPAs, in a data set containing more than 5 million college grades for 167,816 students. Percentage of variance accounted for by SAT scores and high school GPAs was 30 to 40% lower when the criteria were freshman and cumulative GPAs than when the criteria were individual course grades. SAT scores and high school GPAs together accounted for between 44 and 62% of the variance in college grades. This study provides new estimates of the criterion-related validity of SAT scores and high school GPAs, and highlights the care that must be taken in choosing appropriate criteria in validity studies.*

When making admissions decisions, many colleges rely heavily on students' high school grade point averages (GPAs) and scores on standardized tests (e.g., the SAT and ACT; Breland, Maxey, Gernard, Cumming, & Trapani, 2002; Hawkins & Clinedinst, 2006). However, many people have criticized colleges' admissions systems, and especially the use of admissions tests. One prominent criticism is that test scores do not adequately capture applicants' potential for academic performance (e.g., Rooney, 1998; Sternberg, 2004). The same can be said of high school

GPA, as it typically predicts college grades about as well as test scores do (e.g., Bridgeman, McCamley-Jenkins, & Ervin, 2000; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Given the prominence of SAT scores and high school GPAs in college admissions, and given the criticisms leveled against them, accurate assessment of their predictive validity is crucial. The present study demonstrates that previous studies have underestimated the predictive validity of SAT scores and high school GPAs because they have used college GPA as the criterion, and college GPA is contaminated by individual differences in course selection.

Although prediction of college GPA is often implicitly treated as the ultimate goal of college admissions systems, we posit that SAT scores and high school GPAs are actually used with the goal of predicting a construct called *academic performance*, which refers to how well students do in academic pursuits while in college. College GPA has been the typical operationalization of academic performance used by researchers. College GPA certainly reflects academic performance to some degree, but there are also well-known sources of construct-irrelevant variance in GPA—particularly instructors' grading idiosyncrasies and differences between students in course choice (e.g., Elliott & Strenta, 1988; Ramist, Lewis, & McCamley, 1990; Willingham, 1985). Thus, one might ask what the validity of SAT scores and high school GPAs would be if the criterion were a measure of academic performance that is more comparable from student to student.

We examined this question by predicting individual course grades (ICGs), and comparing the results with how well SAT scores and high school GPAs predicted college GPA. Although a 3.0 GPA may carry different meanings for two students (e.g., one student may have earned grades in difficult courses, whereas the other may have earned grades in lenient courses), a grade of 3.0

Address correspondence to Christopher Berry, Wayne State University, Department of Psychology, 5057 Woodward Ave., 7th Floor, Detroit, MI 48202, e-mail: berry@wayne.edu.

(or B) earned in the same course should hold more comparable meaning for two students. Thus, using ICGs instead of GPA should yield a less contaminated criterion by holding differences in course selection constant. Certainly, if one is simply interested in predicting GPA, then predicting ICGs is not appropriate. However, if one is interested in assessing how well SAT scores and high school GPAs predict academic performance, of which GPA is just an imperfect indicator, then analyses using ICGs as the criterion should provide a better estimate. This is not to say that ICGs are a perfect measure of academic performance. Grades are imperfect measures of learning. However, ICGs should provide a better approximation of academic performance than GPA.

Ramist et al. (1990) used ICG criteria in their validity study, correlating SAT scores and high school GPAs with ICGs from 4,680 college courses. The sample-size-weighted average of the 4,680 SAT-ICG correlations (corrected for range restriction) was .49, and the average correlation between high school GPA and ICGs was .47. Because these correlations were calculated for single courses, and because single grades are typically less reliable than the composite of GPA, Ramist et al. applied the Spearman-Brown formula to estimate what the validity of SAT scores and high school GPAs would be if these measures were correlated with a composite of 8.6 ICGs (thus putting ICGs on equal footing with freshman GPA, given that freshman GPA was a composite of, on average, 8.6 courses per student in the data set). This calculation simulated what the validity of SAT scores and high school GPAs would be if the courses making up freshman GPA were more comparable. The resulting validity estimates, correlations of .75 and .72 for SAT scores and high school GPAs, respectively, were quite high compared with the correlations obtained for freshman GPA, which were .53 and .57 for SAT scores and high school GPAs, respectively.

However, there is reason to suspect that the correlations Ramist et al. (1990) obtained for ICGs were overestimates. Ramist et al. excluded negative correlations. Even if the true correlations were above zero, sampling error alone could have caused some correlations to fluctuate enough to fall below zero (especially given that Ramist et al. included courses with as few as five students). Thus, excluding negative correlations may have biased the estimated validity of SAT score and high school GPA upward.

THE PRESENT STUDY

Our purposes in the present study were threefold. First, we calculated correlations of SAT scores and high school GPAs with ICGs using a much larger sample than Ramist et al. (1990; our sample included grades from 145,000+ courses), and without excluding negative correlations. Thus, the present study stands as a more stable and accurate test of the validity of college admissions systems.

Second, we compared how well SAT scores and high school GPAs predicted ICGs with how well they predicted two common GPA criteria: college-freshman GPA and cumulative GPA throughout the college career. Both of these GPA criteria are affected by individual differences in course choice. However, this may be especially true of cumulative GPA because it includes grades earned throughout college, and grades earned later in the college career tend to be less predictable from SAT score and high school GPA (Willingham, 1985).

Third, we assessed the validity of SAT scores and high school GPAs when used in conjunction to predict ICGs, as most colleges use some combination of the two in the admissions process. This also allowed us to examine the incremental validity of each measure over the other in predicting ICGs. A typical finding is that high school GPA accounts for more variance in college GPA than do SAT scores (e.g., Bridgeman et al., 2000). However, Ramist et al. (1990) reported that SAT scores accounted for more variance in ICGs than did high school GPAs. We revisited these incremental-validity analyses using a larger data set and without excluding negative correlations.

METHOD

Participants

Participants were drawn from a College Board data set on 167,816 college students representing the entering classes at 41 U.S. colleges in 1995 through 1997 (see Sackett et al., 2009, for descriptions of the 41 colleges).

Measures

SAT Scores

The College Board provided students' scores on the verbal and math subtests of the SAT. We combined these scores into a unit-weighted composite for each student.

High School GPAs

An item in a questionnaire students completed at the time they took the SAT asked them to report their high school GPA. A meta-analysis by Kuncel, Crede, and Thomas (2005) found a strong relationship (mean $r = .82$) between self-reported and high-school-reported GPAs. Kuncel et al. also made the case that self-reported GPAs tend to predict outcomes as well as school-reported GPAs, especially among students with relatively high GPAs in high school, such as those in the present study.

Cumulative GPAs

Colleges provided cumulative GPAs across all courses taken by participants throughout up to their first 6 years in college.

Freshman GPAs

Colleges provided GPAs for participants' freshman year in college.

ICGs

Each college provided ICGs (scale from 0.0 to 4.3) received by participants throughout up to their first 6 years in college. The resulting data set included more than 5.1 million grades earned in more than 145,000 courses. Associated with each ICG was a college, a course title, and the year of the course. Any ICGs for courses with the exact same title, college, and year were treated as earned in the same course. Note that these data permitted only partial control over differences in instructors' grading standards, as the data set did not differentiate between multiple sections of the same course taught in the same year.

Procedure

We assessed the extent to which individual differences in course selection result in underestimation of the validity of SAT score and high school GPA in predicting academic performance when the latter is measured by two common GPA criteria: cumulative GPA and freshman GPA. Thus, we first estimated the validity of SAT score and high school GPA in predicting cumulative GPA and freshman GPA in college, and then estimated the validity of SAT score and high school GPA in predicting conceptually comparable ICG criteria. Because cumulative GPA reflects courses taken throughout college, and freshman GPA reflects only courses taken in the freshman year, we carried out two separate sets of analyses using ICGs. The first set used all ICGs earned throughout college as criteria and thus represented the ICG analogue to cumulative GPA. The second set used only ICGs earned in the freshman year, and thus represented the ICG analogue to freshman GPA.

GPA Analyses

Within each college, we calculated the correlations of SAT score with cumulative GPA and freshman GPA and the correlations of high school GPA with cumulative GPA and freshman GPA. We then separately meta-analyzed these four sets of within-college GPA correlations using Hunter and Schmidt's (2004) method, the accuracy of which has been upheld in multiple simulation studies and comparative analyses (e.g., Burke, Raju, & Pearlman, 1986; Schmidt, Oh, & Hayes, in press). In the present circumstance, this method involved (a) computing the sample-size-weighted mean and variance of the correlations, (b) computing sampling-error variance, and (c) subtracting sampling-error variance from sample-size-weighted variance to estimate net variance across correlations (i.e., residual variance). Thus, this meta-analysis yielded four mean correlations and their estimated variances.

These four mean correlations were affected by range restriction. Range restriction refers to the reduction in variance when the study sample has been selected on the basis of scores on the variable in question (e.g., computing SAT-GPA correlations in samples of students selected on the basis of SAT scores) or on the basis of a variable correlated with the variable of interest (e.g.,

computing SAT-GPA correlations in samples selected using high school GPAs, which are correlated with SAT scores). Restricted variance in SAT scores and high school GPAs results in lower correlations of these variables with college GPA than would be the case if the correlations were based on the full samples of college applicants. Given that multiple variables (i.e., both SAT scores and high school GPAs) were used in selecting applicants, we used multivariate range-restriction corrections (Sackett & Yang, 2000).

Thus, each college's GPA correlations were corrected for range restriction using the Pearson-Lawley multivariate correction (Gulliksen, 1950, pp. 165–166). In the multivariate restriction scenario, there is a set of variables for which the unrestricted standard deviations and intercorrelations are known, and another set of variables for which only restricted standard deviations and intercorrelations are known. In the present setting, unrestricted data on two variables known prior to college entry (SAT score, high school GPA) were available. However, because college GPAs were available only for students who were selected into the 41 colleges and then enrolled in those colleges, only restricted correlations with SAT and high school GPA were known for the college GPA variables, and the range-restriction correction was used to estimate the unrestricted correlations.

It was therefore important to carefully choose the population from which we estimated the unrestricted standard deviations and correlations. Students choose which colleges to apply to in part on the basis of how their SAT scores and high school GPAs match colleges' standards; therefore, variability in SAT scores and high school GPAs will be smaller within any given college's applicant pool than in the total population of college applicants. Correcting for range restriction using each college's applicant-pool standard deviations and correlations as unrestricted values estimates how well SAT score and high school GPA could be expected to predict grades within the average college's applicant pool. Correcting for range restriction using standard deviations and correlations drawn from the entire population of college applicants estimates how well SAT score and high school GPA could be expected to predict grades if variance in SAT scores and high school GPAs were not reduced by students' self-selection into colleges' applicant pools. Neither correction is necessarily "correct"; both can help one answer important questions about how well SAT score and high school GPA predict college grades. Thus, we present results based on the two types of corrections separately.

Therefore, we obtained two separate sources of information regarding unrestricted standard deviations and correlations. First, we obtained standard deviations and correlations between SAT score and high school GPA among the entire population of individuals taking the SAT in 1995 through 1997 (more than 2.5 million students). This facilitated estimation of what the correlation between SAT score and high school GPA would be if self-selection into college applicant pools had not occurred. So, in

this case, the population correlations we estimated were the correlations among all students who took the SAT. For this *national-population correction*, we used the Pearson-Lawley correction. The GPA correlations for each college were corrected using the standard deviations and correlations for SAT score and high school GPA in the overall population. The resulting within-college GPA correlations were then sample-size-weighted and averaged to arrive at overall corrected correlations between SAT score and cumulative GPA, SAT score and freshman GPA, high school GPA and cumulative GPA, and high school GPA and freshman GPA.

Second, we estimated standard deviations and correlations in the applicant pool for each specific college. So, in this case, the population correlations we estimated were the correlations within each college's specific applicant pool. Although such data were not directly available, we used a reasonable proxy. When taking the SAT, students indicate the colleges to which their scores should be sent; the set of students who requested that their scores be sent to a given college was used as the estimate of that college's applicant pool. For our *college-specific correction*, we again used the Pearson-Lawley correction. The GPA correlations within each college were corrected using the standard deviations and correlations for SAT score and high school GPA within that college's applicant pool. The resulting 41 corrected correlations between SAT score and cumulative GPA, between SAT score and freshman GPA, between high school GPA and cumulative GPA, and between high school GPA and freshman GPA were sample-size-weighted and averaged to arrive at overall corrected correlations.

GPA-Versus-ICG Analyses

To compare results obtained for ICG versus GPA criteria, we needed to create individual-course-level analogues to cumulative GPA and freshman GPA. For every course in our data set that included at least three students, we calculated the correlations between grades in that course and SAT scores and high school GPAs. This resulted in 148,072 correlations between SAT score and ICG and 145,876 correlations between high school GPA and ICG. We meta-analyzed these two sets of correlations separately using Hunter and Schmidt's (2004) method. We corrected the mean correlations for range restriction using the same methods and unrestricted standard deviations and correlations as described for the GPA analyses.

The range-restriction-corrected correlations with ICG represented average correlations with grades in a single course. However, in our data set, cumulative GPA was a composite of grades in an average of 33.36 courses. In order to directly compare validity for ICGs against validity for cumulative GPA, we used Ghiselli, Campbell, and Zedeck's (1981, pp. 163–164) formula for estimating the correlation between a composite and an outside variable to estimate what the correlations with ICGs would be if ICGs were a composite composed of as many grades as cumulative GPA (ICG composite). The formula required the

range-restriction-corrected correlation between SAT score and ICG or between high school GPA and ICG, the number of composite components (33.36), and the intercorrelation of the composite components (i.e., ICGs; $\bar{r} = .29$ uncorrected, $\bar{r} = .35$ range-restriction-corrected). After application of the formula from Ghiselli et al., the resulting correlations with the conceptual ICG composite estimated what the correlations of SAT scores and high school GPA with cumulative GPA would be if cumulative GPA comprised 33.36 parallel courses, thus controlling for individual differences in course choice. Because this ICG composite is an individual-course-level analogue to cumulative GPA and should represent a closer approximation of academic performance, we refer to it as the *cumulative academic-performance composite*, or cumulative APC.

To compare validity for GPA with validity for ICGs in the case of freshman grades, we made two key changes. First, only ICGs earned by students in their freshman year were used as criteria. This selection resulted in 10,117 correlations between SAT scores and ICGs and 9,989 correlations between high school GPA and ICGs. Second, when applying the composite formula from Ghiselli et al. (1981), we used the mean number of courses taken by students in their freshman year (9.92) and the mean intercorrelation of freshman grades ($\bar{r} = .38$ uncorrected, $\bar{r} = .42$ range-restriction-corrected) to create a conceptual APC analogue to freshman GPA, hereafter referred to as freshman APC.

Multivariate Relationships With APC

We also investigated course-level validity of SAT score and high school GPA in combination, using multiple regression. For these regression analyses, we first calculated the correlations between SAT score and high school GPA using the same methods as described for the correlations between SAT score and ICG. Combining this correlation with the results of the previous analyses, we created meta-analytic range-restriction-corrected correlation matrices among SAT score, high school GPA, and APC (four matrices in all: two each for cumulative and freshman APC, one using the national-population correction and one using the college-specific correction). These four matrices were used to regress APC on SAT scores and high school GPAs.

RESULTS

Cumulative GPA Versus Cumulative APC

The validities of SAT score and high school GPA in predicting cumulative GPA were .424 and .472, respectively, when we used the college-specific range-restriction correction and .509 and .546, respectively, when we used the national-population correction (see Table 1). These validities can be directly compared with those obtained in analyses using cumulative APC as the criterion (see Table 1). When we used the college-specific range-restriction correction, SAT score correlated .545 with APC, and high school GPA correlated .592 with APC. Thus, because of individual differences in course choice, the cumu-

TABLE 1

SAT Score and High School Grade Point Average (GPA) as Predictors of College GPA Versus Individual-Course-Grade (ICG) Criteria

Predictor	k	N	$\bar{r}$	SD_r	SD_e	SD_{res}	$\bar{r}_{RR}$: college[a]	$\bar{r}_{RR}$: national[a]
			Criterion: cumulative GPA					
High school GPA	41	158,044	.371	.104	.015	.103	.472	.546
SAT score	41	166,559	.328	.077	.015	.076	.424	.509
			Criterion: all ICGs earned throughout college					
High school GPA	145,876	4,838,923	.254	.201	.171	.115	.361 → .592	.433 → .710
SAT score	148,072	5,119,221	.224	.215	.168	.139	.332 → .545	.410 → .672
			Criterion: freshman GPA					
High school GPA	41	157,307	.382	.102	.015	.101	.485	.561
SAT score	41	165,781	.357	.074	.015	.073	.457	.543
			Criterion: freshman ICGs					
High school GPA	9,989	568,732	.293	.158	.128	.100	.412 → .593	.489 → .704
SAT score	10,117	599,824	.257	.182	.127	.134	.381 → .548	.464 → .668

Note. k = number of samples; N = total sample size; $\bar{r}$ = sample-size-weighted mean correlation; SD_r = sample-size-weighted standard deviation of correlations; SD_e = standard deviation due to sampling error; SD_{res} = standard deviation after subtracting out variance due to sampling error; $\bar{r}_{RR}$: college = mean correlation corrected for multivariate range restriction using each college's unrestricted applicant-pool standard deviations and correlations; $\bar{r}_{RR}$: national = mean correlation corrected for multivariate range restriction using the standard deviations and correlations for the entire population of SAT test takers.
[a]In these columns, when two numbers are connected by an arrow, the first number is the average range-restriction-corrected correlation with ICG, and the second number is an estimate of what the correlation would be if ICG were a composite of as many courses as the corresponding GPA measure. There is only one number for the GPA correlations because no composite correlations needed to be estimated.

lative-GPA criterion underestimated the percentage of variance accounted for by SAT scores by 39.5% ([$.545^2 - .424^2$]/$.545^2 = .395$) and underestimated the percentage of variance accounted for by high school GPAs by 36.4% ([$.592^2 - .472^2$]/$.592^2 = .364$). When we used the national-population correction, the correlations of APC with SAT score and high school GPA were .672 and .710, respectively. In this case, the cumulative-GPA criterion underestimated variance accounted for by SAT scores by 42.6% ([$.672^2 - .509^2$]/$.672^2 = .426$) and underestimated variance accounted for by high school GPAs by 40.9% ([$.710^2 - .546^2$]/$.710^2 = .409$).

Table 2 lists validity results for cumulative GPA versus cumulative APC when SAT scores and high school GPAs were used in conjunction as predictors. SAT scores and high school GPAs together accounted for between 27.7 and 36.3% of the variance in cumulative GPA, depending on whether the college-specific or national-population correction was used. In contrast, SAT scores and high school GPAs together accounted for between 44.4 and 62.2% of the variance in cumulative APC. Thus, the use of cumulative GPA as a criterion underestimated the variance jointly accounted for by SAT scores and high school GPAs by between 37.6 and 41.6% ([$44.4 - 27.7$]/$44.4 = .376$; [$62.2 - 36.3$]/$62.2 = .416$). SAT score and high school GPA accounted for incremental variance beyond each other both when cumulative GPA was the criterion and when cumulative APC was the criterion, but accounted for more incremental variance when cumulative APC was the criterion (cumulative GPA: ΔR^2s =

.054–.104; cumulative APC: ΔR^2s = .094–.151). In either case, high school GPA held a predictive advantage, contrary to the results of Ramist et al. (1990).

Freshman GPA Versus Freshman APC

The validities of SAT score and high school GPA in predicting freshman GPA were .457 and .485, respectively, when we used the college-specific correction and .543 and .561, respectively, when we used the national-population correction (see Table 1). These validities can be directly compared with those for freshman APC (see Table 1). When we used the college-specific range-restriction correction, the correlations of freshman APC with SAT score and with high school GPA were .548 and .593, respectively. Thus, because of individual differences in course choice, the freshman-GPA criterion underestimated variance accounted for by SAT scores by 30.5% ([$.548^2 - .457^2$]/$.548^2 = .305$) and underestimated variance accounted for by high school GPAs by 33.1% ([$.593^2 - .485^2$]/$.593^2 = .331$). When we used the national-population correction, the correlations of freshman APC with SAT score and with high school GPA were .668 and .704, respectively. Thus, the freshman-GPA criterion underestimated variance accounted for by SAT scores by 33.9% ([$.668^2 - .543^2$]/$.668^2 = .339$) and underestimated variance accounted for by high school GPAs by 36.5% ([$.704^2 - .561^2$]/$.704^2 = .365$).

Table 3 lists validity results for freshman GPA versus freshman APC when SAT scores and high school GPAs were used in

TABLE 2

High School Grade Point Average (GPA) and SAT Score as Predictors of the Cumulative Academic Performance Composite (APC) Versus Cumulative GPA: Results of Hierarchical Regression Analyses

Analysis and predictor	β	R	R^2	ΔR^2
	Criterion: cumulative GPA			
Incremental validity of SAT score over high school GPA				
Step 1		.472 (.546)	.223 (.298)	
High school GPA	.472 (.546)			
Step 2		.526 (.602)	.277 (.363)	.054 (.065)
High school GPA	.351 (.383)			
SAT score	.262 (.302)			
Incremental validity of high school GPA over SAT score				
Step 1		.424 (.509)	.180 (.259)	
SAT score	.424 (.509)			
Step 2		.526 (.602)	.277 (.363)	.097 (.104)
High school GPA	.351 (.383)			
SAT score	.262 (.302)			
	Criterion: cumulative APC			
Incremental validity of SAT score over high school GPA				
Step 1		.592 (.710)	.350 (.504)	
High school GPA	.592 (.710)			
Step 2		.667 (.788)	.444 (.622)	.094 (.118)
High school GPA	.433 (.490)			
SAT score	.346 (.407)			
Incremental validity of high school GPA over SAT score				
Step 1		.545 (.710)	.297 (.504)	
SAT score	.545 (.710)			
Step 2		.667 (.788)	.444 (.622)	.147 (.151)
High school GPA	.433 (.490)			
SAT score	.346 (.407)			

Note. Numbers outside parentheses reflect the college-specific range-restriction correction, and numbers inside parentheses reflect the national-population range-restriction correction. Correlation matrices for these regression results were drawn from Table 1, except for the correlations between SAT score and high school GPA, which were .253 (uncorrected), .461 (corrected using the college-specific correction), and .540 (corrected using the national-population correction).

conjunction as predictors. SAT scores and high school GPAs together accounted for between 30.4 and 39.6% of the variance in freshman GPA, depending on whether the college-specific or national-population correction was used. In contrast, SAT scores and high school GPAs together accounted for between 44.7 and 61.3% of the variance in freshman APC. Thus, the use of freshman GPA as a criterion underestimated the variance jointly accounted for by SAT scores and high school GPAs by between 32.0 and 35.4% ([44.7 − 30.4]/44.7 = .320; [61.3 − 39.6]/61.3 = .354). SAT score and high school GPA accounted for incremental variance beyond each other both when freshman GPA was the criterion and when freshman APC was the criterion, but accounted for more incremental variance when freshman APC was the criterion (freshman GPA: ΔR^2s = .069–.101; freshman APC: ΔR^2s = .096–.166). In either case, high school GPAs held a predictive advantage over SAT scores.

DISCUSSION

Four findings of this study are particularly noteworthy. The first concerns the validity of college admissions systems. We made the case that validity has been underestimated because researchers have used GPA criteria contaminated by individual differences in course choice. We accounted for this contamination by using the exact same grades that go into GPA, but calculating validity at the individual-course level and forming APCs. These APCs estimated what the validity of SAT score and high school GPA would be if the GPA criteria comprised parallel courses. Our individual-course-level analyses demonstrated that variance in grades accounted for by SAT score and high school GPA was underestimated by 30 to 40% when cumulative or freshman GPA was the criterion. Depending on the corrections used, validity (*r*) estimates increased from between .42 and

TABLE 3

High School Grade Point Average (GPA) and SAT Score as Predictors of the Freshman Academic Performance Composite (APC) Versus Freshman GPA: Results of Hierarchical Regression Analyses

Analysis and predictor	β	R	R^2	ΔR^2
Criterion: freshman GPA				
Incremental validity of SAT score over high school GPA				
Step 1		.485 (.561)	.235 (.315)	
High school GPA	.485 (.561)			
Step 2		.552 (.629)	.304 (.396)	.069 (.081)
High school GPA	.348 (.378)			
SAT score	.296 (.339)			
Incremental validity of high school GPA over SAT score				
Step 1		.457 (.543)	.209 (.295)	
SAT score	.457 (.543)			
Step 2		.552 (.629)	.304 (.396)	.096 (.101)
High school GPA	.348 (.378)			
SAT score	.296 (.339)			
Criterion: freshman APC				
Incremental validity of SAT score over high school GPA				
Step 1		.593 (.704)	.392 (.496)	
High school GPA	.593 (.704)			
Step 2		.669 (.783)	.447 (.613)	.096 (.117)
High school GPA	.432 (.485)			
SAT score	.349 (.406)			
Incremental validity of high school GPA over SAT score				
Step 1		.548 (.668)	.300 (.446)	
SAT score	.548 (.668)			
Step 2		.669 (.783)	.447 (.613)	.147 (.166)
High school GPA	.432 (.485)			
SAT score	.349 (.406)			

Note. Numbers outside parentheses reflect the college-specific range-restriction correction, and numbers inside parentheses reflect the national-population range-restriction correction. Correlation matrices for these regression results were drawn from Table 1, except for the correlations between SAT score and high school GPA, which were .239 (uncorrected), .461 (corrected using the college-specific correction), and .540 (corrected using the national-population correction).

.56, with GPA criteria, to between .55 and .71, when the less-contaminated APC criteria were used. The use of GPA criteria also underestimated the variance jointly accounted for by SAT scores and high school GPAs by 30 to 40%. Although the individual-course-level validities represent averages across the 41 colleges, we also ran these analyses within colleges (these results are not presented here because of space concerns), and individual-course-level validities were consistently substantial and positive across all 41 colleges.

Therefore, caution must be taken in interpreting GPA correlations as indices of the predictive validity of college admissions systems. Previous studies using GPA criteria have likely underestimated the validity of SAT scores and high school GPA by amounts similar to those determined in the present study. Thus, one policy implication of this study is that correlations with GPA criteria should not necessarily be the gold standards for determining the validity of college admissions systems.

Our results also indicate that SAT scores and high school GPAs have impressive criterion-related validity, as these measures accounted for as much as half, or even more than half, of the variance in college grades. Thus, to the degree that prediction of grades is a goal of college admissions systems, SAT scores and high school GPA are clearly useful tools for deciding which college applicants will achieve the greatest levels of academic performance. However, there is remaining variance left unexplained, and we encourage researchers to continue attempts to account for this remaining variance (e.g., Camara & Kimmel, 2005). Given the levels of validity in the present study, we believe that the search for additional predictors of college grades should be framed in terms of a search for useful supplements to, rather than replacements for, SAT score and high school GPA.

Our second noteworthy finding is that levels of predictive validity were similar for freshman and cumulative grades once

contamination due to individual differences in course choice was controlled. Although SAT score and high school GPA predicted freshman GPA better than cumulative GPA, this predictive advantage was erased when the criteria were freshman and cumulative APC. It makes sense that cumulative GPA would be more contaminated by individual differences in course choice than freshman GPA is, because upperclassmen (a) have more courses to choose from than freshmen and (b) have likely become savvier in choosing courses commensurate with their academic ability (Willingham, 1985). Our results do not support the claim that college admissions systems predict freshman grades only (e.g., Kohn, 2001; Rothstein, 2001).

Our third noteworthy finding is the confirmation that SAT scores and high school GPAs have incremental validity over each other. Evaluators know more about an applicant's potential for academic performance in college when they know both the applicant's SAT scores and his or her high school GPA than when they have only one of these pieces of information. A policy implication is that simultaneous use of the two measures is warranted in college admissions decisions. At the same time, as has typically been reported, and contrary to the results obtained by Ramist et al. (1990), high school GPAs predict grades slightly better than SAT scores do.

Our fourth noteworthy finding is that the results differed depending on whether we used the college-specific or the national-population range-restriction correction. Although the patterns of results did not change, the magnitudes did. Validity estimates were 16 to 23% higher when we used the national correction (e.g., in the first row of Table 1, the national-population correlation of .546 is 16% larger than the college-specific correlation of .472—i.e., [.546 − .472]/.472 = .157). The national-population correction estimates what SAT validity would be if range restriction had not occurred and if each college's applicant pool were the entire national population of SAT test takers, that is, if students did not self-select into certain colleges' applicant pools. The college-specific correction estimates what SAT validity would be in each college if range restriction had not occurred, that is, if all students who applied to a given college had enrolled. Thus, self-selection into college applicant pools reduced estimates of the validity of SAT scores and high school GPA between 16 and 23%. Regardless, neither correction is necessarily better than the other; they address different questions. If one is interested in how well SAT score and high school GPA would predict performance if student self-selection into applicant pools did not reduce variance, then .67 and .71, respectively, are the best estimates, according to our results. Part of the value of SAT scores and high school GPAs is that these indices assist students in selecting the colleges to which they should apply. Thus, .67 and .71 are upper-bound estimates of the "total value" of SAT score and high school GPA. However, if one is interested in how well the average college could expect SAT and high school GPA to predict academic performance in its own applicant pool, then

our results suggest that .55 and .59, respectively, are the best estimates.

Given the prominent role that SAT scores and high school GPA play in college admissions, and given the frequent criticism of the use of these indices, it is crucial to accurately estimate their validity. Previous studies using GPA criteria have underestimated the validity of SAT scores and high school GPAs because of criterion contamination in the form of individual differences in course choice. We controlled for this form of criterion contamination by predicting millions of ICGs across more than 145,000 college courses, and arrived at estimates 30 to 40% higher than when we used freshman or cumulative GPA as the criterion. Our results highlight the need to take care in choosing appropriate criteria in validity studies.

REFERENCES

Breland, H., Maxey, J., Gernard, R., Cumming, T., & Trapani, C. (2002, March). *Trends in college admission 2000: A report of a national survey of undergraduate admission policies, practices, and procedures*. Retrieved June 25, 2008, from http://www.airweb.org/images/trendsreport.pdf

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning Test* (College Board Research Report No. 2000-1). New York: College Entrance Examination Board.

Burke, M.J., Raju, N.S., & Pearlman, K. (1986). An empirical comparison of the results of five validity generalization procedures. *Journal of Applied Psychology, 71*, 349–353.

Camara, W.J., & Kimmel, E.W. (2005). *Choosing students: Higher education admissions tools for the 21st century*. Mahwah, NJ: Erlbaum.

Elliott, R., & Strenta, A.C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement, 25*, 333–347.

Ghiselli, E.E., Campbell, J.P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W.H. Freeman.

Gulliksen, H. (1950). *Test theory of mental tests*. New York: John Wiley and Sons.

Hawkins, D.A., & Clinedinst, M. (2006). *State of college admission 2006*. Retrieved June 25, 2008, from http://www.nacacnet.org/PublicationsResources/Research/Documents/06StateofAdmission.pdf

Hunter, J.E., & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Kohn, A. (2001, March 9). Two cheers for an end to the SAT. *The Chronicle of Higher Education*. Retrieved July 29, 2008, from http://www.alfiekohn.org/teaching/sat.htm

Kuncel, N.R., Crede, M., & Thomas, L.L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A

meta-analysis and review of the literature. *Review of Educational Research*, *75*, 63–82.

Ramist, L., Lewis, C., & McCamley, L. (1990). Implications of using freshman GPA as the criterion for the predictive validity of the SAT. In W.W. Willingham, C. Lewis, R. Morgan, & L. Ramist (Eds.), *Predicting college grades: An analysis of institutional trends over two decades* (pp. 253–288). Princeton, NJ: Educational Testing Service.

Rooney, C. (1998). *Test scores do not equal merit: Enhancing equity and excellence in college admissions by deemphasizing SAT and ACT results*. Cambridge, MA: FairTest.

Rothstein, R. (2001, March 28). The SAT debate ought to be broader. *The New York Times*. Retrieved July 29, 2008, from http://query. nytimes.com/gst/fullpage.html?res=9B0DE5D81F3CF93BA157 50C0A9679C8B63

Sackett, P.R., Kuncel, N.R., Arneson, J.J., Cooper, S.R., & Waters, S.D. (2009). Does socio-economic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, *135*, 1–22.

Sackett, P.R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, *85*, 112–118.

Schmidt, F.L., Oh, I., & Hayes, T.L. (in press). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*.

Sternberg, R.J. (2004). Theory-based university admissions testing for a new millennium. *Educational Psychologist*, *39*, 185–198.

Willingham, W.W. (1985). *Success in college: The role of personal qualities and academic ability*. New York: College Entrance Examination Board.