

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/24001204>

Does Socioeconomic Status Explain the Relationship Between Admissions Tests and Post-Secondary Academic Performance?

Article in *Psychological Bulletin* · February 2009

DOI: 10.1037/a0013978 · Source: PubMed

CITATIONS

198

READS

1,942

5 authors, including:



Nathan Kuncel

University of Minnesota Twin Cities

84 PUBLICATIONS 7,979 CITATIONS

[SEE PROFILE](#)



Justin Arneson

The Myers-Briggs Company

3 PUBLICATIONS 254 CITATIONS

[SEE PROFILE](#)



Shonna Waters

SHRM

11 PUBLICATIONS 458 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Community of curiosity researchers [View project](#)



ASTD Competency Study [View project](#)

Does Socioeconomic Status Explain the Relationship Between Admissions Tests and Post-Secondary Academic Performance?

Paul R. Sackett, Nathan R. Kuncel, Justin J. Arneson, Sara R. Cooper, and Shonna D. Waters
University of Minnesota, Twin Cities Campus

Critics of educational admissions tests assert that tests measure nothing more than socioeconomic status (SES) and that their apparent validity in predicting academic performance is an artifact of SES. The authors examined multiple large data sets containing data on admissions and related tests, SES, and grades showing that (a) SES is related to test scores ($r = .42$ among the population of SAT takers), (b) test scores are predictive of academic performance, and (c) statistically controlling for SES reduces the estimated test–grade correlation from $r = .47$ to $r = .44$. Thus, the vast majority of the test–academic performance relationship was independent of SES: The authors concluded that the test–grade relationship is not an artifact of common influences of SES on both test scores and grades.

Keywords: socioeconomic status, academic performance, admissions testing, test validity

Each year millions of individuals in the United States take post-secondary admissions tests (e.g., SAT [formerly the Scholastic Aptitude Test], the ACT [formerly, American College Testing], the Graduate Record Examination [GRE], the Law School Admission Test [LSAT], the Medical College Admission Test [MCAT], and the Graduate Management Admission Test [GMAT]). Given their prominent role in influencing educational opportunities, these tests are of great interest to the public and undergo considerable scrutiny. A common assertion among test critics is that test scores used for high-stakes decisions (e.g., college admission) measure nothing more than socioeconomic status (SES). Examples of this assertion, drawn from Zwick (2002), include the claim that “in the interest of truth in advertising, the SAT should simply be called a ‘wealth test’” (Guiner, cited in Zwick, 2002), that “the SAT merely measures the size of students’ houses” (Kohn, 2001), and that the “only thing the SAT predicts well now is socioeconomic status” (Colvin, 1997). Implicit in these criticisms is that socioeconomic

status (SES) has an artificial and irrelevant effect on test scores: High SES leads to higher test scores (e.g., through knowledge of test-taking techniques) but not to higher true standing on the characteristic the test is intended to measure (i.e., developed abilities relevant to academic performance). This assertion can be paired with another one, namely, that SES has a similar artificial effect on academic performance measures (e.g., grading is biased in favor of high-SES students) and, thus, that the appearance of test validity (i.e., test–grade correlations) is also an artifact. If SES inflates both test scores and grades of high-SES students and deflates both test scores and grades of low-SES students, then a test that is, in fact, completely invalid as a predictor of academic performance will appear valid as a result of the common effects of SES on both test and grades.

Assertions that the appearance of test validity is an artifact of SES have also been prominently placed within the psychological literature. One claim is that “it has now been documented with massive data sets from the University of California that SAT I scores lose any ability to predict freshman year grades if the regression analyses control for socioeconomic status” (Crosby, Iyer, Clayton, & Downing, 2003). Similarly, “SAT scores used for college admission do not predict freshman year grades when socioeconomic status is controlled” (Biernat, 2003, p. 1023). The most visible critic of the SAT, former president of the University of California system Richard Atkinson (2005), stated that “after controlling for [SES]. . . the relationship between SAT I scores and UC [University of California] grades virtually disappears.” Moving beyond the specific issue of SES and test validity, it is noteworthy that a task force commissioned by APA to examine SES and recommend directions for psychological research and practice has recently issued a report (Saegert et al., 2007). This task force affirmed the criticality of understanding the role of SES.

We concluded that a systematic exploration of the degree to which SES accounts for test–grade relationships was in order. Our goal was to summarize findings from data sets that permit the examination of three relationships: (a) the correlation between SES and scores on cognitively loaded tests, with primary focus on those

Paul R. Sackett, Nathan R. Kuncel, Justin J. Arneson, Sara R. Cooper, and Shonna D. Waters, Department of Psychology, University of Minnesota, Twin Cities Campus.

The order among the latter three authors is alphabetical; all contributed equally to the project.

Justin J. Arneson is now at Target Corporation, Minneapolis, Minnesota, and Shonna D. Waters is now at the Human Resource Research Organization, Alexandria, Virginia.

This research was supported by a grant from The College Board to Paul R. Sackett and Nathan R. Kuncel. Paul R. Sackett serves on The College Board’s SAT Psychometric Panel and the Educational Testing Service’s Visiting Panel on Research. Nathan R. Kuncel serves on the Educational Testing Service’s GRE (Graduate Record Exam) Technical Advisory Committee.

We thank Sarah Hezlett and Jana Rigdon for helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to Paul R. Sackett, Department of Psychology, University of Minnesota, Elliott Hall, 75 East River Road, Minneapolis, MN 55455. E-mail: psackett@umn.edu

of the type used for educational admissions, (b) the correlation between such tests and indices of subsequent academic performance (e.g., grades), and (c) the correlation between SES and these indices of academic performance. With estimates of these three relationships, we can statistically control either test scores or SES to shed light on the nature of the SES–test–academic performance relationships.

Studies using cognitive assessments other than traditional admissions tests (primarily the SAT and the ACT) were included because of considerable evidence that they measure similar constructs. Recently, two separate studies demonstrated strong relationships between both the SAT and the ACT and multiscale measures of general cognitive ability (.82 and .78, respectively; Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008). These correlations are not corrected for unreliability. Although reliability values are not reported for these samples, we would expect reliability of about .90, and correction using this value resulted in correlations of .91 and .87. These results are consistent with Carroll’s (1993) classification of human abilities, in which he placed the SAT–Verbal (SAT-V) and SAT–Mathematics (SAT-M) performance within his taxonomy of human abilities as “verbal ability” and “quantitative reasoning.” He further argued that the “combined total score is probably best regarded as a measure of the second stratum ability 2C (crystallized intelligence)” (Carroll, 1993, p. 705).

The existence of this strong relationship between standardized tests and ability has been consistent over time. Thorndike (1947) provided correlations between a number of different cognitive ability assessments (including IQ tests) and an earlier version of the SAT, with an unreliability-corrected concurrent correlation between the SAT-V and the American Council on Education Verbal measure of .81. Therefore, data in the meta-analysis from different points in time across different types of instruments can be reasonably combined in a single analysis. Given the very strong overlap among these measures and the goal of providing a comprehensive assessment of our research questions, we included studies that evaluated either traditional standardized admissions tests or measures that operationalize the same constructs. As we show below, results were nearly identical across the different types of measures, providing further empirical support for this decision.

We contrasted two conceptual models of the relationships between test scores and grades. Model 1, implicit in the position of the critics noted above, is depicted visually in Figure 1. SES influences test scores, and SES influences grades, but there is no direct relationship between the characteristics measured by the test and grades. Any correlation between test scores and grades is an artifact of the common influences of SES on both test scores and grades. If this model is correct, then the correlation between test scores and grades will drop to zero when statistically controlling

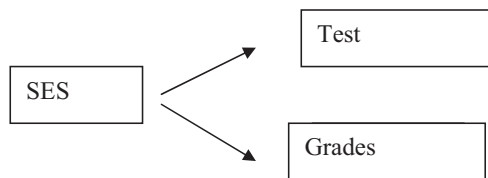


Figure 1. Model 1. SES = socioeconomic status.

for SES. This model is statistically and conceptually consistent with the criticisms discussed earlier. A weaker version of this model would concede the possibility of a weak test–grade relationship after controlling for SES but would nonetheless posit that much or most of the apparent test validity is an artifact of SES. Thus, a comparison of the test–grade correlation with the test–grade correlation controlling for SES can shed light on the feasibility of this model. A finding of a test–grade correlation that changes minimally, if at all, when controlling for SES would be strong evidence against the assertion that the test–grade correlation is an artifact of the joint association of both variables with SES.

Note that the model is a causal one, and the observational data used here did not permit the determination of causality. What we were able to do was to determine whether the observed data are or are not consistent with the model. Large-scale data inconsistent with the model (i.e., a finding that test–grade correlations are minimally affected by statistically controlling SES) would be a major challenge to the assertion that test–grade correlations are an artifact of the common direct effects of SES on test scores and on grades. On the other hand, a finding of data consistent with the model would support the feasibility of the model but would not constitute a definitive test of it.

Figure 2 offers an alternative conceptual model of the relationship between SES, test scores, and grades. Here, SES affects the characteristics measured by tests, which subsequently affect grades. A key feature of this model, however, is that SES is not posited to have a direct relationship with grades; its link to grades is a result of the mediating role of test scores. Whereas SES has an influence on test scores, the test scores truly are predictive of academic performance. The test–grade relationship is not an artifact of the joint influence of SES on both test and grade. The crucial test of this model is a comparison of the SES–grade correlation with the SES–grade correlation after controlling for test score. A finding that the SES–grade correlation is reduced to zero or near zero after controlling for test score, paired with the finding of a substantial test–grade correlation after controlling for SES, would be consistent with this model.

Both of the models articulated above posit test–grade relationships. Model 1 views this relationship as artifactual: Controlling for SES, the test–grade correlation would drop to zero or near zero if this model were correct. Model 2 views the relationship as reflecting a real advantage conferred by high SES: Higher SES leads to higher developed ability, which leads to higher academic performance. Were Model 1 true, continued test use would be inappropriate. Were Model 2 true, then test scores contain meaningful information predictive of academic performance, and the focus would shift to the question of the societal consequences of the fact that being higher in SES confers meaningful advantage. This may lead some to call for interventions to alleviate the advantage conveyed by high SES. It may also lead some to question test use, but it is important to differentiate between criticizing tests on the grounds that they are not valid measures of academically relevant skills and criticizing tests on the grounds that one is not comfortable with the social consequences of using a test, despite its being a valid predictor of academic performance.

To evaluate the credibility of the models, we report three independent investigations. First, we summarize key findings from Sackett, Kuncel, Arneson, Cooper, and Waters’s (2007) analysis of

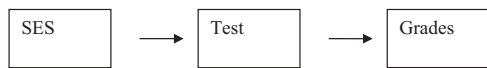


Figure 2. Model 2. SES = socioeconomic status.

a large data set collected by the College Board with the help of 41 colleges and universities, which contains SAT scores (Verbal and Math), SES measures, and freshman grades for over 150,000 students from multiple entering classes at these institutions. Because Sackett et al. also had access to SAT scores and SES measures for a nationwide population of over 2.5 million SAT takers over a comparable 3-year period, it is possible to examine the degree of range restriction that took place on SAT and SES within each of the 41 schools and to perform multivariate psychometric corrections to estimate the correlation of interest in applicant samples rather than in samples selected, in part, on the basis of test scores. Second, we report meta-analyses that we carried out on each of the three relationships of interest (test–SES, test–grade, and SES–grade), which we then used as the basis for computing the partial correlations of interest. These meta-analyses focused on correlations extracted from the psychological and educational literature on the use of both admissions tests and other cognitively loaded tests to predict post-secondary academic performance. The studies reporting relationships with grades were limited to admitted students for whom subsequent grades were available and thus did not permit comparison of applicants and selected students or conducting corrections for restriction of range.

Third, we report findings from a number of existing large-scale publicly available data sets examining tests other than those for college admissions. These include an admissions test for postgraduate study (e.g., law school) and tests of general cognitive ability administered in high school settings. These studies complement the examination of admissions tests in the multi-institution College Board study and in the meta-analytic study. We identified a number of data sets that contain the key variables of test scores, SES, and post-secondary grades. Finally, we present a comparison of findings across these multiple investigations, drawing conclusions about the role of SES in understanding test–grade relationships.

A meta-analysis of relationships between SES and academic achievement was previously conducted by White (1982). There are three important differences between White’s work and ours. First, we included over 2 decades of more recent research published since White’s review. Second, White included students of all ages, including grade school and high school students. Our main focus was on the use of educational admissions tests and, thus, our investigation focused on the prediction of post-secondary school performance. Third, and most critically, White focused on relationships between SES and measures of academic achievement and merged scores on standardized tests with outcome measures such as grades and class rank. Thus, White combined two categories, the separation of which is crucial to our inquiry, namely, test scores and outcome measures. Our focal question was the extent to which the relationship between test scores and educational outcome measures can be explained by SES. Thus, White’s analysis could not answer the questions of interest to us.

We also note that the discussion here is based on the notion of positive effects of high SES. Higher SES does not necessarily have a ubiquitous positive influence on children and their development.

For example, Luthar and Latendresse (2005) and others have demonstrated the presence of elevated substance abuse among children from affluent families, with links to depression and anxiety. Given the links between alcohol dependence and academic failure in college (e.g., Aertgeerts & Buntinx, 2002), the increased prevalence of substance dependence among the children of affluent parents is not unimportant. However, these results should be placed in context. Other variables, including test scores, have shown much larger effects on academic achievement than alcohol abuse (Wood, Sher, Erickson, & DeBord, 1997), and dependence, but not abuse, has been shown to have a deleterious effect on performance. The rate of alcohol-dependent students in the sample was troubling but comparatively small (3.6%). Therefore, nonlinear effects of SES on academic achievement due to substance abuse are not likely to seriously distort SES correlations with grades or tests.

Methodological Issues

Several methodological issues arise as one considers this research domain. The first is the measurement of SES. There is no uniform agreement on measurement of this construct, although most studies have focused on some combination of three measures: parental education, parental income, and parental occupational status. When multiple indicators of SES were obtained, we combined them into an equally weighted composite to create an overall SES measure. The meta-analysis we report here provided the opportunity to determine whether the use of different indicators of SES affects conclusions about the strength of SES–test or SES–grade relationships; findings were consistent across indicators, thus supporting our use of composite measures. Other operationalizations of SES may yield different effects.

A second important methodological issue involves the use of single-institution samples vs. broader samples (e.g., the pooling of data from multiple institutions, or data sets where individuals tested in high school are followed as they attend a wide range of different colleges and universities). On the one hand, analysis of single institution samples or a meta-analysis of single institution studies has the advantage of avoiding confounding level effects. On the other hand, multiple institution and national samples can be affected by level effects but often have the advantage of being based on more representative samples of institutions, including nationally representative samples. One consistent result is that single-institution samples will typically have less variance on SES and on test scores than broader samples. This is due to multiple factors, including self-selection based on test scores (i.e., students have access to information about the test score profiles of admitted students at various universities and target their applications accordingly (Kuncel & Klieger, 2007; Weiler, 1994), institutional selection based on test scores (i.e., test scores are one component of universities’ screening process), and self-selection based on SES (i.e., SES may affect the resources and/or the aspiration for a local vs. national college application process).

Thus, different samples address different research questions. Data on students admitted to a single college or university address the question, Among students admitted to this school, how well do test scores predict subsequent academic perfor-

mance? However, the question of real interest is, Among students who apply to this school, how well do test scores predict subsequent academic performance? Unless schools were to admit students at random in order to permit answering this research question, the best recourse is reliance on psychometric formulas for correction for range restriction. If test variance in the applicant pool and in the admitted class is known, estimates can be made of the correlation between test scores and academic performance in the applicant pool. We were able to include such corrections in our analysis of the College Board data set, as information about both the applicant pool and the admitted class was available.

Data on multi-institutional samples are best viewed as attempts at addressing the question of the relationship between test scores and academic performance among the college-going population in general rather than estimating the relationship within a specific institution. The ideal sample to address this question would be a nationally representative sample of the population of youths applying for college and university entry; some of the national samples we examine fit this category.

In short, single-institution samples address test–performance relationships within the entering cohort of a single school; multi-institutional samples address broader questions about test–performance relationships in the college-going population in general. Both sets of questions are of interest, as is the role of SES in both types of samples.

Third, our investigation was limited to studies at the individual level of analysis (i.e., studies reporting the correlations between individual test scores, SES, and grades). These relationships can also be examined at the aggregate level (e.g., reporting the correlations across schools between mean test scores, mean SES, and mean grade). Individual and aggregate correlations address very different questions (Do students with higher SES have higher test scores? vs. Do schools with higher mean SES have higher mean test scores?), and there is no necessary relationship between the two. White (1982) reported much higher correlations at the aggregate level; we found the same. In the College Board data, for example, the mean SES–SAT correlation at the individual level was .22; the correlation between SES means and SAT means across the 41 schools was .49. Because the mean level correlations treat all individuals in a particular school or SES category as being identical (which they are not), and we are most interested in how SES and ability influences individuals, we examine only individual-level correlations.

Fourth, our analyses all used cross-sectional estimates of SES. Longitudinal research has demonstrated that chronic or persistent poverty has larger effects on the development of children than does transitory poverty (for a review, see McLoyd, 1998). Environmental toxins, less effective parenting, chronic stressors, and less cognitive stimulation all appear to be associated with poor academic achievement. Within our data, low income was therefore a mixture of students who had chronically experienced low incomes combined with those whose family income had only recently decreased to a low level. Those students who have experienced chronically low family income were likely to have lower test scores and poorer academic outcomes than those whose families were experiencing transitory poverty. None of the studies in our meta-analyses

of SES relationships in the context of college admission differentiated between chronic and transitory poverty.

Investigation 1: College Board Data from 41 Colleges and Universities

Method

We first present key findings from Sackett et al.'s (2007) analysis of the multi-institution data provided by the College Board because of the large sample size and the fact that it permits addressing the range restriction issue in multiple ways. We rely on the other data sets to corroborate the multi-institution data findings using a variety of different tests and test settings.

Sample

The College Board collected SAT, SES, and freshman grade information from three entering cohorts (1995, 1996, and 1997) in collaboration with a group of 41 colleges and universities. These were selected to be geographically diverse, to include large and small schools, to include public and private institutions, and to cover a broad range in terms of school selectivity on SAT scores. Twenty-eight schools provided data for all three entering cohorts, eight provided data for the 1995 and 1996 cohorts only, and five provided data for the 1995 cohort only. All schools provided freshman grades, whereas a smaller subset of schools provided cumulative grades for at least 4 years; analyses reported here focused on freshman grades to maximize the number of participating institutions (i.e., to include all 41 schools). For prior research using this data set, see Bridgeman, Pollack, and Burton (2004). Table 1 presents descriptive information about each school, including mean entering class size, public–private status, mean SAT Mathematics plus Verbal scores, and mean SES for entering students at each school. Schools are not identified by name to preserve anonymity. The total sample size was 167,816; all three key variables (SAT, SES, and grade point average [GPA]) were available for 155,191 students, and analyses focused on those students with complete data. Table 1 also includes SES–SAT, SES–GPA, and SAT–GPA correlations for each school.

Measures

SAT-M and SAT-V were obtained from College Board records and summed to form a composite. Three SES variables were obtained from questionnaires completed by students at the time they took the SAT: father's years of education, mother's years of education, and family income. The mean school-specific correlation between the two education variables was .57; father's and mother's education had mean correlations of .43 and .35, respectively, with family income. As detailed below, these three SES variables were available for the national population of SAT takers who reported these data on a questionnaire at the time they applied to take the SAT. Using College Board data about means, standard deviations, and intercorrelations among these variables at the national SAT-taking population level, we created an equally weighted composite of these three variables by standardizing each SES variable, summing the three, and restandardizing the resulting sum. This created an SES variable with a mean of zero

Table 1
Characteristics of the 41 Schools in the Multi-Institution Database

School	Entering class <i>N</i>	Public (0)/private (1)	SAT mean	SES mean	SES–SAT <i>r</i>	SES–GPA <i>r</i>	SAT–GPA <i>r</i>
1	1,911	0	1082.97	0.06	.26	.05	.20
2	1,271	0	951.11	–0.16	.14	.04	.28
3	3,821 ^a	1	1134.71	0.33	.14	.05	.44
4	2,333	0	1090.71	–0.06	.27	.08	.24
5	1,160	1	1250.85	0.61	.20	.10	.23
6	1,007 ^a	1	1294.22	0.81	.34	.21	.35
7	554	1	1101.17	0.51	.21	.16	.49
8	829	1	1084.56	–0.02	.09	.09	.26
9	763 ^b	0	939.54	–0.26	.17	.06	.27
10	961	0	960.61	–0.18	.22	.07	.26
11	372 ^a	1	967.98	–0.07	.05	.05	.48
12	297	1	1058.05	0.07	.18	.18	.47
13	2,007	0	1124.35	0.47	.16	.13	.32
14	371	1	1204.14	0.75	.20	.07	.41
15	1,296	0	948.77	–0.23	.16	.09	.30
16	395	0	1154.45	0.58	.13	.04	.40
17	334	1	1263.91	0.73	.14	.09	.30
18	1,029	0	1024.77	–0.14	.11	.08	.31
19	3,444	0	1190.93	0.74	.27	.20	.34
20	1,011	0	974.27	–0.18	.20	.13	.38
21	1,624 ^a	0	1031.40	0.09	.22	.12	.42
22	453	0	1044.30	0.26	.11	.08	.37
23	652	1	1220.57	0.77	.34	.26	.46
24	1,467	0	971.05	0.00	.17	.10	.30
25	1,146	0	976.36	–0.17	.18	.08	.30
26	1,210	0	941.07	–0.34	.14	.10	.31
27	421 ^a	1	1044.39	0.17	.18	.06	.46
28	365 ^b	1	1070.87	0.44	.05	.02	.47
29	195	1	994.79	–0.10	.20	.07	.41
30	1,604 ^a	0	1151.05	0.19	.42	.13	.34
31	5,544 ^a	0	1060.95	0.32	.21	.17	.41
32	2,836 ^b	0	1004.11	0.18	.18	.12	.38
33	1,152 ^b	1	1190.98	0.57	.08	.06	.37
34	2,031	0	984.17	0.01	.10	.07	.20
35	3,498 ^b	0	1084.44	0.06	.42	.25	.41
36	3,215	0	1107.60	0.46	.35	.26	.46
37	1,898	0	1074.22	0.33	.19	.09	.36
38	6,172	0	1084.36	0.41	.27	.14	.40
39	1,507	1	1114.71	0.51	.12	.03	.42
40	3,529	0	1097.51	0.28	.29	.16	.39
41	1,136 ^a	0	944.20	–0.10	.19	.08	.31

Note. Schools unmarked with an a or b superscript provided data for 1995, 1996, and 1997 cohorts. SES = socioeconomic status; GPA = grade point average. From “Socio-Economic status and the Relationship Between the SAT and Freshman GPA: An Analysis of Data From 41 Colleges and Universities” (Tech. Rep. No. 2007-5), by P. R. Sackett, N. R. Kuncel, J. J. Arneson, S. R. Cooper, and S. D. Waters, 2007. Copyright 2007 by The College Board. Adapted with permission. All rights reserved.

^a School provided data for 1995 cohort only. ^b School provided data for 1995 and 1996 cohorts. No asterisk indicates school provided data for 1995, 1996, and 1997 cohorts.

and a standard deviation of one in the national population. Individual SES scores were then computed using this metric, thus permitting comparisons of each school’s SES with the national test-taking population. Freshman GPA was provided by the college or university. High school GPA was obtained from the student questionnaires and used in multivariate corrections for restriction of range.

Analyses

Differences by entering cohort (1995–1997) were examined. Characteristics of entering classes were very stable from year to year within school. The correlation across schools of SAT means averaged .99 across the three possible comparisons (1995–1996,

1996–1997, and 1995–1997); the correlation across schools of SES means averaged .98. On the basis of these findings, data across the three cohorts for each school were pooled. All analyses were then conducted separately by school.

Meta-analyses of the SES–SAT, SES–GPA, and SAT–GPA correlations were conducted using the Hunter and Schmidt (2004) method. This approach involves (a) computing the sample-size-weighted mean and variance of the correlations, (b) computing sampling error variance, (c) subtracting sampling error variance from observed variance to estimated variance in correlations net of sampling error, commonly referred to as population variance, (d) computing the 95% confidence interval around the mean correlation, and (e) computing a 90% credibility interval, which estimates

the range within which 90% of correlations net of sampling error are expected to fall. The Hunter–Schmidt approach does not include a statistical test for homogeneity of correlations; we computed the Hedges and Olkin’s (1985) Q statistic to test the hypothesis that population variance is zero according to a fixed model for error. The accuracy of the Hunter–Schmidt approach has been upheld in multiple simulation studies (e.g., Burke, Raju, & Pearlman, 1986).

Sackett et al. (2007) also sought to obtain applicant population data in order to estimate correlations among test scores, grades, and SES independent of the effects of range restriction. *Range restriction* refers to the fact that variance is reduced when the sample available for study has been selected in part on the basis of scores on the variable in question (e.g., computing SAT–GPA correlations in samples where SAT scores were part of the selection process) or on the basis of a variable correlated with the variable of interest (e.g., computing SAT–GPA correlations in samples where high school grades, which are correlated with SAT scores, are used in the selection process). Restricted variance on the test results in a lower test–grade correlation than would be the case if the relationship were examined in applicant samples. When multiple variables are used in selecting applicants (e.g., both SAT scores and high school grades), multivariate range restriction corrections ideally should be used (Sackett & Yang, 2000).

In the multivariate range restriction scenario, there is a set of variables for which the unrestricted means, standard deviations, and correlations among the variables are known and another set of variables where only restricted means, standard deviations, and correlations are known. In this setting, as detailed below, unrestricted data on three variables known prior to college entry (SAT, SES, high school GPA) are available. However, because college grades are only available for those who are selected and then enroll, only restricted means, standard deviations, and correlations with SAT, SES, and high school GPA are known for the college grade variable, and range restriction correction is used to estimate the unrestricted means, standard deviations, and correlations. The goal of range restriction corrections is to obtain a better estimate of the unrestricted population correlation. The optimal correction would include all variables that affect application and admissions decisions, a condition never met in applied settings. Thus, whereas admissions test scores and high school performance are major determinants of admissions decisions, they are not the sole determinants.

Sackett et al. (2007) obtained two separate sources of information regarding unrestricted means, standard deviations, and correlations. The first consisted of means, standard deviations, and correlations between SAT, SES, and high school GPA among the entire population of individuals taking the SAT and completing a questionnaire reporting SES in 1995, 1996, and 1997 (over 2.5 million students). Thus, these data describe the population for whom the test is relevant. Table 2 presents these unrestricted means, standard deviations, and correlations.

The second goal was to obtain estimates of the means, standard deviations, and correlations in the applicant pool for each specific college or university in order to obtain unrestricted estimates of the correlation of interest among each school’s applicant population. Whereas such data were not directly available, Sackett et al. (2007) obtained data that provided a reasonable proxy to the school-specific applicant pool. When students take the SAT, they indicate

Table 2
Means, Standard Deviations, and Correlations Among SAT, HS GPA, and SES for National SAT Population

Measure	HS GPA	SAT	SES
SAT	0.54		
SES	0.20	0.42	
<i>M</i>	3.21	1012.77	0
<i>SD</i>	0.66	206.47	1

Note. HS GPA = High school grade point average; SES = socioeconomic status. From “Socio-Economic Status and the Relationship Between the SAT and Freshman GPA: An Analysis of Data From 41 Colleges and Universities” (Tech. Rep. No. 2007-5) by P. R. Sackett, N. R. Kuncel, J. J. Arneson, S. R. Cooper, and S. D. Waters, 2007. Copyright 2007 by The College Board. Adapted with permission. All rights reserved.

the schools to which they wish their scores to be sent; the set of students who asked that their scores be sent to a given school was used as the estimate of the applicant pool for that school. Thus multivariate range restriction corrections were made with the use of both the school-specific estimates of the applicant pool and the entire SAT-taking population as the referent population (Sackett & Yang, 2000). The results of both analyses are presented for comparison purposes, although we believe that the school-specific corrections provide the most appropriate results in that they more closely approximate the setting in which scores are actually used (i.e., to select among applicants to a given school).

Results: Multi-Institution Sample, 1995–1997

Table 3 presents the results of the meta-analyses of SES–SAT, SES–GPA, and SAT–GPA correlations. Table 4 presents the mean correlations and partial correlations of interest, with and without corrections for restriction of range. Table 4 is a summary table that also includes findings from investigations detailed in later sections of this article.

SES–SAT Relationships

Answers to the question, “How strongly are test scores and SES related?” vary as a result of the type of data examined. The observed sample-size weighted mean SES–SAT correlation among students enrolled at a given college or university was .22. Using school-specific range restriction corrections resulted in an estimated mean correlation of .31 among applicants to a specific school; using national SAT population range restriction correction resulted in an estimated mean correlation of .42 in the entire SAT-taking population. The difference between the correlation estimate of .42 in the entire test-taking population and .31 in the population of applicants to a given school reflects self-selection on either or both variables: Both SES and knowledge of typical SAT scores of admitted students may affect student application decisions (e.g., Kuncel & Klieger, 2007). The difference between the corrected correlation of .31 in the school-specific applicant pool and the observed correlation of .22 among enrolled students reflects a combination of these self-selection factors and the school’s use of the SAT scores as a factor in admissions decisions. Thus, correlations computed from samples of enrolled students underestimated SES–test relationships in the college-bound population,

Table 3
Meta-Analysis of Multi-Institution College Board Data

Variables	<i>K</i>	<i>N</i>	Mean <i>r</i>	<i>SD_r</i>	<i>SD_e</i>	<i>SD_p</i>	<i>Q</i>	95% Confidence interval	90% Credibility interval
SES–SAT	41	155,191	.221	.081	.018	.079	830.25*	.196–.246	.091–.351
SES–GPA	41	155,191	.124	.061	.02	.058	381.40*	.106–.142	.028–.219
SAT–GPA	41	155,191	.354	.075	.012	.074	1,601.56*	.331–.377	.232–.476

Note. *K* = number of samples; *N* = total sample size; *SD_r* = observed standard deviation of correlations; *SD_e* = standard deviation expected due to sampling error; *SD_p* = residual standard deviation.

* *Q* test for homogeneity, significant at *p* = .05.

leading to the conclusion that the population of interest must be specified when one estimates the correlation between SES and test scores.

SES–Grade Relationships

The sample-size weighted mean within-school SES–grade correlation was .12. After correcting for range restriction, we found a mean correlation of .19 for applicants to a specific school and a mean correlation of .22 for the full population of SAT test-takers. Thus, institutional or self-selection on SES or on correlates of SES (e.g., test scores) reduced the SES–grade correlation in enrolled student samples. In short, SES was correlated with grades, though the correlation was relatively low and was lower than the correlation between SES and test performance.

SAT–Grade Relationships

The sample-size weighted mean within-school SAT–grade correlation was .35. After correcting for range restriction, we found

mean correlations of .47 for applicants to a specific school and .53 for the full population of SAT test-takers. Institutional or self-selection on the SAT or on correlates of the SAT (e.g., high school GPA, SES) reduced the SAT–grade correlation in enrolled student samples. Thus, the SAT–grade relationship varied as a result of decisions about whether and how to correct for range restriction. We posit that correcting for school-specific applicant pools gives the best estimate of the relationship of operational interest, namely, how well the SAT predicts grades given the set of applicants who present themselves for consideration at a given school. We note that school-specific applicant pool information is often not available, and it is not uncommon to use the SAT-taking population as the reference group in making range restriction corrections. This answers a hypothetical question (i.e., What would the validity of the SAT be if the applicant pool for a given school were a random sample of the SAT-taking population?) rather than the operational question of the validity of the SAT for existing applicant pools. An argument for using broader applicant pools as the basis for correction is that some students decide not to apply to a given school

Table 4
Summary of SES–Test, SES–Grade, and Test–Grade Relationships Across Studies

Sample	No. studies	Sample size	<i>r</i> _{SES–test}	<i>r</i> _{SES–grade}	<i>r</i> _{test–grade}	<i>r</i> _{test–grade, controlling for SES}	<i>r</i> _{SES–grade, controlling for test}
Investigation 1							
Meta-analysis of College Board data	41	155,191					
Observed <i>r</i> s			.22	.12	.35	.33	.04
<i>r</i> s corrected for school-specific range restriction			.31	.19	.47	.44	.05
<i>r</i> s corrected for national population range restriction			.42	.22	.53	.50	–.01
Investigation 2							
Meta-analysis							
Studies using admissions tests, SES composites, and including SES, test, and grade	17	17,235	.15	.09	.37	.36	.03
All studies	55	60,565	.25				
	65	41,829		.09			.00
	37	26,127			.35	.34	
Individual studies							
1995 National Study of Law School Performance	21	3,375	.16	.07	.38	.38	.01
Harvard Study of the Class of 1964–1965		486	.07	.05	.30	.29	.03
LSAC National Longitudinal Bar Passage Study		19,264	.13	.05	.35	.35	.01
NELS88		6,314	.40	.10	.24	.23	.02
NLS-1972		5,735	.30	.04	.31	.31	–.01
Project Talent		749	.18	.05	.30	.29	.01

Note. SES = socioeconomic status; LSAC = Law School Admission Council; NELS88 = National Educational Longitudinal Study of 1988; NLS-1972 = National Longitudinal Study of the High School Class of 1972.

on the basis of knowledge of their own test scores and of the typical scores of students enrolling at a given school. Thus test score variance at a given school is restricted as a result of both the school's selection processes and individual students' self-selection processes. Thus, whereas both estimates are of interest, we focused on current operational validity (i.e., how well the test predicts among those currently applying) while acknowledging that test scores can also play a role in student self-selection.

SAT–Grade Correlations, Controlling for SES

To test the proposition that the SAT–grade relationship was an artifact of the relationships between SES and both test scores and grades, we computed the test–grade correlation partialing out SES to determine the degree to which controlling for SES reduced the SAT–grade relationship. Contrary to the assertion of test critics, observed SAT–grade correlations were, at most, nominally affected when controlling for SES. We view the SAT–grade correlation corrected for school-specific range restriction ($r = .47$) as the best estimate of operational test validity for predicting grades within a given school's applicant pool. This value drops to .44 when controlling for SES. Thus, contrary to the claim that the relationship drops to near zero when controlling SES, our conclusions are that the SAT retains virtually all of its predictive power when SES is controlled and that SAT validity is not an artifact of SES.

SES–Grade Correlations, Controlling for SAT

To examine whether data were consistent with the Model 2 proposition that the observed correlation between SES and grades was mediated by test performance (i.e., that SES did not influence grades other than through its relationship with test performance), we partialled SAT performance from the SES–grade relationship. Consistent with this proposition, SES–grade correlations did drop substantially when controlling for SAT. The SES–grade correlation, corrected for school-specific range restriction of .19, dropped to .05 when controlling for SAT score. This is consistent with the Model 2 position that the relationship between SES and grades is largely mediated by test score.

Investigation 2: Meta-Analysis of SES–Test, SES–Grade, and Test–Grade Relationships

Method

Samples

Articles were identified for potential inclusion in the meta-analysis by conducting electronic searches of computerized databases with the keywords *mother's income*, *socioeconomic status*, *father's income*, *family background*, *mother's education*, *SES*, *father's education*, *parental education*, *mother's occupation*, *father's occupation*, and *parental income* in the Education Resources Information Center (ERIC) database (1966–2004), PsycINFO (1887–2004), and the Dissertation Abstracts online database. Screening requirements were that articles must contain the relevant variables with codable information (zero-order correlations or a statistic that could be converted (e.g., F , t , χ^2), have college applicant or late adolescent–young adult samples, and be written in

English (Cohen, 1988; Hunter & Schmidt, 2004). Studies were coded as to whether they included admissions tests or other cognitively loaded tests. Studies were coded as to whether they contained all three relationships (test–grade, test–SES, and grade–SES) or whether they contained one of the SES relationships (test–SES or grade–SES). Studies containing only the test–grade relationship were not coded because of the availability of existing large-scale meta-analyses of test–grade relationships for educational admissions tests (e.g., Hezlett et al., 2001). Articles were coded by two of the authors. There were very few disagreements, and consensus was reached in discussion. This process resulted in 66 studies, containing 55 independent samples for the SES–test relationship, 65 independent samples for the SES–grade relationship, and 37 independent samples for the test–grade relationship. We excluded a large study (Young & Fisler, 2000; $N = 69,284$), as it examined a considerably smaller sample from the same source that we examined in the multi-institution College Board study described in the earlier section of this article (e.g., SES–SAT correlations were based on students' reporting of SES on a questionnaire administered in conjunction with the SAT). Table 5 contains descriptive information about each study, including sample and measures used, as well as the SES–test, SES–grade, and test–grade correlations for each study.

Variables

SES. The SES measures used in the meta-analysis varied by study, typically involving parental education, earnings, and/or occupational status. We examined the relationships between each of the SES indicators and the criteria of interest. The relationships varied by the SES indicator used, with mean SES–test correlations ranging from .19 to .28 and mean SES–grade correlations ranging from .03 to .09. Separate meta-analyses were conducted for studies using each SES measure. The relationships among SES indicators were also examined for those studies that provided intercorrelations among SES indicators. The mean of 37 reported intercorrelations among individual measures was .43.

Our main analysis, however, focused on studies using composites of SES measures. When multiple indicators for a variable were provided, all were coded, and a composite was created and used in the meta-analysis. For example, when father's education, mother's education, and parental income were all provided as indicators for a single study, we calculated a composite. When text reported that multiple indicators were collected, but only an overall summary index was used for analysis, that summary value was used. We present separate meta-analyses for studies using SES composite measures and studies using single indicators and focus our substantive conclusions on studies using composite measures.

Test. Forty-two percent of the studies used the SAT, 25% used the ACT, 5% used the Preliminary SAT (PSAT), and 29% used other examinations. A composite was computed if multiple test scores were reported (e.g., SAT-M and SAT-V).

Grades. A college GPA measure was obtained from each study. Among the studies, 58% used cumulative grades, 26% used

(text continues on page 13)

Table 5
Descriptive Information About Studies Included in the SES-Test, SES-Grade, and Test-Grade Meta-Analyses

Authors	Sample no.	Study participants	N	Test	SES measures	Grade	SES-test <i>r</i>	SES-grade <i>r</i>	Test-grade <i>r</i>
Alao, Coffey, Ellington, & Wright (1999)		Freshmen Black students at predominantly White colleges	162	SAT	PI		.17		
Allen (1981)		Black students at predominantly White colleges	135		ME, FE	GPA		.07	
Allen (1985)		Black students at predominantly White colleges	327		SES (unspecified)	GPA		.12	
Allen (1992)		Black college students	1,800		ME, FE, MO, FO, PI	GPA		.04	
Allen & Haniff (1991)		Black students at predominantly White colleges	1,462		PI, ME	GPA		.07	
Alnabhan, Al-Zegoul, & Harwell (2001)		Education students	600	Tawjhi (Jordanian national standardized test)	ME, FE	GPA	.06	.02	.23
Anthony, Sanders, & Kahle (1981)		Black freshmen and sophomores	121	Composite of School and College Ability Test and Otis-Lennon Ability Test	PE		.36		
Bachman & O'Malley (1986)		10th graders	1,487	Composite of multiple ability tests	FO, FE, ME, PI		.38		
Baker (1998)		Black males at predominantly White college	69		PE, PO, PI	GPA		.17	
Best (1968)		Nursing students	128		FO, FE, ME, MO	Freshman GPA		.00	
Braddock (1981)	1	Black college students	113		PI	GPA		-.04	
Brigham (1981)	2	Black college students	140		PI	GPA		-.08	
		Scholarship male student athletes	90	ACT	PI	GPA Composite of mathematics and English GPA	.30	.19	.48
Brodnick & Ree (1995)		Traditional students	339	SAT, ACT	PI		.26	.07	.28
Burnham & Hewitt (1967)	1	Yale University students, class of 1931/1932	1,393	SAT	FO, ME	Senior year GPA	.09	.02	.26
	2	Yale University students, class of 1935/1937	1,285	SAT	FO, ME	Senior year GPA	.10	.01	.29
Burton (1976)	3	Yale University students, class of 1959	899	SAT	FO, ME	Senior year GPA	.01	.09	.18
Calkins, Willoughby, & Arnold (1982)		College seniors	229		PI, PE, ME			-.06	
	1	Minority freshmen	46		FE, ME	GPA		.27	
	2	White freshmen	373		FE, ME	GPA		.06	
Carlson (1967)		Nursing students	81	School and College Ability Test	FE, ME	GPA	.07	-.24	.18
Coates (1996)		Student athletes	108	SAT	FE, ME, SES	GPA	.46	.34	.59
Crumb (1986)		Second semester juniors	94	ACT	PI	GPA	.21	.25	.24
Donovan (1983)		Predominantly minority college students	337		PE, PI	Freshman GPA		.01	
Just (1995)		Freshmen	202	SAT	FE, ME, PI		.30		

(table continues)

Table 5 (continued)

Authors	Sample no.	Study participants	N	Test	SES measures	Grade	SES-test r	SES-grade r	Test-grade r
Epps & Jackson (1985)	1	High school senior males	167	Composite of reading, vocabulary, and mathematics tests	PI, FO, MO, FE, ME		.04		
	2	High school senior females	319	Composite of reading, vocabulary, and mathematics tests	PI, FO, MO, FE, ME		.02		
	3	High school senior males	259	Composite of reading, vocabulary, and mathematics tests	PI, FO, MO, FE, ME		.47		
	4	High school senior females	319	Composite of reading, vocabulary, and mathematics tests	PI, FO, MO, FE, ME		.28		
Falb (1995)		Students transferring from community college to 4-year college	149		PI	1st semester GPA		-.14	
Gansmer & Bealer (1977)		High school sophomores	460	Various IQ tests (different schools used different tests)	FE, ME, FO		.18		
Gillie (1972)		Community college females	115		FE, FO	Freshman GPA		.19	
Glenn (1970)		Minority and economically disadvantaged students	149		ME, FE	Freshman GPA		.04	
Greever, Tseng, & Udelle (1973)		Freshmen	188	ACT	Duncan's socioeconomic index and equivalent NORC prestige score	GPA	.15	.03	.39
Guerrero (2000)		Freshmen	1,154	SAT	SES, self-report: poor, working class, middle, upper middle, upper	GPA	.20	.10	.24
Harrison (1981)		Students in a special studies program	240	SAT	PI	GPA	.35	.07	-.26
Hearn (1981)		Freshmen	5,211	SAT	FE, ME, PI		.21		
Hill & Woederhoff (1966)		Students enrolled in beginning German	192		PI	Course grade		-.10	
Horn (2004)		Graduates	743	ACT	PI	Major GPA	.07	.10	.31
Huffman, Sill, & Brokenleg (1986)	1	Native American college students	38		ME, FE, PI	GPA		.03	
	2	White college students	48		ME, FE, PI	GPA		.06	
Ikeda, Wolfe, & Rich (1971)	1	Black students	95	SAT	FO	GPA	.08	.14	.08
		White students eligible for federally aided Educational Opportunity Grants							
	2	White students not eligible for Educational Opportunity Grants	125	SAT	FO	GPA	.14	.39	.23
	3	White students not eligible for Educational Opportunity Grants	138	SAT	FO	GPA	.04	-.13	.27

(table continues)

Table 5 (continued)

Authors	Sample no.	Study participants	N	Test	SES measures	Grade	SES-test r	SES-grade r	Test-grade r
Keiller (1997)		Introductory psychology students	592	ACT	PI, PE	GPA	.16	.02	.28
Kinsella (1995)		Disadvantaged college students	46	SES	Freshman GPA		.04		
Larkin (1981)		Freshmen	100	SAT	PI	Freshman GPA	-.13	-.06	.32
Lee (1986)		College sophomores	148		ME, FE, PI	1st semester GPA		.10	
Lovato (1981)		Students in second semester of enrollment	377	SAT	PO, PE	1st semester GPA	.31	.04	.20
Lowman & Spuck (1975)		Disadvantaged Mexican American students, males	43	SAT	PI	Freshman GPA	.06	.06	.08
Lowman & Spuck (1975)		Disadvantaged Mexican American students, females	32	SAT	PI	Freshman GPA	-.08	.40	.33
McQuary (1951)		Freshmen who sought services of student counseling center	174	Composite—Quantitative Comprehension, Vocabulary	FO, FE, ME	GPA	.48	-.04	.62
Mittal (1989)		Students in teachers education program	406	ACT	PI, FE, ME	GPA	.06	-.06	.37
Moline (1987)		Freshmen	227	PSAT	PI	GPA	.23	.22	.50
Nettles, Thoeny, & Gosman (1986)		Sample of 50% White and 50% Black college students	4,040		PI, PO, PE	GPA		.12	.08
Nonnamaker (2000)		Freshmen	261	SAT	PE, PI	GPA	.14	-.04	
Nottingham (1990)		Freshmen in general biology course	196	SAT	PO, PE	Freshman GPA	.21	.12	.45
Noxel (2000)		Baccalaureate recipients	7,472	ACT	PI, FE, ME	GPA	.19	.14	.49
Oliver, Rodriguez, & Mickelson (1985)	1	Latino college students	63		Social class, FE, ME	GPA	.14		
	2	Black college students	84		Social class, FE, ME	GPA	.11		
Rea (1992)	1	Freshmen	1,280	SAT	PE, PI	Freshman GPA	.02	.10	.19
Reed (1944)		Students in college sociology class	200	Michigan Vocabulary Test	ME, FE		.04		
Reitzes & Mutran (1980)		College sophomores, juniors, and seniors	396		FE, PI	GPA		.12	
Rimmer (1973)		Freshmen	153	SAT	FO, FE, MO, ME		.02		
Rodgers (1984)		Freshmen	424	ACT	PI	Freshman GPA	.14	.12	.39
Rodriguez (1992)	1	Anglo-American freshmen	83	ACT	PI, FE, ME	Freshman GPA	.15	.18	.33
	2	Mexican American freshmen	225	ACT	PI, FE, ME	Freshman GPA	.17	.08	.36
Scarr (1979)	1	Adopted children ages 16–22 years	150	WAIS IQ	PI, FO, FE, ME		.10		
Smith (2002)	2	Biological children ages 16–22 years	237	WAIS IQ	PI, FO, FE, ME		.26		
		College students	153		FE, ME, FI	GPA		.29	
Spector (1966)		Junior college freshmen	777		Class rating (lower class to upper class)	1st semester GPA		.04	

(table continues)

Table 5 (*continued*)

Authors	Sample no.	Study participants	N	Test	SES measures	Grade	SES-test r	SES-grade r	Test-grade r
Suarez (1981)		American Indian students	72	ACT	FE, ME, PI		.22		
Terenzini, Pascarella, & Lorang (1982)		Freshmen High school male seniors from National Longitudinal Study of the Class of 1972	469		PE, PI	Freshman GPA		.11	
Thomas (1977)	1	High school female seniors from the National Longitudinal Study of the Class of 1972	5,104	Composite of math, verbal, letter groups, and reading tests	FE, FO, ME		.36		
	2	Students at predominantly Black colleges	5,416	Composite of math, verbal, letter groups, and reading tests	FE, FO, ME		.36		
Tilford & Allen (1974)		Graduates	470	ACT	PE, PO, PI	GPA	.14	.04	.41
Tollett (1980)		High school seniors from High School and Beyond study	111	ACT	FE, ME, PI, SES	Freshman GPA	.16	.05	.69
Tuttle (1981)		High school seniors from National Longitudinal Study of the Class of 1972	14,687	Composite of 5 ability tests	FE, ME, PI		.32		
Velez (1985)		Freshmen, males	3,169		FE, ME, PI, FO, Household Items	GPA		.13	
Werts (1967)	1	Freshmen, females	2,000	PSAT	FE	GPA	.20	-.05	.39
	2	Middle ranking among 4-year liberal arts colleges	2,000	PSAT	FE	GPA	.25	.06	.37
Willard (1993)		Freshmen	443	SAT	FE, ME, PI	Freshman GPA	.15	.12	.47
Wright & Bean (1974)		Freshmen	1,631	SAT	PI, ME, FO	Freshman GPA	.15	.08	.24

Note. SES = socioeconomic status; SAT refers to a composite of the Mathematics and Verbal subtests; ACT = (formerly American College Testing Program); PSAT = Preliminary SAT; NAIS = Wechsler Adult Intelligence Scale; FE = father's education; ME = mother's education; PI = father's occupation; FO = mother's occupation; MO = mother's occupation; PO = parental occupation; FI = father's income; MI = mother's income; PI = parental income; NORC = National Opinion Research Center; GPA = grade point average.

freshman grades, 6% used first-semester grades, and 9% used another grade measure (e.g., first and second year, major).

Analyses

As in the analyses of the multi-institution College Board sample reported in the prior section of this article, we used the Hunter-Schmidt meta-analytic method (Hunter & Schmidt, 2004). Individual study corrections for measurement error and range restriction were not possible because of the limited number of studies providing such information and the fact that too few studies reported the needed information for an artifact distribution approach to be feasible.

The strongest data for answering the questions of interest come from studies that used admissions tests, that included all three key variables, and that used a composite SES measure. We thus focused initially on studies meeting these three criteria and present these as our central findings. We focused on studies including all three key variables in response to concerns about using a meta-analytically derived matrix for multivariate analyses (Becker & Schram, 1994; Shadish, 1996). Assembling a matrix based on studies that do not contribute information for all of the cells can create inaccurate estimates, particularly if the data come from samples from different populations. Here, all studies examine the population of college students. In addition, the measurement of constructs needs to be consistent across studies to produce a meaningful analysis. Construct measurement is probably not an issue for the current study because ability, SES, and grades were operationalized in similar ways across studies, particularly in our central analysis focusing on admissions tests and composite SES measures. Finally, the presence of variability in study effects within each cell of the study due to artifactual (e.g., range restriction) or substantive sources of variability can also lead to misleading estimates. We cannot fully rule out this issue because not all of the variability across studies was attributable to sampling error. However, as is shown below, standard deviation estimates of the true variability of correlations were smaller (between .04 and .12) than what are typically observed in meta-analyses on individual difference variables, suggesting that effects are comparatively stable across samples and situations. We later demonstrate that substantive conclusions based on studies with all three variables were the same as those in which differing sets of studies were used to estimate relations among SES, tests, and grades.

We supplement these focal analyses with more detailed analyses of the SES–test, SES–grade, and test–grade relationships in which

we examine several potential moderator variables. For each relationship of interest (SES–test, SES–grade, test–grade), meta-analyses are reported for the entire sample. Separate analyses are then reported for studies using admissions tests versus other tests, for studies for which all three key variables (SES, test, grade) were available versus studies for which only two of the variables (SES and test or SES and grade) were available, and for studies using an SES composite versus a single SES indicator. The *Q* statistic (Hedges & Olkin, 1985) was used to test the significance of the difference between the mean correlations in each of these pairs of conditions. As the *Q* statistic is appropriate for independent partitions of the data, it was not applied to comparisons between SES–test correlations for individual SES indicators (e.g. mother’s education, family income), as the comparison of individual indicator correlations was based on a mixture of effect size measures drawn from studies using multiple SES indicators and studies using single indicators.

Results

Studies Conducted With Admissions Tests and Composite SES Measures and Reporting SES–Test, SES–GPA, and Test–Grade Relationships

Table 6 summarizes the results of meta-analyses of SES–test, SES–GPA, and test–GPA relationships among 17 studies, which included over 17,000 students and which met the key criteria outlined above (i.e., use of an admissions test, use of a composite SES measure rather than a single indicator, and inclusion of all three key variables: SES, test, and GPA). SES–test correlations averaged .15 with an estimated population standard deviation of .05. SES–GPA correlations averaged .09, with a population standard deviation of .04. Test–GPA correlations averaged .37, with a standard deviation of .13. The *Q* test for homogeneity is significant for all three of these relationships, which is consistent with the finding of nonzero standard deviation estimates.

Partial correlations were computed to determine the validity of tests for predicting college grades when controlling for SES. After controlling for SES, the test–grade mean correlation of .37 was reduced by .01 to .36. Thus, controlling for SES had a negligible effect on the test–grade relationship in this analysis. Partial correlations were also computed to examine the SES–grade correlation controlling for test. After controlling for test, the SES–grade mean correlation of .09 was reduced to .03. Thus the SES–grade relationship dropped to near zero when controlling for test score.

Table 6
Meta-Analysis of Studies Using Admissions Tests, Composite Measures of SES, and Including All Three Key Variables (SES, Test, and Grade)

Variables	<i>K</i>	<i>N</i>	Mean <i>r</i>	<i>SDr</i>	<i>SD_e</i>	<i>SD_p</i>	<i>Q</i>	95% Confidence interval	90% Credibility interval
SES–Test	17	17,235	.154	.061	.037	.048	46.21*	.127–.181	.074–.234
SES–GPA	17	17,630	.092	.059	.040	.043	36.99*	.067–.117	.021–.163
Test–GPA	17	17,244	.368	.131	.023	.129	551.49*	.305–.431	.156–.580

Note. SES = socioeconomic status; *K* = number of samples; *N* = total sample size; *SDr* = observed standard deviation of correlations; *SD_e* = standard deviation expected due to sampling error; *SD_p* = residual standard deviation.

* *Q* test for homogeneity, significant at *p* < .05.

Table 7
Meta-Analysis of SES–Test Correlations

Variable	<i>K</i>	<i>N</i>	Mean <i>r</i>	<i>SD_r</i>	<i>SD_e</i>	<i>SD_p</i>	<i>Q</i>	95% Confidence interval	90% Credibility interval
Total	55	60,565	.247	.101	.031	.096	583.82*	.221–.273	.089–.405
Admissions	36	26,557	.166	.066	.042	.051	88.90*	.146–.186	.082–.250
Nonadmissions	19	34,008	.311	.075	.021	.072	242.44* (251.66*)	.277–.345	.196–.429
Three-variable	37	25,839	.166	.072	.044	.057	99.07*	.144–.188	.072–.260
Two-variable	18	34,726	.308	.074	.020	.071	246.42* (238.33*)	.274–.342	.191–.425
SES composite	36	52,533	.255	.102	.026	.099	554.06*	.222–.288	.093–.417
SES single indicator	19	8,032	.197	.077	.054	.055	38.63* (0.00)	.165–.299	.107–.288
Father's education	30	50,717	.284	.081	.023	.078	372.08*	.255–.313	.145–.412
Mother's education	31	46,529	.223	.094	.027	.090	375.74*	.190–.256	.075–.371
Father's occupation	19	32,504	.208	.081	.026	.077	184.41*	.172–.244	.082–.334
Income	30	30,980	.186	.104	.035	.098	264.88*	.149–.223	.025–.347
Admissions, SES composite, three-variable	17	17,235	.154	.061	.037	.048	46.21*	.127–.181	.074–.234

Note. *Q* values in parentheses are tests of between-condition differences. SES = socioeconomic status; *K* = number of samples; *N* = total sample size; *SD_r* = observed standard deviation of correlations; *SD_e* = standard deviation expected due to sampling error; *SD_p* = residual standard deviation.

* *Q* test for homogeneity, significant at $p < .05$.

This is consistent with the Model 2 depiction of the SES–grade relationship being mediated by test score.

Supplemental Analyses of SES–Test Relationships

Table 7 summarizes the results of the meta-analyses of SES–test relationships. Results are reported for the full sample, for admissions versus other tests, for studies containing all three variables versus studies with only SES and test scores, for studies using an SES composite versus a single SES indicator, and for individual SES indicators. Finally, for comparison purposes, we also include the results of our central analyses shown earlier in Table 6 for the studies using a combination of admissions tests, those including all three key variables, and those using a composite SES measure.

Whereas the overall mean SES–test correlation was .25, sizable and significant differences were found for studies using admissions tests (mean $r = .17$) versus other tests (mean $r = .31$). Similarly, differences were found between studies containing all three variables (mean $r = .17$) and studies containing only SES and test scores (i.e., studies not including grades; mean $r = .31$). These reflect the same phenomenon, as studies involving admissions tests were virtually the same as the set of studies including all three variables. In essence, studies in college contexts used an admissions test and had grades available. Studies of other types of tests tended to use broader, noncollege samples and, thus, tended not to include college grades as a variable. This phenomenon helps explain the difference in findings between admission and other tests. As the earlier multi-institution College Board study showed, studies within a single school are based on more restricted samples and thus produce lower SES–test correlations.

Studies using an SES composite showed higher mean correlations (mean $r = .26$) than studies using a single SES indicator (mean $r = .20$), though the difference was not statistically significant. This is an expected finding: A composite SES measure should be more reliable than a single indicator and thus would be expected to correlate more highly with test scores.

Finally, as noted earlier, the studies using admissions tests, containing all three variables, and using a composite SES measure

produced a mean correlation of .15. We interpret this as our best estimate from this set of meta-analyses of the SES–admissions test relationship in samples of admitted college students.

Supplemental Analyses of SES–Grade Relationships

Table 8 summarizes the results of the meta-analyses of test–grade relationships. Results are reported for the full sample, for admission versus other tests, for studies containing all three variables versus studies with only SES and grades, for studies using an SES composite versus a single SES indicator, and for individual SES indicators. Finally, for comparison purposes, we also include the results of our central analyses shown earlier in Table 6 for studies using a combination of admissions tests, including all three key variables, and using composite SES measures.

The overall mean SES–grade r was .09. Mean correlations did not differ significantly for studies using admissions tests versus other tests, for studies containing all three variables versus studies containing only SES and grade (i.e., studies not including tests), or for studies using an SES composite versus a single indicator.

As noted earlier, the studies using admissions tests, containing all three variables, and using a composite SES measure produced a mean r of .09. We interpret this as our best estimate from this set of meta-analyses of the SES–grade relationship in samples of admitted college students.

Supplemental Analyses of Test–Grade Relationships

Table 9 summarizes the results of the meta-analyses conducted across all test–grade relationships. Results are reported for the full sample and separately for studies using admissions tests versus other tests. Recall that our search for test–grade correlations was limited to studies that also provided SES–test and SES–grade correlations, as many other studies have collected and meta-analyzed large numbers of studies of test–grade relationships. We note that the findings of the present meta-analysis (mean $r = .35$ across all studies; mean $r = .37$ in the subset of studies using admissions tests and composite SES measures) converge with

Table 8
Meta-Analysis of SES–Grade Correlations

Variable	<i>K</i>	<i>N</i>	Mean <i>r</i>	<i>SD_r</i>	<i>SD_e</i>	<i>SD_p</i>	<i>Q</i>	95% Confidence interval	90% Credibility interval
Total	65	41,829	.087	.065	.051	.040	105.58*	.073–.101	.021–.153
Admissions	31	21,187	.093	.065	.049	.043	54.55*	.073–.113	.023–.163
Nonadmissions	34	20,642	.081	.064	.053	.036	49.58* (1.45)	.063–.099	.022–.140
Three-variable	37	26,634	.087	.067	.048	.047	72.09*	.068–.106	.010–.164
Two-variable	28	15,195	.088	.061	.055	.026	34.44* (0.00)	.069–.107	.045–.131
SES composite	41	31,936	.092	.058	.046	.035	65.18*	.077–.107	.034–.150
SES single indicator	22	9,520	.071	.081	.063	.051	36.67* (3.73)	.042–.100	–.013–.155
Father’s education	24	15,440	.090	.061	.050	.035	35.72*	.069–.111	.033–.147
Mother’s education	26	17,814	.061	.055	.051	.021	30.24*	.021–.044	.027–.098
Father’s occupation	6	5,510	.031	.032	.045	.000	3.03	.004–.057	.031–.031
Income	31	18,457	.057	.052	.054	.000	28.75*	.043–.071	.057–.057
Admissions, SES composite, three-variable	17	17,630	.092	.059	.040	.043	36.99*	.067–.117	.021–.163

Note. *Q* values in parentheses are tests of between-condition differences. SES = socioeconomic status; *K* = number of samples; *N* = total sample size; *SD_r* = observed standard deviation of correlations; *SD_e* = standard deviation expected due to sampling error; *SD_p* = residual standard deviation. * *Q* test for homogeneity, significant at *p* < .05.

other estimates. For example, Willingham, Lewis, Morgan, and Ramist (1990) reported an uncorrected correlation of .37 between SAT and freshman GPA on the basis of data from a sample of 124 colleges.

Test–Grade Relationships, Controlling for SES

On the basis of the results reported in Tables 6, 7, 8, and 9, partial correlations were computed to determine the validity of tests for predicting college grades when controlling for SES. These findings are reported in the middle section of Table 4. We computed these separately for the full sample and for the subset of studies that use admissions tests, include all three key variables, and use a composite SES measure. In the full sample, the mean correlation between test and grades was .35. After controlling for SES, this correlation was reduced by .01 to .34. Similarly, as noted earlier, in the subset of studies using admissions tests, containing composite SES measures, and including all three key variables, the test–grade mean correlation of .37 was reduced by .01 to .36. Thus, controlling for SES had a negligible effect on the test–grade relationship in this study.

SES–Grade Relationships, Controlling for Test

On the basis of the results reported in Tables 6, 7, 8, and 9, partial correlations were computed to examine the SES–grade

correlation controlling for test. These findings are reported in the middle section of Table 4. We computed these separately for the full sample and for the subset of studies that used admissions tests, included all three key variables, and used a composite SES measure. In the full sample, the mean correlation between SES and grades was .09. After controlling for test, this correlation was reduced to .00. Similarly, as noted earlier, in the subset of studies using admissions tests, containing composite SES measures, and including all three key variables, the SES–grade mean correlation of .09 was reduced to .03. Thus the SES–grade relationship dropped substantially, to zero or near zero, when controlling for test score. This is consistent with the Model 2 depiction of the SES–grade relationship being mediated by test score.

Investigation 3: Re-Analysis of Longitudinal Data Sets

Method

Our final investigation focused on re-analysis of existing publicly available data sets, which contained test scores, SES measures, and measures of academic performance. Note that a number of these studies deal with settings other than college admissions (e.g., tests of ability–SES relationships in samples of high school seniors) For each data set, we had access to the primary data. Data sets included Project TALENT, the National Longitudinal Study of

Table 9
Meta-Analysis of Test–Grade Correlations

Variable	<i>K</i>	<i>N</i>	Mean <i>r</i>	<i>SD_r</i>	<i>SD_e</i>	<i>SD_p</i>	<i>Q</i>	95% Confidence interval	90% Credibility interval
Total	37	26,127	0.352	.132	.029	.129	766.57*	.309–.395	.140–.564
Admissions	31	21,045	0.347	.142	.030	.139	693.454*	.297–.397	.119–.575
Nonadmissions	6	5,082	0.372	.077	.025	.073	56.920* (16.20)	.309–.435	.025–.492
Admissions, SES composite, three-variable	17	17,244	0.368	.131	.023	.129	551.490*	.305–.431	.156–.580

Note. *Q* values in parentheses are tests of between-condition differences. *K* = number of samples; *N* = total sample size; *SD_r* = observed standard deviation of correlations; *SD_e* = standard deviation expected due to sampling error; *SD_p* = residual standard deviation. * *Q* test for homogeneity, *p* < .05.

the High School Class of 1972 (NLS-1972), the National Educational Longitudinal Study of 1988, the Harvard Study of the classes of 1964 and 1965, the Law School Admissions Council National Bar Passage Study, and the 1995 National Study of Law School Performance. The SES variables examined in these studies are the same three used across the first two investigations: father's education level, mother's education level, and family income. All studies contained multiple SES indicators; these were combined into unit-weighted composites (Ghiselli, Campbell, & Zedeck, 1981). The log of income was used whenever possible in order to reflect the likely diminishing returns of income and to normalize income variables.

Description of Data Sets

Brief descriptions of the data sets used in this study follow. Project Talent is a longitudinal study run by the U.S. Department of Health, Education, and Welfare. This study includes information on factors that support or restrain the development of human talents. We had access to a publicly available subsample of about 4,000 students (Flanagan et al., 2001); our analyses focused on those students for whom college grades were available. Project TALENT used over 30 cognitive ability measures to assess a large number of cognitive abilities. Estimated in the study was a composite IQ score, which was used in our analyses.

The National Longitudinal Study of the High School Class of 1972 (NLS-1972) consisted of surveys administered to seniors in high school in 1972 and included follow-up surveys in 1973, 1974, 1976, and 1986 (U.S. Department of Education, National Center for Education Statistics, 1999). The NLS-1972 was one of the largest studies done on a single generation of Americans. The sampling process created a stratified two-stage probability sample. Schools were oversampled for race and income; analyses reported here weight the data to be representative of the population. NLS-1972 included SAT-V and SAT-M and ACT total scores, which were used in this study. The SAT subscales were combined into an overall score. If participants had scores on the SAT and the ACT, they were combined after being standardized within test. Because the ACT and the SAT measure similar abilities, this procedure was used to maximize the number of participants available for the analyses, as not all students take both tests. Results based on a combined sample are nearly identical to results obtained within each test separately. Analyses were conducted using probability weights appropriate for the Time 1 and first follow-up data to obtain more accurate estimates.

The National Educational Longitudinal Study of 1988 (NELS88) examined a national sample of eighth graders in 1988. A sample of these respondents was then resurveyed through four follow-ups in 1990, 1992, 1994, and 2000 (U.S. Department of Education, National Center for Education Statistics, 2004). The data from this longitudinal study included standardized tests of verbal and mathematical abilities. These two ability scores were combined into a standardized test composite and served as the ability measure in our analyses.

The 1995 National Study of Law Student Performance invited all American Bar Association–approved law schools to participate in the study; 21 law schools did so (Sander, 1995). We examined the effects of SES on the correlations between the Law School Admissions Examination (LSAT) and overall law school GPA

separately for each of these 21 law schools. These data were aggregated meta-analytically using the Hunter and Schmidt psychometric meta-analytic method (Hunter & Schmidt, 2004).

The Law School Admission Council (LSAC) National Longitudinal Bar Passage Study monitored the 1991 entering cohort of 163 of the 172 accredited U.S. Law Schools (Wightman, 1998). We examined the effects of SES on the LSAT–GPA correlation. Unlike the 1995 National Study of Law School Performance, the LSAC study does not contain identifying information that permits separate analysis by school. Thus, the two law school studies differed in this respect: One was a within-school analysis, and the other was an across-school analysis.

The Harvard Study of the classes of 1964 and 1965 was a comprehensive assessment of two incoming Harvard classes (Harms & Roberts, 2005). SAT-V and SAT-M scores were combined into a total score for our analyses.

Prior Research With These Data Sets

The sheer scope and quality of the publicly available databases has made them the focus of a truly vast number of studies. Research on some databases has been so extensive as to warrant the publication of lengthy research reports that are simply annotated bibliographies compiling studies done during a decade or so of time (Maline, 1993; Taylor, Stafford, & Place, 1981). The results that we present overlap with a number of studies, and we briefly review some of this literature to provide context, background, and support for our analyses. This review illustrates that researchers have utilized diverse approaches to analyze data from these large-scale longitudinal studies. To eliminate the confound of inconsistent methodology, we computed the results reported here.

The NLS-1972, Project Talent, and the NELS88 have all been cut and analyzed for a number of different purposes. Examining educational outcomes and their relationship with ability and SES has been common. However, the most common outcome of interest appears to have been educational and occupational attainment rather than academic performance during school. In addition, a number of studies have cut the data on racial or other demographic variables to examine these effects for specific groups.

In keeping with the results that we present, relatively consistent patterns of outcomes are evident across the databases. For example, using the NLS 1972 study, we found persistence in higher education to be most strongly related to more proximal variables such as ability measures and academic performance, with negligible residual relationships for SES (Alexander, Riordan, Fennessey, & Pallas, 1982; Hilton, 1982). Similarly, other studies, although not focused on the questions we examine, present results with the same patterns of findings, with small correlations between SES and college grades, intermediate correlations between SES and ability, and largest correlations between ability and college performance (Peng & Fetters, 1978).

The NELS88 and Project Talent data have been used to examine the same questions with a heavy emphasis on completion and success in high school and occupational outcomes, respectively. Results indicate positive but modest correlation between SES and ability measured from middle school to the end of high school (Mullis, Rathge, & Mullis, 2003; Stewart, 2006). Analyses of the data from Project Talent have also yielded a similar pattern of modest relationships between SES and ability and the finding of

relationships between ability and occupational and educational attainment, even when background variables were controlled (Jencks, Crouse, & Mueser, 1983). The importance of ability with significant but weaker relationships for SES variables has been reported with the Project Talent data for college completion (Bayer, 1968) and occupational attainment (Austin & Hanisch, 1990).

Thus although these data sets have been extensively examined, and whereas various individual correlations of interest have been reported in various publications and reports, we computed all results reported below from the publicly available data sets rather than extracting them from existing reports. This was done to ensure consistency in treatment of the data across data sets.

Results

Table 10 presents the results of the meta-analysis of the 1995 National Study of Law Student Performance data across 21 law schools. The mean SES–LSAT correlation was .16, the mean SES–GPA correlation was .07, and the mean LSAT–GPA correlation was .38. The partial correlation for LSAT predicting grades when controlling for SES was .38. This partial correlation was the same, to two decimal places, as the correlation between LSAT predicting grades directly without controlling for SES, indicating that LSAT was not synonymous with SES when predicting grades. The relationship between LSAT and grades did not change when controlling for SES. The partial correlation for SES predicting grades when controlling for standardized test scores was virtually zero, indicating that standardized tests scores captured almost everything that SES did, and substantially more.

The correlations for the rest of the longitudinal studies are presented in the bottom section of Table 4. In general, the correlations between SES and grades were smaller than the correlation between test scores and grades and tended to be smaller than the correlations between test scores and SES. SES was generally a weaker predictor of grades than test scores and did not account for the majority of the variance in test scores.

For the Harvard data set, the most restricted sample evaluated, both on ability and SES, the correlation for SES and test scores was .07, and the correlation between SES and grades was .05. However, the correlation for test scores and grades was .30. The diversity of students in SES and race who attended Harvard University when these data were obtained was very limited when compared with the diversity of students used in the other studies. However, the relationship between test scores and grades remained

strong, demonstrating once again that SES and tests were not measuring the same things.

In contrast, the NLS-1972 had a much larger correlation between SES and test scores, more in keeping with other broad samples. The results showed a stronger correlation between test scores and grades, and the SES–test and SES–grades correlations were much larger. Controlling for SES had little influence on the predictive power of ability measures.

Across the data sets, correlations between test scores and grades were very similar to the partial correlations between test scores and grades when SES was controlled. In contrast, the partial correlations for SES and grades, controlling for test scores, answers the question, If everyone has the same test score, what does SES tell us beyond test scores in predicting grades? These data are also presented in Table 4, and all partial correlations were close to zero. SES was found to be a weaker predictor of academic success to begin with and added almost nothing above test scores.

Conclusions and Discussion

Our analyses of multiple large data sources produced consistent findings. First, SES was indeed related to admissions test scores. In broad, unrestricted populations, this correlation was quite substantial (e.g., $r = .42$ among the population of SAT takers). Second, scores on admissions tests were indeed predictive of academic performance, as indexed by grades. Observed correlations in samples of admitted students averaged about $r = .35$ for admissions tests; applying range restriction corrections to estimate the validity for school-specific applicant pools resulted in an estimate of .47 as the operational validity. Third, the test–grade relationship was not an artifact of common influences of SES on both test scores and grades. Partialing SES from the above estimate of the operational validity of admissions tests ($r = .47$) reduced the estimated validity to .44. Fourth, the SES–grade relationship was consistent with a model of a mediating mechanism in which SES influences test scores, which are subsequently predictive of grades. SES had a near-zero relationship with grades other than through this SES–test–grade chain of relationships.

These findings are paralleled closely for cognitively loaded tests other than college admissions test (e.g., tests of general cognitive ability). The meta-analytic mean test–grade correlation was .37; partialing SES from this correlation resulted in an estimate of .36. Analyses of large data sets that include admissions tests for contexts other than undergraduate entry (i.e., law school admission) and cognitive ability tests used with nationally representative samples of high school students also showed very small reductions in test–grade correlations when controlling for SES.

Relationship of Findings to Related Studies

The results presented here are at odds with the critics’ claims presented earlier that large-scale University of California data show that the predictive power of the SAT drops to zero when SES is controlled. This contrast warrants some specific discussion. Whereas such a claim has been made by Geiser and Studley (2002) in their analysis of the University of California data, that work does not, in fact, actually provide data supporting those conclusions, as has been pointed out by others scholars who have reanalyzed their data (Johnson, 2004; Zwick, Brown, & Sklar, 2003).

Table 10
Summary of the 1995 National Study of Law Student Performance

Variables	<i>k</i>	<i>N</i>	\bar{r}	Residual <i>SD</i>
SES–LSAT	21	3,376	.16	.13
SES–grades	21	3,375	.07	.11
LSAT–grades	21	4,142	.38	.14

Note. SES = socioeconomic status; LSAT = Law School Admission Test; *k* = number of studies; *N* = total sample size; \bar{r} = sample size weighted mean observed correlation; Residual *SD* = standard deviation of correlations after controlling for sampling error.

Such a conclusion would require partialing SES from SAT–grade correlations or regression analyses showing that the SAT’s predictive power disappears when SES is added to the model. However, Geiser and Studley’s main focus was a comparison of the SAT I and the SAT II, the latter of which is a composite of specific subject area tests. The authors estimated a regression model with SAT I, SAT II, high school grades, and SES as predictors of GPA. They asked questions about whether controlling for SES reduces the incremental contribution of one test over another, which is very different from asking whether the SAT I used alone remains predictive of grades when SES is controlled. Because SAT I and SAT II are highly correlated, the incremental contribution of either one over the other will be quite small, even if both are predictive of grades. In fact, reanalyses revealed that the SAT I–GPA correlation changed from .38 to .35 when partialing out SES, a finding fully consistent with the data we present here (Johnson, 2004; Zwick et al., 2003). Thus, data from the University of California also do not support critics’ claims.

It is useful to compare our conclusions with those of Rothstein (2004), who also examined SAT–grade relationships net of a set of variables, including SES-related variables, in a large University of California data set. On the surface, the two studies appear contradictory. We report that test–grade relationships are minimally affected by controlling for SES. Rothstein reported that a sizable portion of the SAT’s predictive power disappears once a set of demographic variables is controlled. In fact, the goals and methods of the two studies differ in a number of ways. Most fundamentally, Rothstein’s interest was in identifying the predictive power of the portion of variance in SAT scores that is unique, that is, not shared with other information potentially available to those making admissions decisions. Thus, his analysis removed variance shared with high school grades, with race, and with characteristics of the high school a student attended (e.g., racial group percentages, average level of parental education). Our interest in the present study was not in the role of test variance that is not shared with other variables, but rather in the question of whether the apparent test–grade relationship could be attributed to variance that both test and grade share with SES.

Our study also differs from Rothstein’s (2004) in the variables examined. He studied race, gender, and several measures at the school level. Although he included SES-related variables at the school level (e.g., average level of parental education at the school), he did not include individual student SES measures. Thus the findings of the studies are not mutually contradictory: That test scores share variance with some demographic variables is not inconsistent with our finding that test–grade correlations are not an artifact of both variables sharing variance with SES.

That test scores are related to some degree to both individual- and school-level background variables is not surprising. A long history of research in developmental psychology has demonstrated the importance of parental involvement, parenting habits, absence of environmental toxins, strong schools, good teachers, and positive academic attitudes on the academic performance of children (e.g., Hubbs-Tait, Nation, Krebs, & Bellinger, 2005; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998). Unfortunately, this mixture of helpful and harmful developmental factors also correlates with SES. As demonstrated here, and by many other scholars, measures of academic preparedness, in general, are associated to some degree with SES. Zwick (2004), for example,

presented evidence that academic achievement and preparedness, including curricula-specific tests, high school grades, and teacher perceptions are associated with social class variables. Test scores demonstrate stable SES differences back to the 4th grade. Zwick’s analysis of these low stakes assessments further illustrated that these correlations are not the function of coaching and test preparation. SES differences in parenting behavior go back even further to the time when children are just beginning to learn language. The parents of children who develop large vocabularies speak, on average, millions more words to their children than do the parents of low-verbal children (Hart & Risley, 1995). The volume of parental word production was associated with social class and IQ. Some of these effects in older children are obscured in high school GPA and high school rank data because of within- versus across-school effects (Zwick & Green, 2007), which are partially the result of differential grading practices (Willingham, Pollack, & Lewis, 2002).

The observed differences in cognitive and noncognitive skills across social class are widely documented, and their link to subsequent academic performance and financial outcomes is strong (e.g., Heckman, 2006). What is truly noteworthy is that research both here and elsewhere shows that test scores contain significant predictive information beyond community-level and individual-level SES variables (e.g., Rothstein, 2004; Zwick, 2004). The demographic variables controlled in our research and those of other researchers are proxies for the family, school, neighborhood, and biological factors that influence academic preparedness, as measured by prior grades and test scores. That is, the relationship of developed reasoning ability, verbal skills, and mathematics skills with SES is partially the consequence of positive and negative developmental effects.

Limitations

We believe that the research described here has many strengths. A major one is the scope of the investigation: the use of multiple data sets and the convergence of findings across data sets make it virtually certain that results are not a function of the use of an idiosyncratic data set, a specific test, or a specific operationalization of SES.

In terms of limitations, although we clearly recognize the importance of range restriction, we were only able to correct for range restriction in the 41-school multi-institution data set, as other data sets examined did not contain the data needed for such corrections. However, we note that a lack of needed data for range restriction corrections is a common problem in meta-analytic syntheses of prior literature.

We also note that we rely on self-reports of SES, as these measures are typically obtained from questionnaires completed by students. It is possible that these self-reports may be in error for some students. However, we believe that students are generally in a good position to report parents’ occupation and educational attainment. Looker (1989) reviewed studies of agreement between student and parent reports; focusing on 12th-grade samples, we observe a mean student–parent correlation of .82 for father’s occupation, .86 for father’s education, and .85 for mother’s education. Error in reports of parental income may be more likely. We note, nevertheless, that SES–test and SES–grade correlations were highly similar regardless of the SES indicator used, and we also

note that we used a composite across multiple SES indicators throughout the study.

It should be noted that grades and tests differ in their reliability. Because test scores are consistently somewhat more reliable than a first-year GPA, it is important to consider the implication of differential reliability on our results. Overall, a more reliable GPA would result in a larger correlation with both SES measures and test scores. That is, if GPAs were made to be as reliable as tests, GPAs would be more predictable both by SES and test scores. This would increase the amount of variance in GPA that is uniquely attributable to SES to a minor degree while increasing the total amount of variance attributable to test scores to a larger degree given their comparatively larger relationship with GPA. Overall, the conclusions and implications of our study would not be meaningfully altered.

Finally, we have examined the influence of SES on the test-grade relationships. Although it is clear that these types of tests predict a range of important outcomes (Kuncel & Hezlett, 2007; Kuncel, Hezlett, & Ones, 2004) and that grades are related to important life outcomes (Roth, BeVier, Switzer, & Schippmann, 1995; Roth & Clarke, 1998), it would be valuable to further extend these analyses to other outcome and performance measures.

In conclusion, our work focuses on the predictive power of admissions tests and other cognitively loaded tests in predicting college grades and shows that this power is not an artifact of SES. In fact, tests retain virtually all of their predictive power when controlling for SES.

References

References marked with an asterisk indicate studies included in the meta-analysis reported as Investigation 2.

- Aertgeerts, B., & Buntinx, F. (2002). The relation between alcohol abuse or dependence and academic performance in first year college students. *Journal of Adolescent Health, 31*, 223–225.
- *Alao, S., Coffey, G., Ellington, R., & Wright, H. (1998). *Predicting placement test scores with cognitive and socio-economic variables*. Washington, DC: U.S. Department of Housing and Urban Development. (ERIC Document Reproduction Services No. 435627).
- Alexander, K. L., Riordan, C., Fennessey, J., & Pallas, A. M. (1982). Social background, academic resources, and college graduation: Recent evidence from the National Longitudinal Survey. *American Journal of Education, 90*, 315–333.
- *Allen, W. R. (1981). Correlates of Black student adjustment, achievement, and aspirations at a predominantly White southern university. In G. E. Thomas (Ed.), *Black students in higher education: Conditions and experiences in the 1970s*. Westport, CT: Greenwood Press.
- *Allen, W. R. (1985). Black student, White campus: Structural, interpersonal, and psychological correlates of success. *Journal of Negro Education, 54*, 134–147.
- *Allen, W. R. (1992). The color of success: African-American college student outcomes at predominantly White and historically Black public colleges and universities. *Harvard Educational Review, 62*, 26–44.
- *Allen, W. R., & Haniff, N. Z. (1991). Race, gender, and academic performance in U.S. higher education. In W. R. Allen, E. G. Epps, & N. Z. Haniff (Eds.), *College in Black and White*. Albany, NY: State University New York Press.
- *Alnabhan, M., Al-Zegoul, E., & Harwell, M. (2001). Factors related to achievement levels of education students at Mu'tah University. *Assessment and Evaluation in Higher Education, 26*, 593–604.
- *Anthony, N. R., Sanders, O. H., & Kahle, J. B. (1981). Ability/aptitude, personal, and social characteristics of black college students. *Journal of Negro Education, 50*, 346–353.
- Atkinson, R. C. (2005). College admissions and the SAT: A personal perspective. *APS Observer, 18*, 15–22.
- Austin, J. T., & Hanisch, K. A. (1990). Occupational attainment as a function of abilities and interests: A longitudinal analysis using Project TALENT data. *Journal of Applied Psychology, 75*, 77–86.
- *Bachman, J. G., & O'Malley, P. M. (1986). Self-concepts, self-esteem, and educational experiences: The frog pond revisited (again). *Journal of Personality and Social Psychology, 50*, 35–46.
- *Baker, S. D. (1998). *Factors influencing the college achievement of African-American males at a predominantly White institution*. Unpublished doctoral dissertation, Chicago School of Professional Psychology.
- Bayer, A. E. (1968). The college drop-out: Factors affecting senior college completion. *Sociology of Education, 41*, 305–316.
- Becker, B. J., & Schram, C. M. (1994). Examining explanatory models through research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 357–377). New York: Russell Sage Foundation.
- *Best, W. P. (1968). The prediction of success in nursing education. *Dissertation Abstracts International, 29*(08), 2558. (UMI No. 6902888).
- Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist, 58*, 1019–1027.
- *Braddock, J. H., II. (1981). Desegregation and black student attrition. *Urban Education, 15*, 403–418.
- Bridgeman, B., Pollack, J., & Burton, N. (2004). *Understanding what the SAT reasoning test scores add to high school grades: A straightforward approach* (College Board Research Report No. 2004–4). New York: College Board.
- *Brigham, I. (1981). Selected academic and non-academic variables as predictors of academic success of male student-athletes at Southern Illinois University at Carbondale. *Dissertation Abstracts International, 42*(10), 4417. (UMI No. 8206446).
- *Brodnick, R. J., & Ree, M. J. (1995). A structural model of academic performance, socioeconomic status, and Spearman's *g*. *Educational and Psychological Measurement, 55*, 583–594.
- 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.
- *Burnham, P. S., & Hewitt, B. A. (1967). *Thirty year follow-up of male college students* (Project No. 5–0816). Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare.
- *Burton, G. E. (1976). Prediction of college grades from selected background factors. *College Student Journal, 10*, 10–14.
- *Calkins, E. V., Willoughby, T. L., & Arnold, L. M. (1982). Predictors of performance of minority students in the first two year of a BA/MD program. *Journal of National Medical Association, 74*, 625–632.
- *Carlson, D. C. (1967). *A descriptive and correlational study of variables related to success in the San Bernardino Valley College two-year registered nurse program*. Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare. (ERIC Document Reproduction Service No. ED 016459).
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- *Coates, D. C. (1996). The use of cognitive and noncognitive variables in predicting the cumulative grade point average of student athletes. *Masters Abstracts International, 34*(06), 2135. (UMI No. 1380426).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Colvin, R. L. (1997, October 1). Q & A: Should UC do away with the SAT? *Los Angeles Times*, p. B2.
- Crosby, F. J., Iyer, A., Clayton, S., & Downing, R. A. (2003). Affirmative action: Psychological data and the policy debates. *American Psychologist, 58*, 93–115.

- *Crumb, L. W. (1986). A study of the projected functioning of students in society from an interdisciplinary liberal arts program. *Dissertation Abstracts International*, 47(08), 2991. (UMI no. 8625530).
- *Donovan, R. (1983). Persistence in higher education among low income black youth: A path analysis. *Dissertation Abstracts International*, 44(08), 2880. (UMI No. 8327205).
- *Epps, E., & Jackson, K. (1985). *Educational and occupational aspirations and early attainment of black males and females*. Atlanta, GA: Southern Education Foundation. (ERIC Document Reproduction Service No. 263818).
- *Falb, M. E. (1995). Psychological hardiness as a predictor of academic performance and its relationship to stress and health status of community college transfer students. *Dissertation Abstracts International*, 56(07), 4049.
- Flanagan, J. C., Tiedeman, D. V., et al. (2001). *Project talent public use file, 1960–1976* [ICPSR version]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research. (Original work published 1976).
- Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastics assessment test and general cognitive ability. *Psychological Science*, 15, 373–378.
- *Gansemer, L. P., & Bealer, R. C. (1977). *Family background and occupational attainment: Replication and extension through a 24-year follow-up*. Washington, DC: U.S. Department of Agriculture. (ERIC Document Reproduction Service No. ED 139579).
- Geiser, S., & Studley, R. (2002). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment*, 8, 1–26.
- Ghiselli, E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W. H. Freeman.
- *Gillie, A. C. (1972). *The differential effects of selected programs on the performance, degree of satisfaction, and retention of community college women students* (Report No. OEC-3-710127). Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare. (ERIC Document Reproduction Service No. ED068086).
- *Glenn, H., Sr. (1970). *Predicting academic achievement of "special admit" students at the Pennsylvania State University*. Unpublished doctoral dissertation, Pennsylvania State University.
- *Greever, K. B., Tseng, M. S., & Udell, F. B. (1973). *An Adlerian approach to measuring change in college freshman*. Paper presented at the annual conference of the American Personnel and Guidance Association, San Diego, CA.
- *Guerrero, B. G. (2000). An analysis of academic, demographic, and non-cognitive factors that influence academic performance during the freshman year in college. *Dissertation Abstracts International*, 61(07), 2591. (UMI No. 9979648).
- Harms, P., & Roberts, B. (2005). *Harvard study of the classes of 1964–65*. Unpublished data, University of Illinois at Urbana–Champaign.
- *Harrison, C. E. (1981). Assessing the value of a remedial reading program in higher education and the transferability of successful student performance to content area courses. *Dissertation Abstracts International*, 42(05), 2047. (UMI No. 8123059).
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- *Hearn, J. C. (1981). *Reconceptualizing equity in postsecondary participation patterns*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, CA.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312, 1900–1902.
- Hedges, L. M., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- Hezlett, S. A., Kuncel, N. R., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., & Camara, W. (2001). The predictive validity of the SAT: A meta-analysis. In D. S. Ones & S. A. Hezlett (Chairs), *Predicting performance: The interface of I-O psychology and educational research*. Symposium conducted at the 16th annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- *Hill, W. M., & Woerdehoff, F. J. (1966). Prediction of academic achievement in beginning German. *The Journal of Experimental Design*, 34, 94–99.
- Hilton, T. L. (1982). Persistence in higher education [College Board Research Report 2–5]. Princeton, NJ: The College Board.
- *Horn, R. A. (2004). In search of parsimony: An evaluation of congruence indices used in testing Holland's theory and academic achievement in college students. *Dissertation Abstracts International*, 65(03), 828. (UMI No. 3127331).
- Hubbs-Tait, L., Nation, J. R., Krebs, N. F., & Bellinger, D. C. (2005). Neurotoxicants, micronutrients, and social environments: Individual and combined effects on children's development. *Psychological Science in the Public Interest*, 6, 57–121.
- Huffman, T. E., Sill, M. L., & Brokenleg, M. (1986). College achievement among Sioux and White South Dakota students. *Journal of American Indian Education*, 32–39.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- *Ikeda, K., Wolfe, J. A., & Rich, R. F. (1971). *The association of racial status, socioeconomic status, and measured ability upon academic performance in a liberal arts college*. Washington, DC: National Institute of Education, U.S. Department of Health, Education and Welfare. (ERIC Document Reproduction Service No. ED084319).
- Jencks, C., Crouse, J., & Mueser, P. (1983). The Wisconsin model of status attainment: A national replication with improved measures of ability and aspiration. *Sociology of Education*, 56, 3–19.
- Johnson, J. W. (2004). Not all affirmative action rewards merit: Comment. *American Psychologist*, 59, 123–124.
- *Just, H. D. (1998). Freshman adjustment and retention: Combining traditional risk factors with psychosocial variables. *Dissertation Abstracts International*, 59(09), 5138. (UMI No. 9905763).
- *Keiller, S. W. (1997). Type A behavior pattern and locus of control as predictors of college academic performance. *Dissertation Abstracts International*, 58(05), 2737. (UMI No. 9731564).
- *Kinsella, M. A. (1995). Predictors of performance of graduates of pre-college enrichment programs. *Dissertation Abstracts International*, 56(05), 1688. (UMI No. 9530245).
- Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, 36, 153–160.
- Kohn, A. (2001). Two cheers for an end to the SAT. *Chronicle of Higher Education*, 9, p. B-12.
- Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, 315, 1080–1081.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? [Special section]. *Journal of Personality and Social Psychology*, 86, 148–161.
- Kuncel, N. R., & Klieger, D. M. (2007). Application patterns when applicants know the odds: Implications for selection research and practice. *Journal of Applied Psychology*, 92, 586–593.
- *Larkin, G. J. (1981). *An examination of returning resident freshman students at colleges of business in New England*. Unpublished doctoral dissertation, Boston College.
- *Lee, R. (1986). *An investigation of the personality, academic, and social background characteristics of highly successful, moderately successful, and non-successful second year college students*. Unpublished doctoral dissertation, Georgia State University.
- Looker, E. D. (1989). Accuracy of proxy reports of parental status characteristics. *Sociology of Education*, 62, 257–276.
- *Lovato, C. Y. (1981). Achievement motivation, role stress, and individual characteristics as predictors of college performance among Mexican

- Americans and Anglo Americans. *Dissertation Abstracts International*, 42(11), 4770. (UMI No. 8208205).
- *Lowman, R. P., & Spuck, D. W. (1975). Predictors of college success for the disadvantaged Mexican-American. *Journal of College Student Personnel*, 16, 40–48.
- Luthar, S. S., & Latendresse, S. J. (2005). Children of the affluent: Challenges to well-being. *Current Directions in Psychological Science*, 14, 49–53.
- Maline, M. S. (1993). *The National Longitudinal Study of the High School Class of 1972: Annotated bibliography of studies, 1980–1992* (OR-93–3156). Washington, DC: Office of Educational Research and Improvement.
- McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist*, 53, 185–204.
- *McQuary, J. P. (1951). *Relationships between non-intellectual characteristics and academic achievement*. Unpublished doctoral dissertation, University of Wisconsin.
- *Mittal, N. (1989). Predictors of Pre-Professional Skills Tests (PPST) scores. *Dissertation Abstracts International*, 51(01), 117.
- *Moline, A. E. (1987). Financial aid and student persistence: An application of causal modeling. *Research in Higher Education*, 26, 130–147.
- Mullis, R. L., Rathge, R., & Mullis, A. K. (2003). Predictors of academic performance during early adolescence: A contextual view. *International Journal of Behavioral Development*, 27, 541–548.
- *Nettles, M. T., Thoeny, A. R., & Gosman, E. J. (1986). Comparative and predictive analysis of Black and White students' college achievement and experiences. *Journal of Higher Education*, 57, 289–318.
- *Nonnamaker, J. B. (2000). Pre-college Internet use and freshman year academic achievement in a private college: The effect of sociodemographic characteristics, family socioeconomic status, academic ability and high school experiences. *Dissertation Abstracts International*, 61(07), 2937. (UMI No. 9981407).
- Nottingham, J. H. (1990). *A health careers summer enrichment program and Black college students' freshman year achievement in biology*. Unpublished doctoral dissertation.
- *Noxel, S. A. (2000). The application of investment theory to predict baccalaureate degree enrollment intensity. *Dissertation Abstracts International*, 61(02), 526. (UMI No. 9962438).
- *Oliver, M. L., Rodriguez, C. J., & Mickelson, R. A. (1985). Brown and Black in White: The social adjustment and academic performance of Chicano and Black students in a predominately White university. *The Urban Review*, 17, 3–23.
- Peng, S. S., & Fetters, W. B. (1978). Variables involved in withdrawal during the first two years of college: Preliminary findings from the national longitudinal study of the high school class of 1972. *American Educational Research Journal*, 15, 361–372.
- Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family background, parenting practices, and the Black-White test score gap. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap*. Washington DC: Brookings Institution Press.
- *Rea, D. G. (1992). Student characteristics, institutional characteristics, and undergraduate achievement: A study of Virginia Tech, 1985 to 1989. *Dissertation Abstracts International*, 53(07), 2196. (UMI No. 9233629).
- *Reed, H. B. (1944). Why do some colleges reach a higher level of achievement than others? *Journal of Educational Research*, 38, 161–172.
- *Reitzes, D. C., & Mutran, E. (1980). Significant others and self conceptions: Factors influencing educational expectations and academic performance. *Sociology of Education*, 53, 21–32.
- *Rimmer, J. D. (1973). The association between selected social, academic, and genetic variables, and college student psychiatric illness. *Dissertation Abstracts International*, 34(09), 6112. (UMI No. 7407062).
- *Rodgers, W. P. (1984). The influence of self-help programs on selected academic and demographic characteristics of freshmen students at the University of South Alabama. *Dissertation Abstracts International*, 44(10), 2994. (UMI No. 8401237).
- *Rodriguez, A. N. (1992). Predicting academic achievement and retention of Mexican-American and Anglo-American college students using traditional and nontraditional factors. *Dissertation Abstracts International*, 54(02), 466. (UMI No. 9309269).
- Roth, P. L., BeVier, C. A., Switzer, F. S., & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, 81, 548–556.
- Roth, P. L., & Clarke, R. L. (1998). Meta-analyzing the relation between grades and salary. *Journal of Vocational Behavior*, 53, 386–400.
- Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121, 297–317.
- Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2007). Socio-economic status and the relationship between the SAT and freshman GPA: An analysis of data from 41 colleges and universities. (Tech. Rep. No. 2007–5). New York: The College Board.
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85, 112–118.
- Saegert, S. C., Adler, N. E., Bullock, H. E., Cauce, A. M., Liu, W. M., & Wyche, K. F. (2007). *Report of the APA task force on socioeconomic status*. Washington, DC: American Psychological Association.
- Sander, R. (1995). *The 1995 National Study of Law School Performance*. Retrieved from <http://www.law.ucla.edu/sander/Systemic/Data.htm>
- *Scarr, S. (1979). *Heritability and educational policy: Genetic and environmental effects on IQ, aptitude, and achievement*. Paper presented at the annual meeting of the American Psychological Association, New York.
- Shadish, W. R. (1996). Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychological Methods*, 1, 47–65.
- *Smith, A. P. (2002). *Differences in performance of African American and White students at a large predominantly White institution as a function of progress towards degree completion*. Unpublished doctoral dissertation.
- *Spector, I. L. (1966). An analysis of certain characteristics and the educational success of junior college freshmen. *Dissertation Abstracts International*, 27(03), 640. (UMI No. 6609245).
- Stewart, E. B. (2006). Family and individual level predictors of academic success for African American students. *Journal of Black Studies*, 36, 597–621.
- *Suarez, O. (1981). Persistence of American Indian students at a comprehensive state university. *Dissertation Abstracts International*, 52(01), 47. (UMI No. 8113250).
- Taylor, M. E., Stafford, C. E., & Place, C. (1981). *National longitudinal study of the high school class of 1972 study reports update*. North Research Triangle Park, NC: Center for Educational Research and Evaluation.
- *Terenzini, P., Pascarella, E., & Lorang, W. (1982). An assessment of the academic and social influences on freshman year educational outcomes. *The Review of Higher Education*, 5, 86–110.
- *Thomas, G. E. (1977). *Race and sex effects on access to college*. Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. 143747).
- Thorndike, R. L. (1947). The prediction of intelligence at college entrance from earlier test. *Journal of Educational Psychology*, 38, 129–148.
- *Tilford, M. P., & Allen, D. E. (1974). Science and non-science majors in three predominantly Black colleges. *Journal of Negro Education*, 43, 117–126.
- *Tollett, J. R. (1980). Relationship of selected variables to measure of success in a teacher education program. *Dissertation Abstracts International*, 41(05), 2071. (UMI No. 8026630).

*Tuttle, R. M. (1981). *A path analytic model of the college going decision*. (ERIC Document Reproduction Service No. 224434).

U.S. Department of Education, National Center for Education Statistics. (1999). *National longitudinal study of the class of 1972* (ICPSR version). Ann Arbor, MI: Inter-University Consortium for Political and Social Research. (Original work published 1972).

U.S. Department of Education, National Center for Education Statistics. (2004). *National education longitudinal study: Base year through fourth follow-up, 1988–2000* (ICPSR version). Ann Arbor, MI: Inter-University Consortium for Political and Social Research. (Original work published 2002).

*Velez, W. (1985). Finishing college: The effects of college type. *Sociology of Education*, 58, 191–200.

Weiler, W. C. (1994). Transition from consideration of a college to the decision to apply. *Research in Higher Education*, 35, 631–646.

*Werts, C. E. (1967). *The study of college environments using path analysis* (Report No. NMSC-RR-VOI-3-MO-4–1967). Evanston, IL: National Merit Scholarship. (ERIC Document Reproduction Service No. 016301).

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91, 461–481.

Wightman, L. F. (1998). *The LSAC National Longitudinal Bar Passage Study*. Newtown, PA: Law School Admissions Council.

*Willard, J. S. (1993). The relationship of problem-solving ability to persistence/withdrawal behavior and to academic success at a private university. *Dissertation Abstracts International*, 54(09), 3201. (UMI No. 9404877).

Willingham, W. W., Lewis, C., Morgan, R., & Ramist, L. (1990). *Predict-*

ing college grades: An analysis of institutional trends over two decades. Princeton, NJ: Educational Testing Service.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, 39, 1–37.

Wood, P. K., Sher, K. J., Erickson, D. J., & DeBord, K. A. (1997). Predicting academic problems in college from freshman alcohol involvement. *Journal of Studies on Alcohol*, 58, 200–210.

*Wright, R. J., & Bean, A. G. (1974). The influence of socioeconomic status on the predictability of college performance. *Journal of Educational Measurement*, 11, 277–284.

*Young, J. W., & Fidler, J. L. (2000). Sex differences on the SAT: An analysis of demographic and educational variables. *Research in Higher Education*, 41, 401–416.

Zwick, R. (2002). Is the SAT a “wealth test”? *Phi Delta Kappan*, 84, 307–311.

Zwick, R. (2004). Is the SAT a “wealth test”? The link between educational achievement and socioeconomic status. In R. Zwick (Ed.), *Rethinking the SAT* (pp. 203–216). New York: Routledge Falmer.

Zwick, R., Brown, T., & Sklar, J. C. (2003). *California and the SAT: A reanalysis of University of California admissions data*. Berkeley, CA: Center for Studies in Higher Education.

Zwick, R., & Green, J. G. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, 44, 23–45.

Received July 9, 2007
 Revision received June 17, 2008
 Accepted June 17, 2008 ■

ORDER FORM

Start my 2009 subscription to *Psychological Bulletin*
 ISSN: 0033-2909

___ \$99.00	APA MEMBER/AFFILIATE	_____
___ \$249.00	INDIVIDUAL NONMEMBER	_____
___ \$650.00	INSTITUTION	_____
	<i>In DC add 5.75% / In MD add 6% sales tax</i>	_____
	TOTAL AMOUNT DUE	\$ _____

Subscription orders must be prepaid. Subscriptions are on a calendar year basis only. Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN
 PSYCHOLOGICAL
 ASSOCIATION

SEND THIS ORDER FORM TO
 American Psychological Association
 Subscriptions
 750 First Street, NE
 Washington, DC 20002-4242

Call **800-374-2721** or 202-336-5600
 Fax **202-336-5568** :TDD/TTY **202-336-6123**
 For subscription information,
 e-mail: subscriptions@apa.org

Check enclosed (make payable to APA)

Charge my: Visa MasterCard American Express

Cardholder Name _____

Card No. _____ Exp. Date _____

 Signature (Required for Charge)

Billing Address

Street _____

City _____ State _____ Zip _____

Daytime Phone _____

E-mail _____

Mail To

Name _____

Address _____

City _____ State _____ Zip _____

APA Member # _____

BULA09