The Role of Admissions Test Scores, Socioeconomic Status, and High School Grades in Predicting College Achievement

El rol de los exámenes de admisión, de las notas de educación secundaria y de la situación socioeconómica en la predicción del desempeño universitario

Rebecca Zwick

Educational Testing Service, USA

Abstract

Although test scores are widely used in college admissions in the United States, their use is the subject of ongoing debate, partly because of the association between test performance and socioeconomic status (SES). Although test critics have argued that this association is due to the particular content of admissions tests or to the differential availability of coaching, large socioeconomic effects are also found in assessments that are tied to school achievement and for which coaching is not available, such as the National Assessment of Educational Progress and in other academic measures. Some commentators have argued, however, that high school grade-point average has a smaller correlation with SES than admissions test scores and is therefore a superior admissions criterion. In this paper I examine the association between SES and test scores, as well as the association between SES and high school grades, and discuss the relevance of this complex web of associations to college admissions research. While the perennial finding that socioeconomic inequities manifest themselves as educational inequities is disheartening, the analysis of performance differences can point the way toward possible remedies.

Keywords: admissions, standardized tests, admissions testing, predictive validity

Resumen

Aunque el puntaje obtenido en pruebas estandarizadas es ampliamente utilizado en los procesos de admisión a instituciones de educación superior norteamericanas, su uso es objeto de debate actualmente, en parte debido a la asociación entre el desempeño en la prueba y el nivel socioeconómico (NSE). Aunque las críticas de las pruebas han sostenido que esta asociación se debe al contenido particular de las pruebas de admisión o a la disponibilidad diferencial de entrenamiento, también se han encontrado numerosos efectos socioeconómicos en las evaluaciones que están ligados al logro del colegio y por el cual el entrenamiento no está disponible, tal como la National Assessment of Educational Progress y otras medidas académicas. Sin embargo, algunos comentaristas han debatido que el promedio escolar de la enseñanza secundaria tiene una relación más pequeña con el NSE que los puntajes de las pruebas de admisión y es, por tanto, un criterio de admisión superior. En este artículo se examina el vínculo entre el NSE y los puntajes de las pruebas, así como también la asociación entre el NSE y las notas de la educación secundaria, y discute la relevancia de esta compleja red de asociaciones para la investigación de las admisiones a la educación superior. Mientras el descubrimiento permanente de que las inequidades socioeconómicas manifiestan en sí mismas como inequidades educacionales es decepcionante, el análisis de las diferencias de desempeño puede señalar el camino hacia posibles soluciones.

Palabras clave: admisión, pruebas estandarizadas, pruebas de admisión, validez predictiva

In the United States, most colleges and universities use standardized tests, either the SAT or the ACT, in admissions. A recent survey shows that ninety percent of responding institutions considered the test scores to be of "moderate" or "considerable" importance in the admission process (Clinedinst & Hawkins, 2010). But although these scores are viewed as useful predictors of college performance, the use of standardized admissions tests in the United States has always been controversial. One reason is the association between admissions test scores and socioeconomic status (SES).

Here I will consider the association of socioeconomic status with admissions test scores and with high school grades and, more generally, the role of socioeconomic factors in the prediction of college performance. Since the debate about admissions tests typically centers on the SAT rather that the ACT, I will focus on the SAT. The SAT is sponsored by The College Board and administered by Educational Testing Service. The SAT consists of three sections –Critical Reading, Math, and Writing– and is taken by almost three million students per year (College Board, n.d.).

The association between SAT scores and socioeconomic status

Test critics have often suggested that the SAT measures *only* socioeconomic status, to the exclusion of the academic skills it is claimed to assess. For example, Harvard Law Professor Lani Guinier stated that "the SAT should simply be called a 'wealth test'" (Guinier, 2000) and test critic Alfie Kohn claimed that "the verbal section of the SAT merely measures the size of students' houses" (Kohn, 2001, p. 2). Similarly, testing opponent Peter Sacks asserted that "one can make a good guess about a child's test scores by looking at how many degrees her parents have and what kind of car they drive" (Sacks, 1997, p. 27).

What do the data tell us about the association of SAT scores with SES? Table 1 shows an example from the 2011 SAT administration, based on nearly a million test takers who provided a response to a question about family income. Family income categories, expressed in thousands of dollars per year, are shown in the first column, followed by the percentage of test-takers in each category. The next three columns give the average scores corresponding to the income categories. The table shows that for all three sections of the SAT, scores increase with annual income. The penultimate row of the table shows the difference in average score between the highest and lowest income category, and the final row shows the difference in terms of standard deviation units. The three sections show a large and consistent SES effect, slightly exceeding one standard deviation. (I made an analogous computation using the two sections of the SAT that existed in 2001 and obtained a nearly identical result; see Zwick [2004, p. 204].)

Table 1
SES Effects for 2011 SAT (N=951,694)

Family income in thousands of dollars	Percentage of test-takers	Critical reading	Math	Writing
Over 200	7	568	586	567
140-200	9	541	555	533
80-140	27	521	535	511
20-80	44	483	496	472
0-20	13	434	460	429
Highest minus lowest (points)		134	126	138
Highest minus lowest (SD units)		1.2	1.1	1.2

Note: Computations are based on Table 11 from College Board (2011). The overall means for Critical Reading, Math, and Writing were 497, 514, and 489, respectively. The standard deviations were 114, 117, and 113. The source table includes finer income categories and also provides score data for those who did not respond to the income question.

The table clearly illustrates that SAT scores are associated with socioeconomic status. But what is the reason for this association? Two hypotheses that are often put forward are what I call the content and coaching hypotheses (Zwick, 2004). The content hypothesis, which many testing commentators in the United States have proposed, is that the test questions are not well tied to the high school curriculum and focus on material that is more familiar to students from wealthier families. Therefore, according to this line of reasoning, test-takers do not have equal opportunities to learn the material. Similar assertions have been used, both in the United States and Chile, to argue for using achievement tests, rather than aptitude tests in admissions (Koljatic, Silva, & Cofre, 2012). The coaching hypothesis says that coaching, or test preparation, is more likely to be available to wealthier test takers, creating an association between the test scores and socioeconomic status.

A pertinent question is whether we still see an association with SES for tests that are based on the school curriculum or for tests that do not have coaching or test preparation available. The National Assessment of Educational Progress (NAEP) provides an ideal example. NAEP is a U.S. Department of Education assessment program that, according to the National Center for Education Statistics, "provides results on subject-matter achievement, instructional experiences, and school environment for populations of students … NAEP does not provide scores for individual students or schools. NAEP results are based on representative samples of students" (National Center for Education Statistics¨ [NCES], n.d., p. 1). NAEP, then, allows for a test of both the content and coaching hypotheses because it is based on school achievement and involves samples of students who are not identified. Because NAEP does not produce student, classroom, or school-level scores, it is a low-stakes assessment for both students and schools and creates no incentive for coaching.

Nevertheless, large socioeconomic effects exist in NAEP. Table 2 shows results from the 2011 NAEP assessment. Fourth- and eighth-graders were assessed in Reading and Math, and eighth- and twelfth-graders were assessed in Writing. The SES indicator used in NAEP is eligibility for the National School Lunch Program, a federal government program that funds free or reduced-price lunches for students of low socioeconomic status. The table gives the overall number of students for each grade and subject, followed by the mean and percentage of students corresponding to the ineligible (higher SES) and eligible (lower SES) students. The rightmost columns show the difference in means between the higher and lower SES students in points and in standard deviation units. The differences range from 0.7 to 0.8 standard deviation units—again, a large SES difference. Because these assessments involve school-based content

and because there is no known coaching for NAEP, the results of table 2 are inconsistent with the content and coaching hypotheses.

In fact, all types of academic measures are found to be correlated with socioeconomic status in the

Table 2
2011 NAEP scores by eligibility for National School Lunch Program

Subject	Grade	N of students	Ineligible (higher SES)		Eligible (lower SES)		Difference in means	
			Mean	% of Students	Mean	% of Students	Points	SD Units
Reading	4	213,000	235	46	207	48	28	0.8
Reading	8	168,000	275	50	252	44	23	0.7
Math	4	209,000	252	46	229	48	23	0.8
Math	8	175,200	296	50	269	44	27	0.8
Writing	8	24,100	161	52	134	42	27	0.8
Writing	12	28,100	157	59	133	34	24	0.7

Note: Results were obtained through the NAEP Data Explorer tool on www.nces.ed.gov and from NAEP reports (National Center for Education Statistics (2011a, 2011b, 2012). Percentages of eligible and ineligible students add up to less than 100 because eligibility information was unavailable for some students.

US. Comparable associations hold for test scores at all ages, number of course credits completed, teacher rating of students, participation in extracurricular academic activities, and also grades (Zwick, 2004). In terms of test scores, similar patterns are found in Chile (Koljatic, Silva, & Cofre, 2012; Manzi, Flotts & Preiss, 2012) and elsewhere in the world. As the report of the 2009 Programme for International Student Assessment (PISA) notes, "many countries… show a strong relationship between the socio-economic background of students and their success at school and, in some of these countries, the disparities are magnified by large variations in the schools' socio-economic backgrounds…" (Organisation for Economic Co-operation and Development [OECD], 2010, p. 13). We know that educational quality varies with SES and so we should not be surprised to see these relationships.

Some testing commentators have claimed that high school grades are less highly correlated with socioeconomic status than are SAT scores and that, partly for this reason, high school grades are the best predictor of college grades. A prominent supporter of this point of view is Richard Atkinson, formerly the president of the University of California. During his tenure, he recommended that the SAT be eliminated as an admissions criterion at the university. This recommendation was not carried out, but it helped to spur a revision of the SAT, which took effect in 2005. According to Atkinson and his coauthor, Saul Geiser (Atkinson & Geiser 2009), "family income and parents´ education… are correlated both with SAT scores and also with college outcomes, so that much of the apparent predictive value of the SAT actually reflects the 'proxy' effects of socioeconomic status… High-school grades, on the other hand, are less closely associated with students´ socioeconomic background" (p. 3). Their discussion is in keeping with the belief, expressed in some recent critiques of testing (e.g., Geiser & Studley, 2004; Rothstein, 2004), that correlations between admissions test scores and subsequent grades are, to a substantial degree, an incidental result of the common influence of SES on admissions test scores and college grades. However, Sackett, Kuncel, Arneson, Cooper, and Waters (2009) found, based on a meta-analysis, that the association between SAT scores and college grades was virtually undiminished when SES was partialed out (or "controlled for"), and concluded that "the vast majority of the test-academic performance relationship was independent of SES" (Sackett et al., 2009, p. 1). Atkinson and Geiser (2009) then criticized the Sackett et al. study for failing to incorporate the effects of high school grade point average (HSGPA). Thus, debate on this subject is still ongoing.

In two studies conducted at the University of California, Santa Barbara, my colleagues and I examined the role of SES in the prediction of college performance. The College Board (2011) provided data for both studies and funded the later one. In one study (Zwick & Green, 2007), we used a multilevel analysis

to examine the question of whether HSGPA has a smaller association with SES than SAT scores do. The results showed that, within high schools, HSGPA and SAT have a fairly similar association with SES. A second study (Zwick & Himelfarb, 2011) explored the role of high school-level SES in predicting college success. Results indicated that high school-level SES can serve to reduce the ethnic-group prediction errors that typically occur when forecasting college performance. These studies are briefly described in the following sections.

The relationship of high school grades to socioeconomic status

My collaborator Jennifer Green and I used a nationally representative dataset consisting of nearly 100,000 students from more than 7,000 high schools to compare the correlation between SAT scores and family income with the correlation between HSGPA and income.

As we demonstrated, the usual way of computing these correlations can be misleading. Consider the correlation between SAT scores and family income. Assuming that the required data were available, the correlation between SAT and income could be calculated within each high school and the average of these correlations across high schools could then be obtained. This correlation has a clear interpretation: It is the average association between SAT scores and income *at the student level.* However, the typically reported correlation is not obtained in this way. Ordinarily, data from multiple high schools are combined and the correlation is computed without taking into account which high school each student attended. The size of these ordinary correlations is affected not only by the average *within-school* relationships between SAT scores and income, but also by the *between-school* relationship, which is the association between high school SAT score averages and high school income averages. We argued that the average within-high-school correlations are of primary interest because they represent the associations between variables at the student level.

Using a specially prepared data set from the College Board that linked students to their high schools of origin, we were able to obtain the average within-high school correlations of income with SAT scores and income with HSGPA. In essence, this analysis made use of a multilevel statistical model (Raudenbush & Bryk, 2002), which takes into account the unit of analysis. When we compared the SAT-income correlation to the HSGPA-income correlation using the ordinary computations, the SAT-income correlation appeared to be much higher. But when we used the average within-high school correlations, the two were much more similar.

For example, in the results for the more than 60,000 White students (which were most consistent with our hypothesis), the ordinary correlations of SAT with income were .21 and .24, respectively, for the verbal and math sections (the two sections of the SAT before 2005), while the correlation between HSGPA and income was only .10. By contrast, the average within-high-school correlations with income were .12 and .15 for the verbal and math sections, respectively, while the correlation between HSGPA and income remained at .10. These results suggest that at the student level, high school grades and SAT scores actually have a similar association with socioeconomic status. It is also relevant that high school grades are less precise than SAT scores. If standard adjustments for reliability were made, the correlations would be even more similar. Complete results can be found in Zwick and Green (2007). Our findings were replicated by researchers at the College Board in 2008 using a later year of SAT data (Mattern, Shaw, & Williams, 2008).

School-level SES and the prediction of college performance

My collaborator Igor Himelfarb and I investigated whether high school-level SES could help to explain a recurrent finding in the admissions testing literature. Research conducted over the last five decades has shown that African-American and Latino students tend to earn lower college grades than predicted when admissions test scores and HSGPA, combined, are used as predictors (Young, 2004). Various conjectures have been offered to explain this persistent finding, including measurement error in the predictors (Zwick, 2002, pp. 117-124; Zwick & Himelfarb, 2011). These explanations, however, leave open the question of whether these systematic prediction errors could be reduced by using a more complete regression model.
One clue comes from the finding that that the ethnic-group prediction errors are more severe when

HSGPA alone is used as a predictor than when SAT is included in the equation. A plausible explanation is that grading is less rigorous at inferior schools, so that using HSGPA alone as a predictor tends to yield misleadingly high predicted college grades for students from these schools. If African-American and Latino students are more likely than other students to attend lower-quality schools, this could explain the typical configuration of prediction errors.

Using data from 34 colleges, based on approximately 71,000 students from roughly 5,700 high schools, we investigated whether these systematic prediction errors could be reduced by the inclusion of a high-school-level SES index. Although adding more predictors to a regression model is guaranteed to increase the proportion of explained variance, R^2, it will not necessarily reduce the average prediction errors for particular groups.

We used regression analysis to predict first-year college grade point average using three different models: Model 1, advocated by some test critics, used only HSGPA as a predictor. Model 2, which is typically used in both institutional research and test validity studies, included HSGPA and SAT Critical Reading, Math, and Writing scores. Model 3 included HSGPA, SAT scores and an index of high school SES that we developed. Because of data limitations, our index was, by necessity, a simple one: a composite of average parent education for the high school and the poverty rate in the school district. All students who attended a particular school shared the same index value.

The results for Model 1 showed that African-American and Latino students did not perform as well as predicted. The average prediction errors were large: one quarter of a grade point for African-American students and one-fifth of a grade point for Latino students on a scale that ranged from zero to four. These were the same groups who were most likely to have attended low-SES high schools: Roughly 70% of African-American and Latino students had attended such schools. By contrast, fewer than half of Asian-American and White students, for whom prediction was very accurate, had attended such schools. For Model 2, which included both HSGPA and SAT scores, prediction errors were smaller: For African Americans and Latinos, average predicted grades exceeded average earned grades by .12 and .09, respectively. For Model 3, which included HSGPA, SAT, and the high school SES index, prediction errors were further reduced, reaching values of .09 and .03 for African Americans and Latinos, respectively.

It is notable that our simple high school SES index further reduced prediction errors even after HSGPA and SAT scores were included in the regression equation. However, in the US, "socio-economically disadvantaged schools also tend to be deprived in terms of basic resources, such as favourable student-staff ratios" (OECD, 2010, p. 101). We believe that on average, lower-SES high schools provide lower-quality education and that grading is less stringent in these schools. As a result, students from these schools do not perform as well in college as would be predicted from their high school grades and test scores.

Our research was intended to help explain observed patterns in prediction, rather than to develop admissions procedures. Blindly applying a model similar to Model 3 in admissions decisions would have the unfortunate effect of providing "affirmative action for high-SES children," as Rothstein (2004) noted in a similar context. However, it might be possible to use a high school SES index in combination with other criteria or to use such an index to identify high schools that need additional resources in order to improve.

Discussion

Socioeconomic status plays a complex role in the analysis of academic success. SES is related not only to scores on standardized admissions tests like the SAT, but to performance on achievement surveys like the National Assessment of Educational Progress and to other academic measures, including high school grades. Although some test critics have contended that high school grades have a smaller association with SES than do admissions test scores, recent findings from a national study suggest that within high schools, SAT scores and high school grades may be similarly associated with SES. Another study showed that differences in high school-level SES may help to explain patterns of systematic ethnic-group prediction errors typically observed in the US.

These research outcomes, as well as the 2011 SAT and NAEP results, are consistent with the recurrent finding that academic performance is influenced both by student-level SES and by school-level SES. The report of the 2009 PISA assessment notes that "although poor performance in school does not automatically follow from a disadvantaged socio-economic background, the socio-economic background of students and schools does appear to have a powerful influence on performance" (OECD, 2010, p. 13).

While the perennial finding that socioeconomic inequities manifest themselves as educational inequities is disheartening, the analysis of performance differences within and between countries can point the way toward possible remedies. Directions proposed by the Organisation for Economic Co-operation and Development (OECD, 2010, pp. 104-105) include focusing on low-performing schools and students, regardless of socioeconomic background; targeting disadvantaged children through special curricula, supplemental instructional resources, or economic assistance; and reducing between-school socioeconomic segregation. It is hoped that steps of this kind can increase student achievement and college readiness while also reducing their dependence on socioeconomic factors.

References

Atkinson, R. C., & Geiser, S. (2009). *Reflections on a century of college admissions tests.* Retrieved from http://cshe.berkeley.edu/

Clinedinst, M. E., & Hawkins, D. A. (2010). *2010 state of college admission.* Retrieved from http://www.nacacnet.org

College Board (n. d.). *SAT fact sheet.* Retrieved from http://press.collegeboard.org/sat/fact-sheet

College Board (2011). *College-bound seniors.* Retrieved from http://professionals.collegeboard.com/data-reports-research/sat/cb-seniors-2011

Geiser, S., & Studley, R. E. (2004). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. In R. Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions* (pp. 125-153). New York, NY: RoutledgeFalmer.

Guinier, L. (2000, September 29). *Lecture at University of California, Santa Barbara.*

Kohn, A. (2001, March 9). Two cheers for an end to the SAT. *Chronicle of Higher Education,* p. B12.

Koljatic, M, Silva, M., & Cofre, R. (2012). Achievement versus aptitude in college admissions: A cautionary note based on evidence from Chile. *International Journal of Educational Development, 33*(1), 106-115.

Manzi, J., Flotts, P., & Preiss, D. D. (2012). Design of a Collegel-Level Test of Written Communication. Theoretical and Methodological Challenges. In E. L. Grigorenko, E. Mambrino, & D. D. Preiss (Eds.), *Writing: A mosaic of new perspectives (*pp. 385-398). New York, NY: Psychology Press.

Mattern, K. D., Shaw, & Williams, F. E. (2008). *Examining the relationship* between *the SAT, high school measures of academic performance, and socioeconomic status: Turning our attention to the unit of analysis.* College Board Research Note No. RN-36. New York, NY: The College Board.

National Center for Education Statistics (n. d.). *NAEP overview.* Retrieved from http://nces.ed.gov/nationsreportcard/about/

National Center for Education Statistics (2011a). *The nation's report card: Mathematics 2011(NCES 2012–458).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics (2011b). *The nation's report card: Reading 2011(NCES 2012–457).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics (2012). *The nation's report card: Writing 2011(NCES 2012–470).* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Organisation for Economic Co-operation and Development (2010). *PISA 2009 results: Overcoming social background – equity in learning opportunities and outcomes* (Volume II). Retrieved from http://dx.doi.org/10.1787/9789264091504-en

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Rothstein, J. M. (2004). College performance and the SAT. *Journal of Econometrics, 121,* 297-317.

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

Sacks, P. (1997). Standardized testing: Meritocracy's crooked yardstick. *Change, 29,* 25-31.

Young, J. W. (2004). Differential validity and prediction: Race and sex differences in college admissions testing. In R. Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions* (pp. 289-301). New York, NY: RoutledgeFalmer.

Zwick, R. (2002). *Fair game? The use of standardized admissions tests in higher education.* New York, NY: RoutledgeFalmer.

Zwick, R. (2004). Is the SAT a "wealth test?" The link between educational achievement and socioeconomic status. In R. Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions* (pp. 203-216). New York, NY: RoutledgeFalmer.

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.

Zwick, R., & Himelfarb, I. (2011). The effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. *Journal of Educational Measurement, 48,* 101-121.