Using SATs or SOLs for School Performance Evaluation Report FAC/FCA-143

Frederick A. Costello
February 1, 2015

Introduction: We now have numerous types of tests that might be used to evaluate a school or school system: (1) the SAT I[1], (2) SAT 2 subject tests; (3) PSAT; (4) the ACT[2]; (5) Virginia Standard of Learning (SOL) scores[3]; (6) Advanced Placement (AP) exam scores; and (7) International Baccalaureate (IB) exam scores.

Of these tests, the two that are taken by the largest percentage of students are the SAT I and the required SOL tests. The SAT I generally is considered to test general ability and is taken by over 70% of FCPS high school seniors. SOL exams test how much students know about specific subjects, and many End of Course (EOC) SOL tests are taken by all or almost all FCPS students.

The purpose of the work reported herein was to determine whether the SAT I or the SOL's are the better measure of student performance

Summary: The mean SAT I and mean SOL test scores for 25 high schools in Fairfax County are closely correlated (Exhibit 1). The lowest correlation coefficient is 0.820, between the SAT math and the end-of-course SOL for Virginia and US History. The correlation coefficient is almost independent of whether SAT I math, SAT I reading, or SAT I writing is used. Most correlation coefficients are above 0.9.

	SAT read	SAT write	SAT math
ENGRead	0.941	0.937	0.940
ENGWrit	0.968	0.966	0.966
Math	0.937	0.932	0.953
VA & US History	0.835	0.822	0.820
World History I	0.843	0.832	0.843
World History II	0.849	0.841	0.854
Biology	0.908	0.905	0.912
Chemistry	0.923	0.919	0.931
Earth Science	0.853	0.844	0.833

Exhibit 1: Correlations between SAT and SOL Scores (2012-2013 school year)

If the SAT I largely measures student ability, as the College Board implies[4], then the correlation is not surprising. High-ability students should perform better on most standardized tests that measure some type of academic ability or academic achievement.

These results also suggest that instruction generally will not significantly change a student's measured ability score[5] or achievement score. If so, these results also imply that schools should focus on providing every student with instruction that maximize his/her potential, rather than attempting to equalize academic outcomes.

The question would then be how most cost-effectively to provide all students with instruction that maximizes their individual potential. In a system with limited resources and large class sizes, one obvious answer is to assign students to classes according to their academic accomplishments (a combination of ability and industriousness.) Teachers would have an easier job because they would be teaching a more academically homogeneous group.

---

[1] http://sat.collegeboard.org/home
[2] http://www.act.org/products/k-12-act-test/
[3] http://www.doe.virginia.gov/testing/
[4] http://professionals.collegeboard.com/testing/sat-reasoning/scores/compare
[5] Approximately 60 years ago, when conclusions could be drawn, a person's ability (IQ) was considered to be constant over the person's life. Now, in this age of skepticism, modern researchers think that ability (IQ), if it can even be measured, can be changed. Some suggest a 20-point change is possible, either increased or decreased.

Discussion: The correlation between mean SAT I scores and mean SOL end of course (EOC) scores was generated by gathering SAT I and SOL EOC scores for the 25 Fairfax County Public Schools (FCPS) high schools for the school year 2012-2013[6]. This school year was chosen because all of the data was readily available. We did not compare data for this school year with data for other school years.

EOC SOL tests are taken at various grade levels, depending on the subject. Thus, EOC SOL scores during 2012-13 are not all for the same group of students. We have implicitly assumed that the ability of the students at each school is the same, regardless of grade. EOC Algebra 1 exams generally are taken during high school only for students who have much difficulty with mathematics. In addition, the strongest math students do not take geometry EOC SOL exams in high school. Therefore, EOC Math SOL scores during high school oversample weak math students. Nonetheless, there is a high correlation between mean SAT I scores and EOC math SOL scores (Exhibit 2, showing also trend lines).

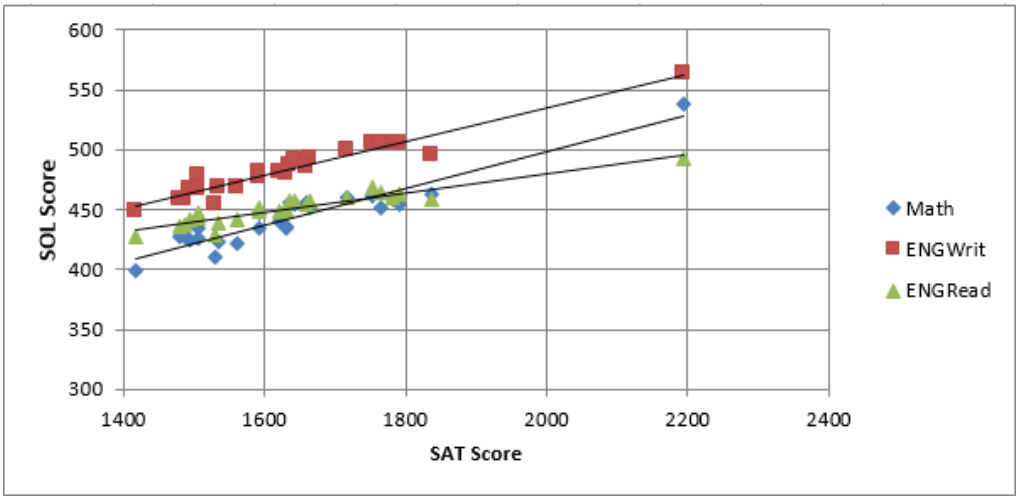

Exhibit 2: The Close Relationship between SOL Scores and SAT Scores

Note that, for a correlation coefficient of 0.820 (Exhibit 1), 82% of the changes (the variance) in the dependent variable are "explained" by the changes in the independent variable. The other 18% is attributed to random occurrences. A curve fit of the dependent variable as a function of the independent variable would be nearly linear when the correlation coefficient is this high (Exhibit 2).

The minimum correlation coefficient between mean high school SAT I subtest scores is 0.981 (SAT I Math vs. SAT I Writing). Although we do not have access to SAT scores for individual students, this result nonetheless implies that a student who does well on one SAT I subtest generally will do well on all three SAT I subtests. This high correlation indicates why the SAT I scores have almost the same correlation coefficient with any one of the SOL tests, as can be seen in Exhibit 1.

In our study, we noticed that mean SAT I scores increased by 15 points from 2000 to 2014[7] (Exhibit 3), so we used the data to determine if we could see the cause of the increase in the mean SAT I scores. We failed. The percent belonging to each ethnic group (Asian, Black, Hispanic, and White) and the total number of students varied almost linearly with year (Exhibit 3), so that the statistical analysis was thwarted. The percent of students who were classified as FRM (those receiving free-and-reduced-price meals) deviated most from linearity. The correlation coefficient between FRM seniors and SAT scores was high (0.854), but the correlation between FRM seniors and year was even higher (0.882). So we could draw no conclusion. The College Board states that, in comparing two students, a difference in SAT of 60 is considered insignificant (per reference in Footnote 4), so the 15-point increase in the average may also be insignificant.

[6] http://www.doe.virginia.gov/statistics_reports/research_data/index.shtml and http://www.fcps.edu/fs/budget/wabe/
[7] The 15-point increase in the SAT could be the result of a shift in the norm.

SOL tests are designed to measure knowledge, which is learned in school. The College Board does not state what SAT I measures, although it seems to measure ability because there is little material in the test that measures knowledge about specific subjects such as biology or history. Students with high ability should be expected to score higher in SOL tests if these students are as industrious as students with low ability[8]. Therefore, the strong correlation between SOL scores and SAT I scores is not surprising.

Both SAT I and the SOL tests measure student performance. How might this data be used?

One use would be in structuring the delivery of instruction, to make it easier for teachers to challenge but not frustrate all the students in their class, by creating more levels of academically homogeneous classes.

Although I have not done the research to confirm the following, I have been told by a parent who is highly involved in the FCPS that, in the FCPS, the trend for the past dozen years has been to move towards one or two levels of instruction, with more academically diverse classes. Elementary school "Level IV" classes now include students with much wider ranges of ability than the GT Center classes during the 1990s. Middle and high-school "honors" classes often have even wider ranges of student ability. Teachers point out that with 20 to 35 students in their classes, they do not have the time to provide differentiated instruction let alone useful feedback to all their students.

This testimony implies that students at all levels of achievement would benefit, and teacher workloads would be reduced, if students were placed in classes of nearly uniform achievement. The teacher could focus on helping all students in the room learn as much as possible, rather than on concentrating on students who might fail their SOL tests. In classes with students who will have difficulty passing their SOL tests, teachers could adapt their curriculum and teaching methods to maximize the amount learned by those students. (Students should be re-assessed each year so students would not be locked into a single achievement level for their entire school career.)

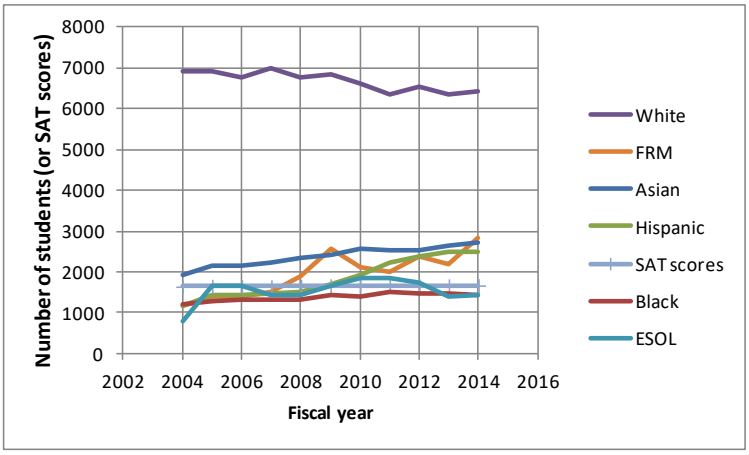

Exhibit 3: History of Countywide SAT I Scores and Student Demographics

[8] Some believe that every person has the same ability; however, this belief is not substantiated by data. Some of the people who so believe also believe, again without supporting data, that every person could be as good at basketball as was Michael Jordan if they worked as diligently at basketball as Michael Jordan did.