COLLEGE READINESS

ISSUES IN COLLEGE READINESS

Using ACT Data as Part of a State Accountability System

Introduction

Accountability in the nation's public schools has become a matter of considerable urgency since the 2001 amendment of the federal Elementary and Secondary Education Act (ESEA) of 1965.[1] Under this most recent version of the ESEA, as of the 2007–08 academic year states and districts must assess the mathematics and reading proficiency of students each year in grades 3 to 8 and at least one year in grades 10 to 12, and must also assess the science proficiency of students at least one year in grades 3 to 5, at least one year in grades 6 to 9, and at least one year in grades 10 to 12. By means of these assessments, each school must demonstrate "adequate yearly progress" (AYP) toward the goal of achieving proficiency of all its students in all subjects by the year 2014, although the definitions of and standards for proficiency can vary substantially by state.

The data derived from subject-area assessments and used by schools to demonstrate progress are referred to as *accountability measures*. Depending on the system a school chooses to use to indicate progress, its accountability measure will typically be one of four types: status, improvement, growth, or value-added. Simply put, status measures are indicators of a school's current degree of proficiency; improvement and growth measures are predictions of a school's degree of progress in the future; and value-added measures show the contribution a school makes to its students' academic achievement over and above the contribution made at a typical school.

This brief summarizes a study in which data from one or more of ACT's three college readiness assessments—EXPLORE® (grade 8 or 9), PLAN® (grade 10), and the ACT® test (grade 11 or 12)—were used to generate all four types of accountability measure. (See the sidebar on p. 2 for more information about these assessments.) The study was based on a sample of 73,240 students in 1,019 cohorts at 485 high schools across the United States who graduated from high school between 2002 and 2006. At least half the students in each cohort had taken all three assessments.[2]

For the status, improvement, and growth measures, we defined proficiency in a subject as meeting or exceeding a particular score on each EXPLORE, PLAN,

[1] The 2001 amendment is also known popularly as the No Child Left Behind Act.

[2] Because most schools in the U.S. whose students take all three of ACT's college readiness assessments are located in the Midwest and Southwest, the 485 schools in this sample were also located primarily in these regions and are therefore not representative of all U.S. schools. Nevertheless, the 73,240 students in the sample were typical of students nationally in terms of their average performance on the three assessments.

IC 050801100

EXPLORE, PLAN, and the ACT

ACT's College Readiness System includes a sequence of three longitudinal college readiness assessments: EXPLORE, PLAN, and the ACT. Based on extensive research into what postsecondary educators expect from entering college students, each assessment measures what students are able to do with what they have learned in school.

EXPLORE, for students in grade 8 or 9, provides baseline information on the academic preparation of students that can be used to plan high school coursework. PLAN, for students in grade 10, provides a midpoint review of students' progress toward their education and career goals while there is still time to make necessary interventions. The ACT, for students in grade 11 or 12, measures students' academic readiness to make successful transitions to college and work after high school.

ACT's College Readiness Benchmarks

The College Readiness Benchmarks for the ACT represent the level of achievement required for students to have a high probability of success in selected credit-bearing first-year college courses. Success is defined as approximately a 75 percent chance that a student will earn a grade of C or better, or approximately a 50 percent chance that a student will earn a grade of B or better. The courses are the ones most commonly taken by first-year college students in the areas of English (English Composition), mathematics (College Algebra), social sciences (History, Psychology, Sociology, Political Science, and Economics), and natural sciences (Biology). Data from 98 postsecondary institutions and more than 90,000 first-year college students were used to establish the Benchmarks, which are median course placement scores that reflect a high probability of student success in a college course.

Based on these Benchmarks, ACT has also established College Readiness Benchmarks for EXPLORE and PLAN. These scores indicate whether students, based on their performance on EXPLORE or PLAN, are on target to be ready for first-year college-level work by the time they graduate from high school.

or ACT subject test. These scores are known collectively as ACT's College Readiness Benchmarks (see sidebar).

In this brief, we explain and illustrate each of the four accountability measures generated from EXPLORE, PLAN, or ACT data. We also evaluate the appropriateness of each type of measure in assessing the academic effects of schools on student learning, by examining its associations with factors that are outside the school's control.

The study shows that data from EXPLORE, PLAN, and the ACT can be used as the basis for the four common accountability measures, three of which are consistent with the current requirements of the ESEA. The study also found that the measures consistent with the ESEA—status, improvement, and growth—also run a greater risk of inappropriately sanctioning schools with certain characteristics than does the fourth, value-added measure.

For clarity and simplicity of illustration, examples of each type of accountability measure presented in the following sections are all based on the same data from a single high school in the sample. This school, which we have given the fictitious name Eastville High School, has a high poverty rate (50 percent of the students are eligible for free or reduced lunch) and a high concentration of racial/ethnic minority students (50 percent). Eastville has two cohorts in the sample: 128 students who graduated from high school in 2003, of whom 64 had each taken EXPLORE, PLAN, and the ACT; and 175 students who graduated from high school in 2006, of whom 93 had each taken EXPLORE, PLAN, and the ACT. Status (as of 10th grade), improvement (projected status of 10th graders in 2013–14), growth (as of 10th grade), and value-added (as of 11th or 12th grade) measures are presented, along with their respective percentile ranks. For the status, growth, and value-added measures, the percentile ranks give the position of the 2006 Eastville cohort among the 1,019 cohorts in the study; for the improvement measures, which require multiple cohorts within a school, the percentile ranks give Eastville's position among the 272 high schools that had more than one cohort in the sample.

1. Status Measure: Percentages of 10th-Grade Students on Target to Be Ready for College-Level Work by High School Graduation

A status measure is an aggregated assessment result from a single school year, so called because it is one indicator of the *status* of the school during that year.

To construct status measures for proficiency in the four subject areas at Eastville High School, we used the College Readiness Benchmarks for PLAN as cutoff scores. The proficiency rates and percentile ranks for 10th-grade Eastville High School students in each subject in academic year 2003–04 are shown in Table 1.

Table 1: Status Measures for 10th-Grade Students at Eastville High School (2003–04)

Subject	% Proficient	Percentile Rank
English	84	54
Mathematics	31	32
Reading	65	67
Science	17	32

Table 1 shows that 84 percent of the Eastville High School 2006 cohort were on target in 2003–04 to be ready for first-year college English coursework by the time they graduated from high school, and that the percentages of these students who were similarly on target in Mathematics, Reading, and Science were 31 percent, 65 percent, and 17 percent, respectively. With respect to students at other high schools, the percentile ranks show that the Eastville cohort's status was about average in English (percentile rank of 54 among all the cohorts in the sample), slightly above average in Reading (percentile rank of 67), and below average in Mathematics and Science (percentile ranks of 32).

Appropriateness of Status Measures as Indicators of School Effectiveness

In the context of the ESEA, status measures serve as snapshots of a particular school's progress in a given year. And because one of the principles of the ESEA is to "set expectations for annual achievement based on meeting grade-level proficiency, not on student background or school characteristics" (U.S. Department of Education, 2007), schools with very different populations or characteristics are, under the requirements of the ESEA, being held to the same standard. For example, a school with a high proportion of students for whom English is not their native language would be expected to perform as well as a school whose students are all native English speakers.

Critics of status measures argue that such measures are unsound as a measure of school effectiveness because they necessarily reflect contextual factors that are beyond a school's control, such as students' entering achievement level or poverty level. And in fact, numerous studies have shown that aggregate school achievement is strongly related to school poverty level and proportion of racial/ethnic minority students (e.g., Howley, Strange, & Bickel, 2000; Linn, 2001). Therefore, it is not surprising that school sanctions resulting from status measures are frequently perceived as less fair than other measures.

Our examination of status measures based on EXPLORE, PLAN, or ACT data shows that poverty level and proportion of racial/ethnic minority students are the two characteristics most strongly related to the student population's initial status (i.e., 8th-grade proficiency as measured by attainment of the College Readiness Benchmarks for EXPLORE), with poverty level having the strongest association. So schools with higher poverty levels and, to a lesser extent, higher proportions of minority students are more likely to have lower initial status. Further, students' initial status is the strongest predictor of their later status in 10th and 12th grade (i.e., proficiency as measured by attainment of the College Readiness Benchmarks for PLAN and the ACT, respectively[3]). This suggests

[3] Performance on EXPLORE may be compared directly with performance on PLAN and the ACT because scores on the three assessments are reported using the same scale.

that high-poverty and high-minority schools such as Eastville High School would have a higher likelihood of being sanctioned inappropriately under an accountability system based on status measures.

2. Improvement Measure: Predicted Percentages of 2013–14 10th-Grade Students on Target to Be Ready for College-Level Work by High School Graduation

An improvement measure is a predicted assessment result for students in a particular grade based on the results in that grade from prior school years, so called because it is one indicator of the *improvement* a school is expected to show by the future year. Essentially, improvement measures use two or more past status measures to predict future status.

To construct improvement measures for proficiency in the four subject areas at Eastville High School, we again used the College Readiness Benchmarks for PLAN as cutoff scores; proficiency rates from 2000–01 and 2003–04 were extrapolated to 2013–14. The predicted proficiency rates and percentile ranks for 10th-grade Eastville High School students in each subject in academic year 2013–14 are shown in Table 2.

Table 2: Predicted Status of Eastville High School 10th-Grade Students in 2013–14, Based on 10th-Grade Results from 2000–01 and 2003–04

Subject	% Proficient	Percentile Rank
English	82	51
Mathematics	0	29
Reading	50	38
Science	2	33

Improvement measures are commonly employed to determine AYP under the ESEA, using full proficiency in all subjects by 2013–14 as the goal. Because the ESEA requires 100 percent proficiency among all students in all subjects by 2014, it is clear from Table 2 that Eastville High School is not making AYP with respect to its 10th-grade population.[4] And while the percentile ranks show that Eastville High School is below average in three of the four subject areas when compared to the other 271 schools in our sample that had multiple cohorts, only 9 of the 272 schools—just 3 percent—were projected to reach full proficiency in all four subject areas at grade 10. Thus, Eastville is not unlike the overwhelming majority of schools in the sample.

[4] The zero percent proficiency rate in mathematics is based on the fact that the mathematics proficiency of Eastville High School 10th graders decreased from 2000–01 to 2003–04, to such an extent that extrapolating this decrease ten years into the future led to a negative result. As the percentile rank shows, this result occurred for 29 percent of the schools in the sample.

In the previous section we saw that, when using status measures, initial student achievement levels at grade 8 are influenced by a school's poverty level and its proportion of racial/ethnic minority students, and that this influence suggests that accountability systems based on status measures are likely to sanction schools inappropriately. Because improvement measures are derived from two or more past status measures, it would stand to reason that improvement measures are also likely to sanction schools less fairly than other measures. Our examination of improvement measures shows that this is indeed the case.

The results show that poverty level and proportion of racial/ethnic minority students are related to the student population's initial status in grade 8: schools with higher poverty levels and higher proportions of racial/ethnic minority students are more likely to have lower initial status. And students' initial status influences their later status in 10th and 12th grade. As with status measures, this suggests that high-poverty and high-minority schools such as Eastville High School would have a higher likelihood of being sanctioned inappropriately under an accountability system based on improvement measures.

3. Growth Measure: Percentages of 12th-Grade Students Expected to Be Ready for College-Level Work by High School Graduation, Based on Their Performance in Grades 8 and 10

A growth measure is a predicted assessment result for students in a future school year based on a set of results for those same students in two or more prior school years, so called because it accounts for the *growth* these students made during the prior school years. Growth measures differ from improvement measures in that growth measures predict results for a set of students based on their own past performance, whereas improvement measures predict results for a school at a particular grade based on the performance of its students in that same grade in the past.

To construct growth measures for proficiency in the four subject areas among a group of students at Eastville High School, we predicted their proficiency at grade 12 based on their assessment results at grades 8 and 10. The predicted proficiency rates and percentile ranks for 12th-grade Eastville High School students in each subject in academic year 2005–06[5] are shown in Table 3.

[5] Constructing growth measures for students in academic year 2013–14 would require EXPLORE results from 2009–10 and PLAN results from 2011–12. Interestingly, the predicted 2005–06 results were highly correlated with the actual results of the students in the Eastville High School 2006 cohort (correlations of 0.85 for English and Mathematics, 0.80 for Reading, and 0.79 for Science).

Table 3: Predicted College Readiness of Students in Grade 12 (2005–06), Based on Their Degree of Progress toward College Readiness in Grades 8 (2001–02) and 10 (2003–04)

Subject	% Proficient	Percentile Rank
English	80	68
Mathematics	23	32
Reading	54	56
Science	14	40

Under the U.S. Department of Education's Growth Model Pilot Program, growth measures have been approved for some states as an alternative to improvement measures in demonstrating AYP. The percentile ranks in Table 3 show that the predicted proficiency of 12th-grade students at Eastville High School during 2005–06 was above average in English and Reading and below average in Mathematics and Science. However, if the ESEA had required 100 percent proficiency in all subjects by 2005–06, Eastville would not have met this requirement in any of the four subjects.

Appropriateness of Growth Measures as Indicators of School Effectiveness

Our examination of growth measures shows that, as with status and improvement measures, school poverty level and proportion of racial/ethnic minority students are related to the student population's initial status at grade 8, and that students' initial status influences their later status in 10th and 12th grade. This again suggests that high-poverty and high-minority schools such as Eastville High School would have a higher likelihood of being sanctioned inappropriately under an accountability system based on growth measures.

4. Value-Added Measure: School Contribution to Student ACT Score

A value-added measure is an estimate of the number of score points by which attending a particular school increases students' test scores in a given subject (beyond the average expected increase associated with attendance at a typical school), based on two or more years of assessment results at the school for these students—so called because it indicates the *value added* by a school to its students' academic achievement.

To construct value-added measures for Eastville High School we used a method that estimates the school's effect on its students' ACT scores, explicitly controlling for student factors (annual family income, race/ethnicity, and incoming performance level as represented by the same students' EXPLORE scores in grade 8) and school factors (enrollment, proportion of students tested, poverty level, and proportion of racial/ethnic minority students). The estimated school effect shown by this measure can be interpreted as an estimate of the school's contribution to students' academic performance, adjusted for their incoming performance level, students' family income and race/ethnicity, school size, proportion of students tested, school poverty level, and proportion of racial/ethnic minority students.

Table 4 shows the value-added results for Eastville High School. Because the average number of score points attributed to a typical school in the sample is set at zero, the figures in the second column represent not absolute score points but the *differences* between the average number of score points attributed to Eastville and those attributed to a typical high school in the sample.

Table 4: Estimates of School Effects on Students' ACT Scores (2005–06)

Subject	Effect on ACT Score (Deviation from Typical School)	Percentile Rank
English	+ 1.2	92
Mathematics	– 0.3	32
Reading	+ 0.9	95
Science	+ 0.4	83

Currently, value-added measures are not accepted as a means of determining AYP under the ESEA. This is because, as mentioned earlier, one of the core principles of the ESEA is that expectations for annual achievement must be based on whether or not a school is meeting grade-level proficiency; value-added measures focus on the change in achievement level (either positive or negative) from one grade to another without regard for whether such change—even a positive change—results in greater proficiency. The value-added results suggest that Eastville High School has an above-average effect on its students' English, Reading, and Science scores, and a below-average effect on their Mathematics score. The percentile ranks indicate that Eastville's effects on English, Reading, and Science scores are in fact greatly above average—92nd, 95th, and 83rd percentiles, respectively.

Value-added measures may in fact be the most informative of the four types of accountability measure. The fundamental purpose of value-added measures is to isolate and estimate the effects of teachers, schools, or academic programs. Because status, improvement, and growth measures do not account for students' entering academic proficiency or contextual factors such as student and school poverty level, policymakers have expressed interest in value-added measures as a means to measure school and teacher effectiveness for high-stakes accountability (i.e., as the basis for rewards or sanctions) and low-stakes accountability (i.e., to improve practice or identify teachers' and schools' strengths and weaknesses).

There is disagreement about the extent to which value-added measures truly measure a school's effectiveness. But value-added models can, at a minimum, be used to produce descriptors of school effectiveness that are more meaningful than those produced by status, improvement, or growth measures. If we had considered only Eastville High School's status, improvement, or growth measures, we might have concluded that its performance was average or perhaps slightly below average. But by considering the value that the school added to its students' academic achievement, we were able to see that Eastville actually has shown above-average effects in all subject areas except mathematics. This information should lead the school to study its mathematics curriculum and devise a strategy for improvement. Further, Eastville can take pride in its strong effect on English, Reading, and Science scores and can identify the teachers and practices that have contributed to this success.

Appropriateness of Value-Added Measures as Indicators of School Effectiveness

In contrast to the other three accountability measures, the value-added measures are by definition unrelated to school or student characteristics such as enrollment, proportion of students tested, poverty level, and proportion of racial/ethnic minority students. Our examination of value-added measures supports the contention that they are the least likely of the four accountability measures to inappropriately sanction high-poverty, high-minority schools such as Eastville High School.

Conclusion

This study illustrates that data from EXPLORE, PLAN, and the ACT can be used as the basis for state accountability systems. However, the study also found that the accountability measures currently accepted under the ESEA run a risk of sanctioning high-poverty, high-minority schools inappropriately.

Maximizing student representation is a crucial element of any accountability system. If data are not available for a significant portion of students in a school, there could be concern that the resulting accountability measures are not an accurate reflection of the school's effects. Therefore, ACT recommends that the standard errors of accountability measures should be reported, especially when the measures are used as the basis for rewarding or leveling sanctions against a school.

Our research highlights how accountability measures can lead to different conclusions about school effectiveness. Clearly, status, improvement, and growth measures can be heavily influenced by factors outside of the school's control: specifically, the entering achievement level and socioeconomic status of the students who attend the school. For this reason, ACT recommends that policymakers preparing to reauthorize the ESEA consider accepting value-added measures as another means of determining AYP. By using value-added measures, the school's effect on its students' academic achievement is better isolated and measured.

References

Howley, C. B., Strange, M., & Bickel, R. (2000). *Research about school size and school performance in impoverished communities* (ERIC Digest). Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Document Reproduction Service No. ED 448 968).

Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems*. CSE Technical Report (CSE_TR-539). Los Angeles: Center for Research and Evaluation, Standards, and Student Testing.

U.S. Department of Education. (2007). *No Child Left Behind. Growth models: Ensuring grade level proficiency for all students by 2014*. Retrieved July 17, 2009, from www.ed.gov/admins/lead/account/growthmodel/proficiency.pdf

500 ACT Drive P.O. Box 168 Iowa City, Iowa 52243-0168 319/337-1000 www.act.org