

Feasibility of the Use of the ACT and SAT in Lieu of Florida Statewide Assessments

Volume 1: Final Report



Assessment Solutions Group

**Ed Roeber
John Olson
Barry Topol**

January 1, 2018

Executive Summary

In the 2017 Florida legislative session, HB 7069 was passed and signed into law on June 15, 2017. The legislation states, in part:

“The Commissioner of Education shall contract for an independent study to determine whether the SAT and ACT may be administered in lieu of the grade 10 statewide, standardized ELA assessment and the Algebra 1 end-of-course assessment for high school students consistent with federal requirements under 20 U.S.C. s. 6311(b)(2)(H). The commissioner shall submit a report containing the results of such review and any recommendations to the Governor, the President of the Senate, the Speaker of the House of Representatives, and the State Board of Education by January 1, 2018.”

The Florida Department of Education (FDOE) issued RFP 2018-48 to solicit vendors to independently conduct studies, research, analyses, and Florida educator and expert meetings, and produce a final report to this effect. Assessment Solutions Group (ASG) and its team of subcontractors – Wisconsin Center for Education Products & Services (WCEPS), University of Minnesota’s National Center on Educational Outcomes (NCEO), and University of Kansas’s Center for Assessment and Accountability Research and Design (CAARD) – were selected to carry out the following studies:

1. Alignment – Evaluate the degree to which the ACT and the SAT align with Florida content standards and are, therefore, suitable for use in lieu of the grade 10 statewide, standardized English Language Arts (ELA) assessment and the Algebra 1 End-of-Course (EOC) assessment (Florida Standards Assessments, FSA).
2. Comparability – Conduct studies, research, and analyses required to determine the extent to which ACT and SAT test results provide comparable, valid, and reliable data on student achievement as compared to the Florida statewide assessments for all students and for each subgroup of students.
3. Accommodations – Determine whether ACT and SAT provide testing accommodations that permit students with disabilities and English learners the opportunity to participate in each assessment and receive comparable benefits.
4. Accountability – Conduct analyses to determine whether ACT and SAT provide unbiased, rational, and consistent differentiation among schools within the state’s accountability system.
5. Peer Review – Conduct evaluations to determine whether the ACT and SAT meet the criteria for technical quality that all statewide assessments must meet for federal assessment peer review.

Results and Conclusions

1. Alignment (Criteria 1 and 2)

- a. Algebra 1 EOC – The analysis focused on the degree to which the assessments were aligned with the 45 Florida Algebra 1 content standards, which are a subset of the high school Mathematics Florida Standards (MAFS). Based on the results of the test forms analyzed, neither the SAT nor the ACT assessment is fully aligned to the Florida Algebra 1 standards. Both the ACT and SAT assessments would need to be augmented to assess the full breadth and depth of the Algebra 1 standards as called for by federal regulations. The analysis indicated that the ACT would need slight adjustment to attain the minimum level of full alignment: seven or eight items would need to be added to each ACT test form. The SAT test forms were found to have conditional alignment, depending on the test form: One test form was found to be acceptably aligned, needing four items added; and the other SAT test form was found to need slight adjustment, needing seven items added to the form to attain the minimum level of full alignment according to the criteria used in this study.
- b. Grade 10 English Language Arts – The results show that the ACT would need major adjustments – needing 10 or more items revised or replaced to be fully aligned with the Florida Grade 10 LAFS. The SAT test forms were found to have conditional alignment, depending on the test form: One SAT test form was found to need five items revised or replaced for full alignment while the other SAT test form was found to need slight adjustments – seven items revised or replaced – for full alignment.

While augmenting the ACT or SAT to gain an acceptable level of alignment is certainly possible, augmentation adds cost and complexity to the administration of the tests, since items used to augment a test need to be developed annually and administered separately from the college entrance test. Without such augmentation, the ACT and possibly SAT tests might not meet the United States Department of Education (USED) peer review criteria for aligned tests, thus jeopardizing the federal approval of Florida’s plan to offer choice of high school tests to its school districts.

2. Comparability (Criterion 3)

There were two significant concerns with the dataset available for analysis (students who took both the FSA and either the ACT or SAT) in the comparability studies.

- a. First, only about half of Florida’s tenth or eleventh grade students take either the ACT or the SAT before they graduate from high school, meaning that the matched samples of students who participated in the FSA assessments and took one or both of the college entrance tests seriously underrepresent the full sample of FSA test-takers (presumably the lower-scoring students that only took the FSAs).
- b. Second, the data provided by the FDOE indicates that 83% of students took the Algebra 1 EOC in eighth or ninth grades, one to three years before they took the ACT or SAT (in spring of tenth grade, fall of eleventh grade, and/or spring of eleventh grade). Large

distances between time of testing in the two tests increases measurement error as learning likely occurred between those test administration times.

In short, the data provided for this study are neither representative of the full population of students nor were the tests taken close enough in time to assume little to no learning occurred. Several different statistical analyses were run to try to account for the data issues and determine if the tests are comparable. An important analysis performed was how often students would be placed at the same level of performance based on their scores on the three tests. The results of this classification consistency analysis indicate that many students would be placed at different performance levels on the three tests, some by as much as four out of the five performance levels. Thus, districts using the FSA option may have very different results than districts using either the ACT or SAT options. This casts serious doubt on the interchangeability of the three tests, and the soundness of making accountability decisions based on them. At this point, it appears the ACT and SAT do not produce results comparable to the FSA and should not be considered alternatives to them. This also indicates that the ACT and SAT will, likely, not meet USED peer review requirements.

3. Accommodations (Criterion 4)

This study concluded that in many ways, in terms of the provision of accommodations, the ACT and SAT could provide comparable benefit to the FSA for purposes of school accountability and graduation, although this was less evident on the SAT for English Learners (ELs). In general, both the ACT and the College Board indicated that they would provide greater numbers of the accommodations in the standard list of accommodations used in this study (previously developed by NCEO) than were provided for the FSAs. Whether these differences were appropriate for the Florida standards was not addressed in these studies.

Comparability in the process for accommodations requests was less clear and often relevant more to the use of the tests for college entrance; comparability to the FSA cannot be judged here because the FSA does not provide a score that can be used for college entrance. Still, if a district based a decision to use one of these tests in lieu of the Florida assessments on the possibility of having college entrance scores for all its students, this goal is unlikely to be realized for some students with disabilities and ELs.

The lack of transparency in the decision-making process about which specific accommodations would result in a college reportable score for which specific students is likely to result in non-comparability for some student groups compared to other student groups, which could be a concern when making the decision about whether to allow Florida districts to use either the ACT or the SAT in lieu of the Florida assessments.

4. Accountability (Criterion 5)

Using the sample of students with two years of data from the FSAs and an ACT or SAT score, simulated schools were created to examine the effects of calculating school-level indicators using the different tests.

Overall, differences are shown across all three indicators. The results show that the numbers going into the accountability determination would differ for many schools by the test selected.

Richer calculations can be done for ELA as there exists a state test for grades 8, 9, and 10. For mathematics, the time at which the Algebra 1 EOC test is taken varies by students and for many, there is no prior year score with which to base a growth calculation.

There are two important findings to consider from this accountability study: one data-based and one more theoretical. First, the differences shown for ELA vary by type of school. Larger schools with a greater number of lower performing students are advantaged by using the alternate tests (ACT/SAT). This finding has implications for policy, as districts could use these results to select a test, rather than making a more holistic determination about its students and what test best fits the population.

Second, there will often be very different students being compared in the growth models. For example, in mathematics, the learning gain using the FSA will be calculated based on grade 8 math and grade 9 Algebra 1. However, using the alternate test, a similar high school would be evaluated based on the learning gain between Algebra 1 EOC in grade 10 and the ACT or SAT in grade 11. Likewise, for the value added model, only schools using the FSAs will have a VAM score for ELA and only some of those for mathematics. With the elimination of the FSA Grade 10 ELA and Algebra 1 EOC tests, two years of prior data will not exist for students taking the ACT or SAT in grade 11.

Both of these findings indicate that the answer to the question on fairness is: “no – it is not fair to compare schools that use the state tests in their accountability system to those that use the alternate tests.”

5. Peer Review (Criterion 6)

To test the acceptability of Florida’s plan to offer its schools the option of using the ACT or SAT in lieu of the FSAs, ASG conducted a mock peer review, using evidence provided by ACT, the College Board, and the Florida Department of Education, as well as from ASG’s studies of alignment, comparability, and accommodations. Experienced peer reviewers examined the evidence and prepared written notes similar to an actual peer review. A summary of the peer review results is shown in Table 1.

Table 1. Peer Review Critical Element Determinations by College Entrance Test

Peer Review Determinations	ACT	College Board
Met Mock Peer Review Requirements	23	20
May Not Meet Mock Peer Review Requirements	1	6
Did Not Meet Mock Peer Review Requirements	6	3
TOTAL	30	29*

*One Critical Element - related to online assessment - is not applicable to the current College Board SAT.

As can be seen, the ACT was judged to not meet 6 of 30 Critical Elements, and possibly not meet the peer review requirement for one additional element. The SAT was judged to not meet 3 of 29 peer review Critical Elements, and possibly not meet the peer review requirements for six additional elements.

Overall Conclusion

It is the opinion of ASG and its partners that due to the alignment, comparability, and accountability system issues associated with the ACT and SAT tests, allowing districts to pick which of the three tests to administer to its students is not appropriate and likely will not meet federal ESSA peer review requirements.

Detailed results from the five studies and a summary of the findings are provided in the following sections of the report and in the Summary and Conclusions chapter at the end. Details of the data and analyses used in the studies are provided in a separate document “Volume II - Appendices.”

Background and Purpose of the Project

In the 2017 Florida legislative session, HB 7069 was passed and signed into law on June 15, 2017. The legislation states, in part:

“The Commissioner of Education shall contract for an independent study to determine whether the SAT and ACT may be administered in lieu of the grade 10 statewide, standardized ELA assessment and the Algebra 1 EOC assessment for high school students consistent with federal requirements under 20 U.S.C. s. 6311(b)(2)(H). The Commissioner shall submit a report containing the results of such review and any recommendations to the Governor, the President of the Senate, the Speaker of the House of Representatives, and the State Board of Education by January 1, 2018.”

The federal law referenced above is commonly known as the Every Student Succeeds Act (ESSA). The law provides flexibility for a state to approve a school district to administer, in lieu of the statewide high school assessment, a “locally selected,” “nationally recognized” high school academic assessment that has been approved for use by the state, including submission for the U.S. Department of Education’s (USED) assessment peer review process. At a minimum, ESSA requires that the state must determine that an assessment used for this purpose meets the following criteria:

1. Is aligned to and addresses the breadth and depth of the State’s content standards;
2. Is equivalent in its content coverage, difficulty, and quality to the statewide assessments;
3. Provides comparable, valid, and reliable data on student achievement as compared to the Florida Standards Assessments (FSA), the statewide assessments used with all students and each subgroup of students. Final ESSA assessment regulations of December 8, 2016, clarify that comparability between a locally selected, nationally recognized high school academic assessment and the statewide assessment is expected at each academic achievement level;
4. Provides accommodations that permit students with disabilities and English learners the opportunity to participate in the assessment and receive comparable benefits;
5. Provides unbiased, rational, and consistent differentiation among schools within the state’s accountability system; and
6. Meets the criteria for technical quality that all statewide assessments must meet (i.e., those specified by USED’s assessment peer review).

The purpose of the project was to conduct a study to determine whether the state of Florida can allow districts to choose to offer its students the FSAs, ACT, and SAT and still have an assessment system that meets the above criteria.

ASG Partners and Plan of Study

In September 2017, Assessment Solutions Group (ASG) and its subcontractors were awarded the contract under Florida RFP 2018-48: *Feasibility of Use of ACT and SAT in Lieu of Statewide Assessments*. The ASG Team consisted of some of the most recognized names in their respective fields of statewide assessment and their expertise aligned perfectly with the six criteria that Florida outlined in the RFP. The ASG team and their assigned areas of study were as follows. The ASG individual assigned to coordinate the work is shown in parentheses for Criteria 1-5.

Criteria	Responsible Organization/(ASG Support Person)
1 and 2 - Alignment	WCEPS - Norman Webb/Sara Christopherson (Olson)
3 - Comparability	CAARD - Marianne Perie (Roeber, Olson)
4 - Accommodations	NCEO - Sheryl Lazarus/Martha Thurlow (Roeber)
5 - Accountability	CAARD - Marianne Perie (Olson)
6 - Peer Review	ASG - Ed Roeber/John Olson (Thurlow, Lazarus, Perie, Webb, and Christopherson)
ASG Team Members	John Olson Edward Roeber Barry Topol

The overall work by ASG and its partners was performed collaboratively with multiple groups working together to evaluate and respond to each of the criteria outlined in the RFP. There was a lead group and ASG support for each criterion, thereby allowing for multiple points of review and expertise to be brought to each area. The overall project direction and management was provided by ASG’s Barry Topol, with technical support provided by ASG’s John Olson and Ed Roeber.

The plan of study for each criterion is outlined below.

Criteria 1 and 2 - Alignment (Wisconsin Center for Education Products and Services)

The proposed methodology for the alignment analysis was based on processes developed and refined by Norman Webb over the past 20 years. These processes have been used to analyze curriculum standards and assessments in around 30 states to satisfy or to prepare to satisfy the Title I compliance as required by the United States Department of Education (USED). Many states that have met the USED requirements used this process to evaluate the alignment of their standards and assessments. The alignment analysis conducted in this study was designed to answer two key research questions:

1. How does content coverage in the ACT and SAT, for both mathematics and English language arts (ELA), compare with the content coverage of the current Florida grade 10 statewide standardized ELA assessment and the Algebra 1 EOC assessment for students?
2. What is the degree of alignment of the ACT and the SAT with the high school Language Arts Florida Standards (LAFS) and the mathematics Florida Standards (MAFS) with regards to the satisfying the federal requirements within the Every Student Succeeds Act (ESSA)?

The proposed alignment analyses involved two stages:

- **Stage I:** An analysis of assessment framework documents
- **Stage II:** An in-person content alignment institute

The final deliverables for the work outlined were two alignment reports, one for mathematics and one for ELA. The two reports provided findings from a reliable and replicable process using an item-level analysis. Results described the overlap in content targeted by Florida's FSAs and the ACT and SAT, as well as any content unique to each assessment, considering refined content topics and levels of content complexity. In addition, the results described the degree of alignment of the ACT and the SAT with the Florida standards for both language arts and mathematics.

The two final reports on alignment included a comparison of the blueprints and item specifications, or equivalent documents, for each of the ACT, SAT, and Florida assessments to provide more in-depth description of the content coverage by each assessment.

Criterion 3 – Comparability (Center for Assessment and Accountability Research and Design)

The main goal of the comparability analyses was to determine if the technical characteristics of the different measures under consideration (ACT and SAT; FSAs in Grade 10 ELA and Algebra 1 EOC) would provide comparable, valid, and reliable data on student achievement, for all students and each subgroup of students, to permit them to be used in lieu of the current FSAs in Florida's assessment system.

In carrying out this comparability work, the following research questions and associated analyses were addressed:

1. How similar are item types across the three tests? Item types were evaluated to determine whether the tests were comparable in terms of the types of items used.
2. How similar are the ranges of difficulty across the three tests at the item level? Average and range of item difficulties of the tests were examined to see if they were significantly different from one another.
3. How similar are the reliabilities across the three tests? Technical information for each assessment (e.g., standard errors of measurement and internal consistency) and the overall reliability of each measure were reviewed and compared.
4. When looking at matched samples, how similar are the distributions of performance across the tests? Comparisons were made of the distribution of test takers by county school system with comparable demographic information about student enrollments and, where necessary, sub-samples of students to best represent the entire state were selected. Once

matched/representative samples were obtained (adjusted if necessary for student subgroup underrepresentation), the performances of students on the different tests were examined.

5. What percentage of students can be expected to be categorized into different achievement levels with each test (and in which direction)? Comparisons were completed on whether the same percentages of students perform in different achievement levels on each test.
6. What is the probability that a student would be placed in the same Florida category of achievement when taking the FL test, the ACT, or the SAT? Analyses were completed to evaluate the comparability of whether every student would be placed in the same category of achievement on the ACT or the SAT as they are on the Florida tests.

To complete these studies, data were collected from recent ACT and College Board Technical Reports, test files, matched datafiles from Florida, other state data (those using ACT or SAT), and state reports of student performance and achievement level results. The intent was to build a body of evidence to be used to evaluate the degree of comparability of the tests.

Criterion 4 – Accommodations (National Center on Education Outcomes)

The National Center on Educational Outcomes (NCEO) organized and conducted in-person studies with Florida educators to evaluate the degree to which the ACT and SAT provide testing accommodations that permit students with disabilities and English learners (ELs) the opportunity to participate in each assessment and receive comparable benefits to participation in the FSA.

NCEO staff worked with the FDOE to identify 8 panelists who had familiarity with the Florida mathematics standards and the ways in which the FSA Algebra 1 EOC assessment is administered to students with disabilities and ELs. Similarly, NCEO also worked with the FDOE to identify 8 panelists who had familiarity with the Florida language arts standards and the ways in which the FSA grade 10 ELA assessment is administered to students with disabilities and ELs.

Each panel included a special educator (Exceptional Student Education); English language learner educator (English for Speakers of Other Languages); blind/low-vision educator; deaf/hard of hearing educator; and content educator. Each group of panelists went through a systematic study process during the in-person meetings and reviewed accommodations for the ACT, SAT, and FSAs with respect to:

1. The process and ease of signing up for accommodations
2. The availability of the accommodations themselves
3. The testing context for the accommodations provided to reach its conclusion about the suitability of the ACT or SAT to replace the current Florida assessments of grade 10 ELA and Algebra 1.

Criterion 5 – Accountability (Center for Assessment and Accountability Research and Design)

The accountability simulations called for in Criterion 5 were in many ways the “bottom line” of establishing the comparability of the college entrance tests with the FSAs in ELA and Algebra 1 EOC because of the consequences for schools and districts for the performance of their

secondary students. Florida includes ELA, math, science, and social studies scores in its school accountability system. In this study, the ACT or SAT scores substituted for the Grade 10 FSA ELA scores or the FSA Algebra 1 EOC scores. Other achievement scores and all other variables included in the Florida accountability system remained the same. In Florida's A-F accountability rating system, the ACT or SAT scores were included in the two indicators for high schools – Achievement and Learning Gains.

Using an adjusted matched sample, CAARD replicated the procedures used by Florida to hold schools accountable for the performance of their students in Florida's school accountability system. This included the substitution of the ACT- or the SAT-derived student scores for the FSA in ELA or the Algebra 1 EOC scores.

A series of simulations were conducted using adjusted matched samples of FSA and ACT or SAT student files. Results from the simulations were then carefully evaluated and the findings from using ACT and SAT were compared to those for FSA ELA and Algebra 1 EOC tests that are currently used in the accountability system reports for Florida schools. These simulations determined whether ACT and SAT provide unbiased, rational, and consistent differentiation among schools in Florida's accountability system.

Criterion 6 – Peer Review (Assessment Solutions Group)

The ASG team's approach under Criterion 6 included the following steps:

1. Summarizing current information and evidence from states that have used the ACT and the SAT as their high school assessments on how they have addressed the requirements in the USED peer review, in particular those Critical Elements related to alignment, test development, accommodations, technical quality, and validity. When this research had concluded, a determination was made that information for the College Board SAT was not yet available, while information for the ACT preceded ESSA and its peer review elements.
2. Creating a hybrid peer review template for use by the FDOE, ACT, and the College Board to submit their evidence of adequately addressing each peer review critical element.
3. Reviewing and commenting on each of the pieces of evidence submitted by ACT and the College Board in support of the use of these assessments in lieu of the FSAs.
4. Providing a professional judgment on the likelihood of ACT and/or SAT, when used as *an optional high school test* in place of the state's test, being approved by USED following peer review. This included providing comments on the strengths and weaknesses of the evidence that was provided and recommendations on the areas where improvements or additional evidence may be needed. Note: Not all evidence for peer review will be provided by ACT or the College Board. There are peer review critical elements for which evidence in support of the use of the ACT or the SAT will come from school districts after the initial administration of these assessments in Florida's districts. Collecting this "local-use" data will add to the complexity of FDOE's ultimate peer review submission to the USED.
5. Preparing the relevant parts of the actual peer review document for submission for the Department.

ASG and its partners used a "peer review-like" process to accomplish the purposes outlined above. The ASG team gathered evidence from the FDOE, ACT, and the College Board,

including the draft of the pertinent sections of the actual *State Assessment Peer Review Submission Cover Sheet and Index Template*, the document that each state uses to submit its evidence of the technical qualities of its proposed assessments. The peer review evidence compilation was split into two tracks: 1) the technical criteria for the Critical Elements responded to by ACT or the College Board, and 2) evidence related to supporting these Critical Elements from the work that ASG and its partners carried out in Criteria 1-5.

Sections outlining the analyses conducted and conclusions reached for each of the six criteria appear in the following sections of the report. Appendices providing detailed data and other information appear in a separate volume of the report.