



Council of Language Arts Supervisors

September 28, 2018

Vince Verges
Assistant Deputy Commissioner
Accountability, Research, and Measurement



FLORIDA DEPARTMENT OF
EDUCATION
fldoe.org

www.FLDOE.org

Topics

- Scoring of Florida's statewide assessments
- Released test timeline updates
- Spring 2019 reporting
- Concordant score updates
- SAT 10 updates
- FLKRS – updated resources
- State leadership and legislative timeline updates

Scoring of Florida's Statewide Assessments

- Measurement – Psychometrics is a type of measurement
- Item Response Theory – IRT
- “Pattern Scoring,” rather than raw scores
 - Pattern scoring is used across the country and around the world, including in statewide and national assessments, professional licensure exams, and international assessments

Measurement Can Be General or Specialized

Measurement	Psychometrics
Assigns numbers to objects or events	Assigns numbers to psychological characteristics
Examples: hurricanes, earthquakes, time, stock market, height, weight	Examples: achievement, personality, IQ, opinion, interests

Psychometrics of Scoring Florida's Tests

Item Response Theory – 3 Parameters (“Traits”)

a. Item discrimination values

- How well items differentiate among students of differing ability levels

b. Item difficulty values

c. Guessing (pseudo-guessing) values

- Test items constructed specifically to limit examinees ability to get the answer correct simply by guessing
- Not all items include this trait (e.g., responses to writing prompts)

Different Methods of Scoring

Pattern Scoring	Number-Correct Scoring
<ul style="list-style-type: none">• Complex Mathematics	<ul style="list-style-type: none">• Simple Mathematics
<ul style="list-style-type: none">• Maximum likelihood estimates, e.g., Item statistics, student's answer pattern, standard error of measurement (SEM)	<ul style="list-style-type: none">• Raw scores (# of points): mean, standard deviation (SD), SEM, % correct
<ul style="list-style-type: none">• “Theta” scale (mean=0, standard dev=1)	<ul style="list-style-type: none">• Number-right scale
<ul style="list-style-type: none">• Score conversions (e.g., scale scores, percentile ranks, etc.)	<ul style="list-style-type: none">• Score conversions (e.g., scale scores, percentile ranks, etc.)

Comparison: Number Correct and Pattern Scoring

Similarities

- The relationship of derived scores is the same
- For example:
 - High correlation of number-right scores and scale scores (we'll refer to this later!)
 - Scale score has the same percentile rank for both methods

Differences

- Methods of deriving scores
- The number of scale scores possible
 - Number right = limited to the number of items
 - IRT = unlimited, or is limited by the scale (e.g., 425–575)

Choosing the Scoring Method

- Which model?
- Simple vs. complex?
- Best estimates?
- Advantages/disadvantages?
 - Example: How can the same number correct get different scale scores?

Advantages of IRT and Pattern Scoring

- Better estimates of an examinee's ability
 - The score that is most likely, given the student's responses to the questions on the test (maximum likelihood scoring)
- More information about students and items are used
- More reliability than number right scoring
- Less measurement error (SEM)

Disadvantages of IRT and Pattern Scoring

- Technical: Complex Mathematics
 - Difficult to understand
 - Difficult to explain
- Not common in classroom tests: “Not like my experience!”
- Sometimes perceived as ‘hocus pocus’:
“Why is the writing score such a mystery?!”

Examples

Effect of Item Difficulty			
Item No.	Parameter		
	a	b	c
1	0.0150	250.000	0.1
2	0.0150	275.000	0.1
3	0.0150	300.000	0.1
4	0.0150	325.000	0.1
5	0.0150	350.000	0.1

Response Patterns (1=correct)		
Pattern (Items 1-5)	SEM	Scale Score
11100	43	300
01110	43	305
10101	43	305
10011	43	310

Answering more difficult items (b-parameter) can result in higher scores.

Examples (cont.)

Effect of Item Difficulty			
Item No.	Parameter		
	a	b	c
1	0.0050	300.00	0.2
2	0.0100	300.00	0.2
3	0.0150	300.00	0.2
4	0.0200	300.00	0.2
5	0.0250	300.00	0.2

Response Patterns (1=correct)		
Pattern (12345)	SEM	Scale Score
11001	94	280
11100	61	270
01110	46	300
00111	39	305

Correctly answering more discriminating items (a-parameter) can result in higher scores, even if the items are all of the same difficulty. There is also less measurement error for the more discriminating items

“Weight” of the Writing Prompt

- In scoring, the writing responses are treated like three separate test items, much like in the preceding examples
- The item parameters associated with the responses each year are different from prior years, so their impact on scoring is also slightly different each year
- **Reminder: High correlation between raw score and scale score**
 - Students who score high on writing are much more likely to score Level 3 and above.
 - Students who score no more than 0, 1, or 2 are *extremely* likely to score only in Level 1 or 2.



Released Tests

Released Tests – Timeline of Activities

- Currently ongoing: Review released test interpretive products (IP) from other states (CA, NY, TX, OH, VA). We are also reviewing IP for previously-released FCAT tests.
- November: Review IP findings with Florida's Technical Advisory Committee; gather feedback on appropriateness of each with respect to Florida's current tests
- Spring 2019: Content Advisory Committees convene to gather further IP recommendations.

Released Tests – Timeline of Activities

- Summer 2019-Spring 2020: Finalize format of released test IP
- Summer 2020: Select operational test forms for use in spring 2021; begin parallel production of released test forms and accompanying IP
- June 30, 2021: Report all scores; release test forms in Grades 3 and 10 ELA (including Grade 10 Writing prompt), Algebra 1, and middle grades ELA (w/prompt) and Mathematics (TBD), science (TBD) and social studies (TBD)
- July 2021: Districts receive individual student score reports, which will include individual writing responses in released grades only.



Spring 2019 Reporting

FSA and NGSSS Spring Reporting

In accordance with House Bill (HB) 7069, the following reporting deadlines are specified in [1008.22\(7\)\(a\)](#), F.S.:

“Assessment results for the statewide, standardized [English Language Arts] ELA and mathematics assessments and all statewide, standardized [End-of-Course] EOC assessments must be made available no later than **June 30**, except for results for the grade 3 statewide, standardized ELA assessment, which must be made available no later than **May 31**.”

Please keep these dates in mind as you communicate with students, teachers, parents, military recruiters, colleges, universities, and any other stakeholders throughout the year and into the spring.



Concordant/Comparative Scores

Concordant/Comparative Scores

- New concordant and comparative scores for the Grade 10 FSA ELA and FSA Algebra 1 EOC assessments became effective on June 28, 2018.
- The [revised rule language](#) establishes new concordant and comparative scores beginning with students who enter grade 9 in the 2018–2019 school year (i.e., those scheduled to graduate from high school in the spring of 2022).
- The previous concordant and comparative scores remain in effect for students who were in high school prior to 2018–2019 (i.e., those scheduled to graduate in either the spring of 2018, 2019, 2020, or 2021). These students are also permitted to use the new concordant scores, if those new scores satisfy their assessment graduation requirements.
- **The scores available for students to use if they cannot pass the statewide assessments are not changed for students enrolled in high school prior to this school year.**

Concordant/Comparative Scores (cont.)

- PERT will continue to be available for dual enrollment or college placement purposes. The vendor, McCann, has committed to maintaining the \$0.94 unit price for the foreseeable future.
- The rule language and an FAQ is posted on the Standard Setting page of the FDOE website at <http://www.fldoe.org/accountability/assessments/k-12-student-assessment/stard-setting.stml>.
- ACT and College Board recently released their own concordance tables. The tests that were chosen for concordance match those chosen for FSA, i.e. SAT EBRW to ACT Reading and English. Concordance scores reported for SAT and ACT are very similar to those adopted by the State Board in June. The ACT/SAT Concordance Guide is here: <https://collegereadiness.collegeboard.org/pdf/guide-2018-act-sat-concordance.pdf>.



SAT10 Update

Stanford Achievement Test, Tenth Edition (SAT 10) Norms Update

- As previously mentioned, norms for the SAT 10 assessment have been updated.
- No SAT 10 test content has changed, nor has the State Board Rule changed with respect to the minimum percentile rank required for good-cause promotion.
- The effect of the norms update is such that students will need to attain a slightly higher scale score than previously in order to place in the 45th percentile.
- Beginning this school year, districts must use the updated norms for good-cause promotion decisions.



FLORIDA DEPARTMENT OF
EDUCATION
fldoe.org

FLKRS

FLKRS Administration: July 9–October 12, 2018

- The FLKRS window is currently open.
- Schools are required to test all kindergarten students within the first 30 days of each school's respective start dates.
- In August, new guidance was posted regarding appropriate use of sub-domain scores.
- More information is available at <http://www.fldoe.org/accountability/assessments/k-12-student-assessment/flkrs>.



State Leadership and Legislative Timelines

State Leadership & Legislative Timelines

- November 6th - Election Day: Impacts on state leadership roles and policies under the authority of the new Governor can only be determined following the election.
- December 2018-February 2019 (tentative): Interim committee meetings
- January 25th, 2019: Deadline for bills to be filed
- March 5th, 2019: Begin legislative session
- May 3rd, 2019: End legislative session (unless extended, or special session is called)

Questions?

Vince Verges

Vince.Verges@fldoe.org