

Making Sense of the New Summative Assessments

Deb Sigman

San Francisco, CA | December 10, 2015



THE CENTER ON
**STANDARDS &
ASSESSMENT
IMPLEMENTATION**

WestEd  CRESST

This document is produced by the The Center on Standards and Assessment Implementation (CSAI). CSAI, a collaboration between WestEd and CRESST, provides state education agencies (SEAs) and Regional Comprehensive Centers (RCCs) with research support, technical assistance, tools, and other resources to help inform decisions about standards, assessment, and accountability. Visit www.csai-online.org for more information.

This document was produced under prime award #S283B050022A between the U.S. Department of Education and WestEd. The findings and opinions expressed herein are those of the author(s) and do not reflect the positions or policies of the U.S. Department of Education.

Topics for Discussion

1. What About Those Scores
2. Ever-Changing Assessment Landscape
3. Making the Most of the Summative Data
4. Aligning Policies with Best Assessment Practices

Now that Scores are Here

- It may not be about a decline...
- However, it's definitely about
 - different standards
 - different expectations
 - different instructional strategies
 - ensuring our students are prepared for post-secondary success
- Reinforce the differences

View in the Current Context

- Peer Review Guidance
- 2% Suggestion
- Every Student Succeeds Act
 - Regulatory Process
 - Fully enacted 2017-18
 - State authority
- Multiple measures
- Parent opt-outs
- Testing overload

Opportunity for Getting it Right

- “Too often, testing is treated narrowly, rather than as a flexible tool to obtain information about important questions.”

U.S. Congress, Office of Technology Assessment, Testing in American Schools: Asking the Right Questions, OTA-SET-519 (Washington, DC: U.S. Government Printing Office, February 1992)

Purpose –Use –Response

“... it is not clear that some of the tests that school districts administer were designed for the purposes for which they are used. The most controversial example is the use of state summative exams to evaluate school district staff when most of these tests were designed to track district and school progress, not individual staff-member proficiency.”

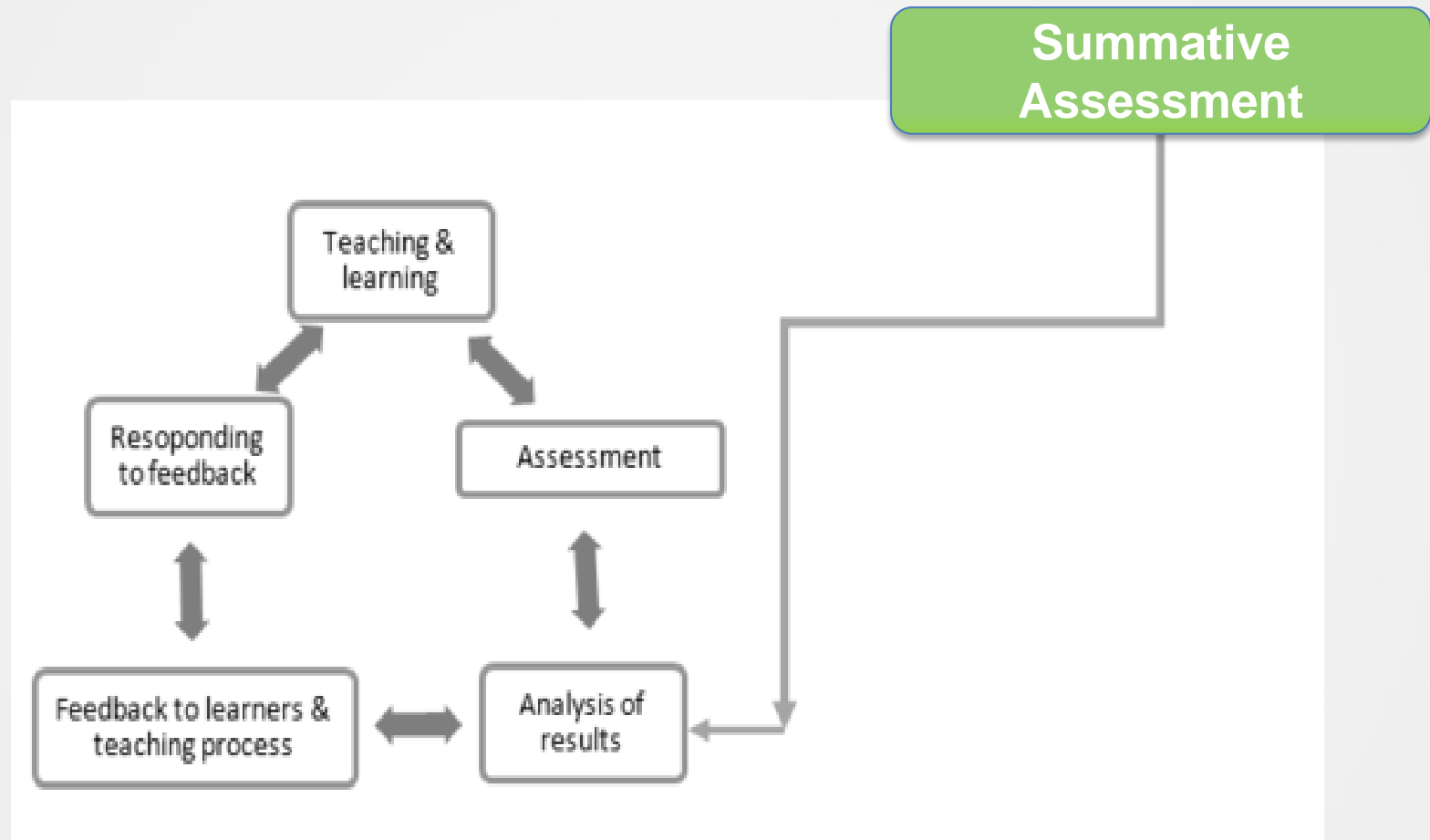
Student Testing in America's Great City Schools: An Inventory and Preliminary Analysis, October 2015

Moving Forward – Every Student Succeeds Act

“...while states are still required to test students annually...they have a freer hand in designing those tests...civil right groups have been tepid in their support for the legislation because they fear that some states will revert to the neglect of minority students that drove Congress to pass No Child Left Behind...”

Commentary from the New York Times, David L. Kirp, Professor of Public Policy at the University of California, Berkeley, December 10, 2015

Making the Most of Results



Making the Most of Summative Results

- Rarely provide definitive answers, but raise many questions that can facilitate reflection on practice and programs
- Provide a general information/direction and a necessary story
- Results are necessary, but not sufficient
- Not the reason for everything
- Remember the purpose
- Most useful at the aggregate (large groups)

The graphic consists of two overlapping blue speech bubbles. The top bubble is circular and contains the text "Let's". The bottom bubble is rectangular with a tail pointing towards the top bubble and contains the text "Talk".

Let's Talk

- ✓ Think about assessment issues that might be facing the states that you serve?
- ✓ How would you help them think about their new assessment results as they are responding to constituent concerns.

Ways to Measure



Different assessments – different scores – different scales

- Not all assessments are created equally
- Beware of too much generality
- Know the technical aspects
- Scales represent different performances



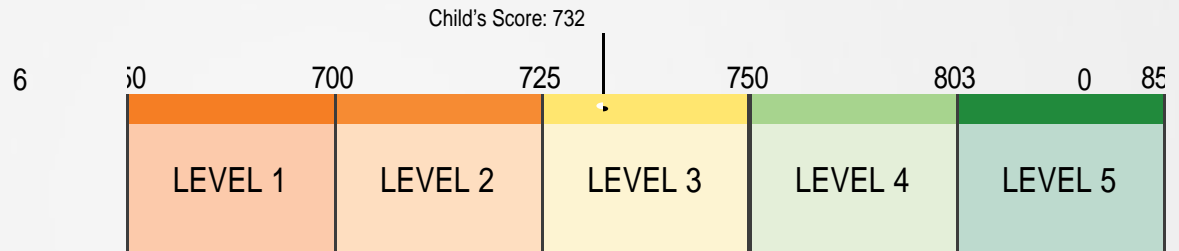
New Jersey – PARCC

MATHEMATICS PERFORMANCE

Level 3

Your child performed at Level 3
and earned a score of **732**

Students performing at levels 4 and 5
met or exceeded expectations.
*For a description of each performance
level, see page 2.*



School average	District average	State average	PARCC average
761	759	758	747

ADDITIONAL INFORMATION ABOUT YOUR CHILD'S MATHEMATICS SCORE


Florida - FCAT



**Reading Retake Student Report
Spring 2015
Grade 11**

APPLEBEE, CARMINE
ID 9901234567
School 9999 - SUNSHINE MIDDLE SCHOOL
District 99 - OCEAN

1



**Reading Retake Student Report
Spring 2015
Grade 11**

APPLEBEE, CARMINE
ID 9901234567
School 9999 - SUNSHINE MIDDLE SCHOOL
District 99 - OCEAN

Page 2

This report shows your student's results from the Spring 2015 FCAT 2.0 Reading Retake assessment. Passing Grade 10 FCAT 2.0 Reading is a requirement for earning a Florida high school diploma for a specific cohort of students. All students who entered grade 9 in the 2009-10 school year or before must earn a score of 241 or above, which is equivalent to a Grade 10 FCAT Reading Score of 300 (FCAT Reading Developmental Scale Score of 1926). All students who entered grade 9 in the 2010-11 school year through the 2012-13 school year must earn a score of 245 or above.

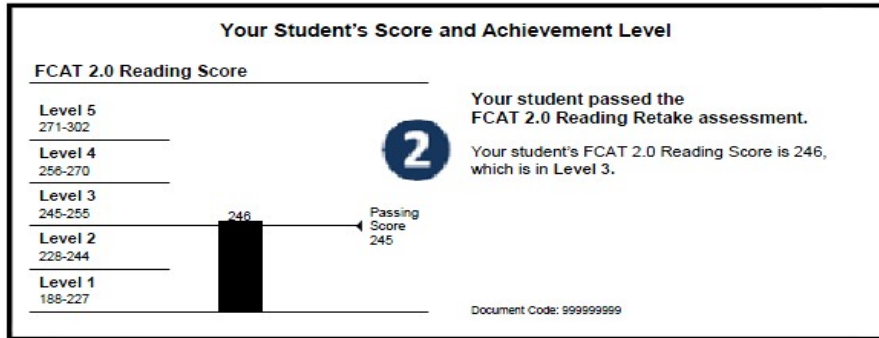
The FCAT 2.0 measures your student's performance on selected benchmarks in Reading as defined by the Next Generation Sunshine State Standards (NGSSS).

Your Student's Reading Content Area Scores

Your student passed the FCAT 2.0 Reading Retake assessment.

The table below gives more specific information about your student's performance in the content areas tested in Reading.

Content Areas	Points Earned	Points Possible
Vocabulary – Uses multiple strategies to determine the meaning of grade-appropriate vocabulary words.	3	7
Reading Application – Uses a variety of strategies to comprehend text suitable for the grade level.	7	14
Literary Analysis: Fiction and Nonfiction – Identifies, analyzes, and applies knowledge of the elements of a variety of literary texts, both fiction and nonfiction.	5	11
Informational Text and Research Process – Comprehends and interprets informational text from a variety of sources.	12	13



Achievement Levels

Five categories describe the success students have with the content tested on FCAT 2.0 Reading.

- Level 5 – Indicates mastery of the most challenging content of the NGSSS.
- Level 4 – Indicates an above satisfactory level of success with the challenging content of the NGSSS.
- Level 3 – Indicates a satisfactory level of success with the challenging content of the NGSSS.
- Level 2 – Indicates a below satisfactory level of success with the challenging content of the NGSSS.
- Level 1 – Indicates an inadequate level of success with the challenging content of the NGSSS.

3

4

Additional information about the FCAT 2.0 is available from your student's school, or at www.fldoe.org, the Florida Department of Education's website.

Florida's official online student advising system, www.flvc.org, is provided free to students by the Florida Department of Education. With this system, your student can start thinking about careers, receive online academic advising to evaluate college and career readiness, learn about higher education opportunities in Florida, and more.

rmddy 20000001-99999-0000000

Copyright © 2015 State of Florida, Department of State

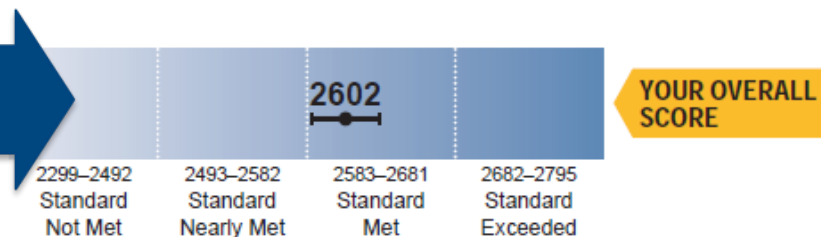
California – Smarter Balanced

The **Scale Score** is the student's score on the test. Scores can be between 2000 and 3000; the range varies by grade.

Chen's Results on California's Assessments

ENGLISH LANGUAGE ARTS/LITERACY

Chen's overall score is: **2602**

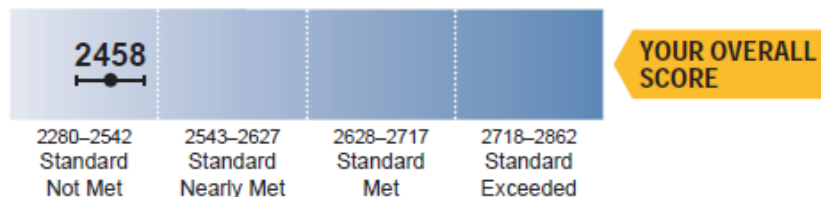


Chen met the achievement standard and demonstrated the knowledge and skills in English language arts/literacy needed for success in entry-level credit-bearing college coursework after high school.

Chen's performance on the four areas that comprise this overall score can be seen on the back of this report.

MATHEMATICS

Chen's overall score is: **2458**



Chen did not meet the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in mathematics needed for success in entry-level credit-bearing college coursework after high school.

Chen's performance on the three areas that comprise this overall score can be seen on the back of this report.

Examination of Assessment Data

- Comprehensive, complex, difficult process
- Every state/district has its own political and policy context
- Multiple ways to look at data
 - Different displays can elicit different responses
 - Use of data to inform policy
- Needs of state policy makers, advocacy groups, parents, students, teachers.

Encourage Policies that Align with Best Assessment Practices

- Why Does It Matter?



THE CENTER ON
**STANDARDS &
ASSESSMENT
IMPLEMENTATION**

WestEd  CRESST



GOOD INTENTIONS

BAD RESULTS



Aligning Policies with Practice

- Standards for Educational and Psychological Testing
 - Use and interpretation of Educational Assessment
 - Peer Review
 - Testing Action Plan



Let's Talk

- ✓ Think of an example where the best of intentions, with regard to assessment and or assessment results resulted in negative, unintended consequences.
- ✓ How might you advise states to consider both the intended and unintended consequences of assessment policy decisions.

For more information please contact:

Deborah Sigman
dsigman@wested.org



THE CENTER ON
**STANDARDS &
ASSESSMENT
IMPLEMENTATION**

WestEd  CRESST