Louisiana Believes

Raising Expectations and Improving Comparability



Agenda

- 2014-2015 Assessment Overview
- 2014-2015 Assessment Creation, Administration, Scoring and Reporting Process
- Cut Scores and Preliminary Louisiana Results

Progress toward Higher Expectations and Improved Comparability

Louisiana has steadily increased the level of expected performance on state tests and has steadily improved its ability to make comparisons with other states.

1999

• Grade 4 and 8 LEAP assessments designed to be as challenging as NAEP. However, results are not comparable with other states. "Approaching basic" (level 2) and levels above earn schools performance score points.

2006

• Grade 3, 5, 6, and 7 iLEAP assessments designed to be as challenging as NAEP. However, results are not comparable with other states.

2013

• Grades 3 – 8 and high school English language arts and math transitional assessments align to Louisiana's new standards. Only "basic" and above earn school performance score points. High schools achieve comparability through ACT 11th grade assessment.

2015

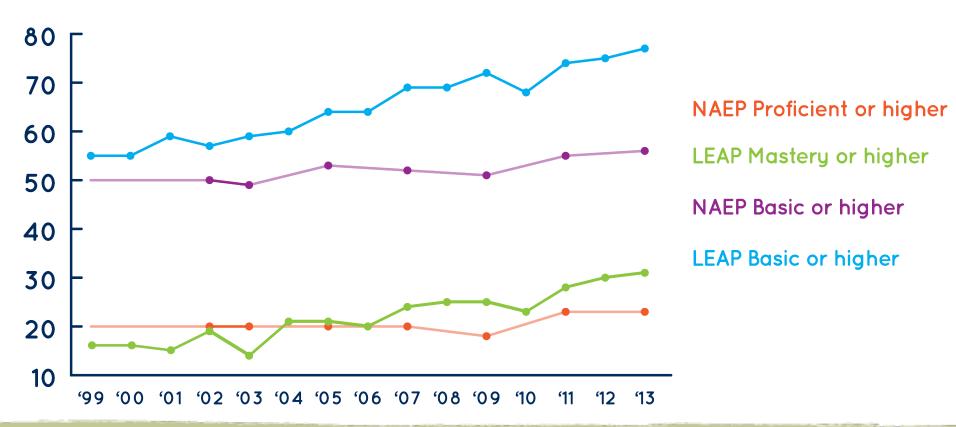
• Grades 3 – 8 English language arts and math exams fully aligned to Louisiana's standards. Results are significantly comparable with other states for the first time.

2025

• By 2025 schools earning ratings of 'A' will average "mastery" performance rather than "basic."

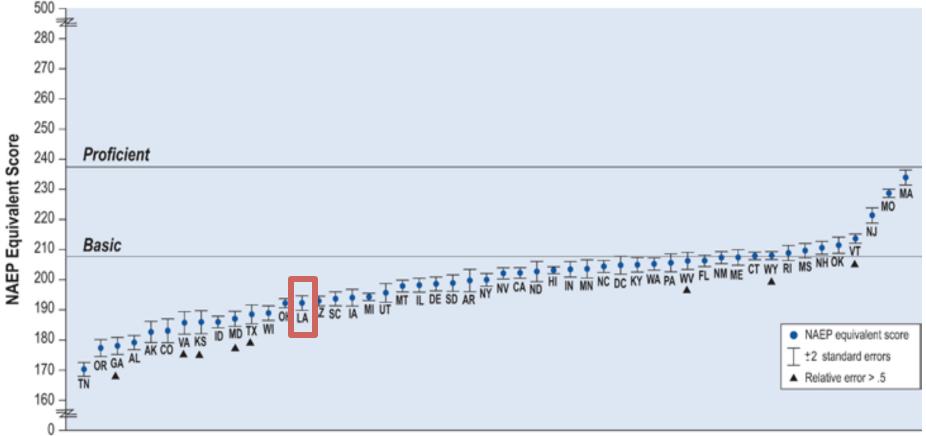
The Case for Raising Expectations

Since making "basic" (level 3 of 5) a standard expectation in Louisiana, the number of students achieving "basic" has grown significantly. Growth at the "mastery" level, however, has been modest. The result is a great number of students called "proficient" in Louisiana but actually not proficient according to NAEP, ACT, and institutions of higher learning. While we should be proud of our progress in getting more students to "basic," we should recognize that "basic" can represent a false promise of readiness.



The Case for Improving Comparability

The false promise is compounded when Louisiana's "basic" is compared with other states' generally accepted proficiency levels. States have often masked low expectations for performance. Comparable performance expectations ensures states cannot mask low expectations.



NAEP scale equivalents of state grade 4 reading standards for proficient performance, by state: 2009

The Center for Assessment

Higher Expectations and Improved Comparability in the Law

The Louisiana Legislature in 2012 recognized the problem of false promises, and thus placed into the law additional requirements for high expectations and improved comparability.

- (a) Standards-based **assessments in English language arts, mathematics, science, and social studies based on state content standards** and rigorous student achievement standards set with reference to test scores of students of the same grade level nationally shall be implemented by the State Board of Elementary and Secondary Education. Such tests shall be administered, at a minimum, in grades three through eleven.
- (b) Beginning with the 2014-2015 school year, standards-based assessments implemented by the State Board of Elementary and Secondary Education in English language arts and mathematics shall be based on nationally recognized content standards that represent the knowledge and skills needed for students to successfully transition to postsecondary education and the workplace. Rigorous student achievement standards shall be set with reference to test scores of the same grade levels nationally.
- (c) The rigor of each standards-based assessment, at a minimum, shall be comparable to national achievement tests, including but not limited to the National Assessment of Education Progress.

* RS 17:24.4: F.(1)

Agenda

- 2014-2015 Assessment Overview
- 2014-2015 Assessment Creation, Administration, Scoring and Reporting Process
- Cut Scores and Preliminary Louisiana Results

Designing a Test for Higher Expectations and Improved Comparability

PARCC was a collaborative process whereby states sought a test aligned to the NAEP, with inter-state comparability.

The following groups played key roles:

- PARCC Consortium: The group of states working together to build and administer the PARCC assessment (Louisiana was a consortium member during the test's creation)
- PARCC Inc.: The nonprofit project manager for the PARCC Consortium
- PARCC Educator Leader Cadre (ELC): Louisiana educators and their peers from other states
- Data Recognition Corp. (DRC): The LEAP vendor for publishing, distributing, and scoring
- Department of Education Staff: Content, assessment, analytics and accountability experts served on test design teams

Field Test

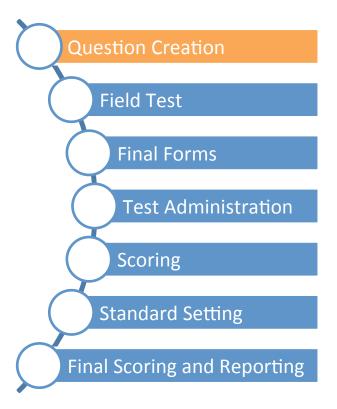
Final Forms

Test Administration

Scoring

Final Scoring and Reporting

Standard Setting



2012 - 2014

The Educator Leader Cadre and Louisiana Department staff worked as a part of the PARCC consortium for over two years to create assessment questions aligned to Louisiana's standards and shared with other states.

Sample of Louisiana's Participants:

Renee Casberque, Associate Professor/Interim Associate Dean, Louisiana State University; Dawn Cassady, Assistant Professor of Curriculum, Instruction, and Leadership, Louisiana Tech University; Clayton Delery, English Instructor, Louisiana School for Math, Science, and the Arts; Kaycee Eckhardt, Teacher, Collegiate Academies: Science Academy; Demetria Gaines, Teacher, School for the Deaf; Kathleen Judy, ELA Assessment Consultant, Louisiana Department of Education; Sandy Landry, Teacher, Jefferson Parish Public School System; Jackie Lewis, Inclusion Teacher, South Grant Elementary/Grant Parish School Board; Carol Price, High School Math Teacher & K-12 District Math Curriculum Specialist/ Math Trainer, Zachary Community School System; Carolyn Sessions, CCSS Math Consultant, Louisiana Department of Education; Whitney Whealdon, ELA Program Coordinator, Louisiana Department of Education: Doris Williams-Smith. Professor - Curriculum & Instruction. Gramblina State University: Martha Younger, Teacher, Central Community School System; Alana Benoit, Teacher, Vermilion Parish; Rachel Gifford, Curriculum Coach, Bossier Parish; Princesses Hill, Teacher, Caddo Parish; Devan Trahan, Teacher, St. Mary Parish; Brandan Trahan, Teacher, St. Mary Parish; Shavela Harvey, Teacher, Calcasieu Parish Schools; Emma Jordan, Supervisor of Curriculum, 6-8, Bossier Parish Schools; Jan Sibley, Assessment Development Section Leader, Louisiana Department of Education; Michelle McAdams, Mathematics Assessment Coordinator, Louisiana Department of Education; Lynne Nielsen, Assistant Professor, Louisiana Tech University; Chanda Johnson, EAGLE Math Developer, Louisiana Department of Education; Sharon Necaise, Education Program Consultant, Louisiana Department of Education; Beth Strange, Education Program Consultant, Louisiana Department of Education; Lyndelle Theriot, Assistant Principal, Vermilion Parish; Serena White, Education Program Consultant, Louisiana Department of Education

Question Creation Field Test **Final Forms Test Administration** Scoring **Standard Setting** Final Scoring and Reporting

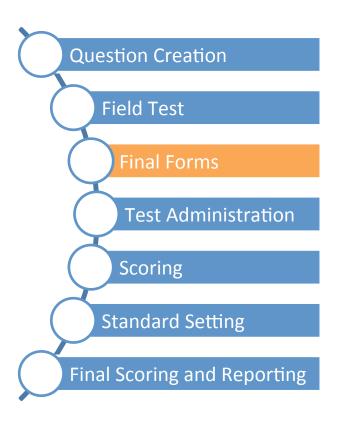
SPRING 2014

Louisiana field tests the PARCC exams.

- Districts practiced the assessment.
- Students experienced the newquestions.
- Educators learned about the accessibility and accommodations features.
- Louisiana teachers, students, and families provided feedback to improve the assessment.
- The PARCC consortium gathered information to confirm question quality and scoring.

More than 45,000 Louisiana students took the field test.

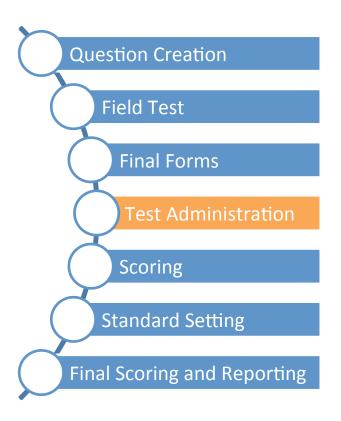
- No major technology issues were reported.
- Students found the test to be easy to navigate if they had engaged in the tutorial items.



SUMMER – FALL 2014

The Louisiana Educator Leader Cadre and Department staff, as a part of the PARCC consortium, constructed test forms for the spring 2015 administration.

- Each question was reviewed to confirm effectiveness during the field test.
- Ineffective questions were removed.
- Effective questions were put together in final and complete forms for the spring 2015 assessments.
- Forms mix difficult questions with simpler questions, based on information gathered from the field test.
- A rubric is finalized to score each question.
- Mix of questions should yield distribution of student scores similar to that of the NAEP.



FALL 2014 - SPRING 2015

Winter 2014: The Department released the 2015 results timeline (November), practice tests (December) and assessment checklist (December). Districts received additional support through administration manuals, guides, webinars, in-person meetings, online office hours, the assessment hotline and assessment@la.gov.

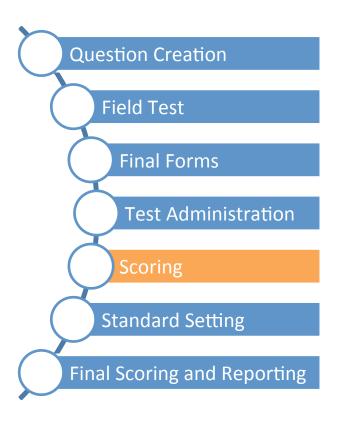
Approximately 320,000 students in grades 3-8 completed PARCC testing in the spring of 2015. 98.5% of students in grades 3-8 statewide participated in the tests.



JUNE – SEPTEMBER 2015

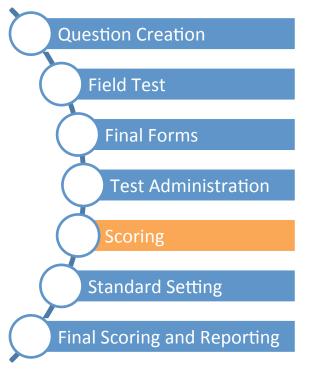
All assessments were scored by DRC.

- DRC scored all constructed response questions.
- DRC scored all multi-select responses.
- Department staff quality checked more than 640,000 individual student responses and scores to make sure that scorers' responses were accurately reflected in each student's raw score, which is the total number of points each student achieved out of the total number available.



How the test is scored is a function of how its tasks are designed.

- PARCC assesses the full scope of reading, writing, and math standards in grades 3-8.
- The test asks students to demonstrate mastery of standards in combination with one another by completing multi-step "tasks." This is different from standardized tests of the past, which tended to ask students to show one step or to fill in a multiple choice bubble only.
- To ensure the tasks measure the scope of the standards, test makers group standards and create descriptions of the skills students should demonstrate on each task. These "evidence statements" guide the design of the task.



This third grade task is aligned to one evidence statement combining two standards .

16. Part A

What is the number with the **least** value that can be made with the digits 6, 7, and 5 using all the digits only once?

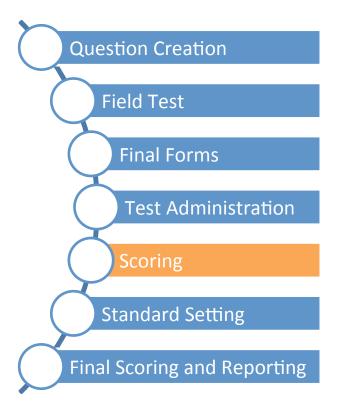
- A 576
- ® 657
- @ 675
- © 567

Part B

Daniel says the number with the **greatest** value he can make with the digits 5, 7, and 6 using the digits only once is 657 because the 7 is in the place with the greatest value.

- Explain why Daniel is **not** correct.
- What is the number with the greatest value he can make using all the digits only once?
- · Explain how you know this number has the greatest value.

Enter your answer and your explanations in the space provided.



Bundling standards within multi-step tasks makes scoring the assessment very different from how we typically think about earning a grade on a test.

Teachers give quizzes regularly, for example, to check for understanding of a specific standard at a specific moment in time. In these cases, they are not asking a student to show all that she has learned across a year.

Sample teacher quiz:

2.
$$6-2 =$$
 (1 point)

3. If Jose has 4 apples and 2 oranges how many pieces of fruit does he have? (1 point)

PARCC tasks allow students to show a wide range of skills, rather than just a quick snapshot. They are not scored on a "percentage right" basis, as with a quiz.

Sample state assessment:

- 1. Jose went to the market. He bought 4 apples, 2 oranges, 4 carrots and 7 potatoes.
 - How many pieces of fruit did Jose buy? (1 point)
 - How many more vegetables did Jose buy than pieces of fruit? (1 point)
 - Jose's friend Angela asked why he bought 10 pieces of fruit. What did Angela do incorrectly? (2 points)
- 2. Kumar had a birthday party. He invited 6 friends. His Dad bought 4 cupcakes and 3 ice cream cones for the party.
 - 2 of Kumar's friends could not attend. How many friends attended? (1 point)
 - Kumar told his dad that he did not get enough treats for the party. Explain why Kumar was wrong. (2 points)

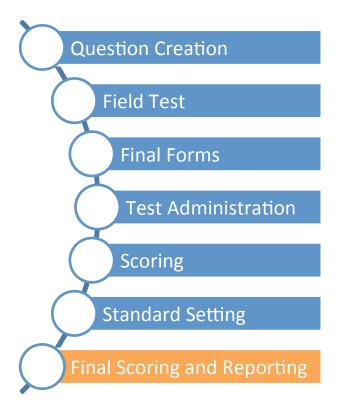
In this example, a student who achieved all points on the teacher quiz may only earn half of the points on the task. This does not mean that the student failed. It means that he has only partially mastered the standards measured on this task.

Question Creation Field Test Final Forms **Test Administration** Scoring **Standard Setting** Final Scoring and Reporting

AUGUST – SEPTEMBER 2015

The PARCC cut scores represent student performance at 5 levels, like the LEAP.

- Educators confirmed the skills required by the standards to be fully prepared for the next grade (identified as level 4).
- Educators assigned an achievement level of 1-5 (basic, mastery, etc.) for performance on each task.
- After tallying raw scores (total points scored out of total available), test makers established conversions to scale scores (650-850). Scales ensure consistent reporting across varying forms, grades, and years.
- Cut scores represent the points between 650 and 850 at which a student has consistently shown a certain achievement level of 1-5 (basic, mastery, etc.).



ACT and AP exams go through similar raw to scale to achievement level conversions.

OCTOBER – NOVEMBER 2015

The Department verifies and reports results to students, parents, schools, and districts.

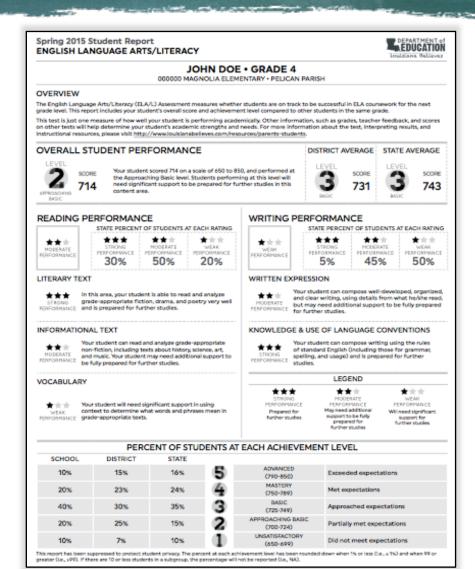
- Raw to scale scores: Department staff convert raw scores for 320,000 students into approximately 4,000,000 scale scores, including sub-categories (e.g. literary text, written expression).
- Scale scores to achievement levels: Once BESE has approved cut scores and correlating achievement levels (Advanced, Mastery, Basic, Approaching Basic, Unsatisfactory), Department staff applies cut scores to approximately 640,000 individual scale scores.
- **Student reports**: Department staff produce 640,000 individual student reports; reports are double checked
- School reports: Department staff validate school and district rosters for 2014-2015 and 2015-2016 school years; each school and district report is generated and double checked for accuracy.

Timeline for Development of 2014-2015 Assessment & Accountability Results

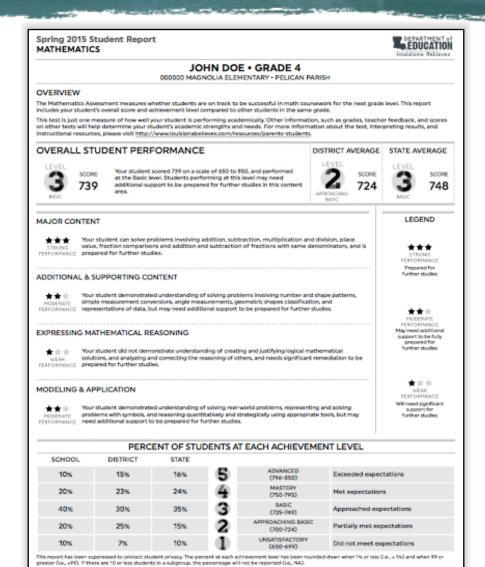
This timeline provides specific dates and weeks at which individual student raw scores, scale scores, cut scores, and skill reports will be created and reported.

Dates	LDOE Action
November 2014	Department announces 2015 test score release schedule
March/May 2015	320,000 students take PARCC tests
June – August	Individual test questions scored by LEAP vendor
Aug –Sept	PARCC state "standard-setting" verifies that questions were as challenging as anticipated before students completed test.
Sept. 28 – Oct. 2	Individual student raw scores (total points out of total available) available to requesting districts
Oct. 5 – 9	Statewide briefings from technical experts on standard setting, scale scores, cut scores, and comparability among Louisiana and other states, in advance of BESE meeting.
Oct. 12	Public release of preliminary statewide scale scores (state-level only; not by LEA level or school level)
Oct. 13	BESE considers cut score levels to determine mastery, advanced, basic, approaching basic, and unsatisfactory
Oct. 14	Department begins applying cut scores to scale scores
Oct. 19 – 23	Public release of LEA scores by cut level
Oct. 26 – 30	Public release of high school performance scores and letter grades (this is the latest date; may be completed earlier).
Nov. 9 – 13	Individual student reports for LEAs, teachers, and families detailing scores and skills for every student
December	Elementary and middle school performance scores and letter grades released

2015 Student Reports - English



2015 Student Reports - Math



School and District Information

Schools, districts, and academic committees, such as the Standards Review Committees and the Accountability Commission, will also receive detailed information in November. Schools and districts will be provided with data that illustrate performance on groups of standards for individual students.

English language arts analysis:

- Reading: literary text
- Reading: non-fiction text
- Reading: vocabulary
- Writing: written expression
- Writing: knowledge and use of language conventions

Mathematics:

- Major content: grade level core standards
- Additional and supporting content: grade level supporting standards
- Expressing mathematical reasoning
- Modeling and application

The Department will release guides to using these data, resources to support instruction for areas of weakness, and training at the November Teacher Leader event.

Agenda

- 2014-2015 Assessment Overview
- 2014-2015 Assessment Creation, Administration, Scoring and Reporting Process
- Cut Scores and Preliminary Louisiana Results

Cut Scores

- Cut scores are the points along the test's scale that indicate students have generally demonstrated performance levels (levels 1-5) on tasks throughout the tests.
- States use the same process and formulae for converting raw scores into scale scores. They also use the same cut scores. This means that "level 4," or "mastery" represents a comparable level of performance in all participating states.
- Cut scores allow the state to classify student performance within categories (basic, mastery). Only after cut scores have been determined can the state produce the following:
 - Reports for parents and teachers
 - School and district performance scores (SPS)
 - Guidance for principals and teachers setting evaluation targets
 - School and district letter grades
 - Charter school renewals and scholarship school eligibility based on SPS or grades
 - Student and family eligibility for school choice
 - School listings in OneApp materials

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
790 – 850	810 – 850	5	Advanced
750 – 789	750 – 809	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
796 – 850	790 – 850	5	Advanced
750 – 795	750 – 789	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
790 – 850	799 – 850	5	Advanced
750 – 789	750 – 798	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
788 – 850	790 – 850	5	Advanced
750 – 787	750 – 789	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
786 – 850	785 – 850	5	Advanced
750 – 785	750 – 784	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

- Every task is scored to show a proficiency level from 1 to 5. The total raw score is converted to a scale score. The cut scores proposed below are the places on the scale at which students typically demonstrated a given performance level on tasks.
- These are the same cut scores as have been or will be used in other states.

Scale Score Range MATH	Scale Score Range ELA	Performance Level	Achievement Level
801 – 850	794 – 850	5	Advanced
750 – 800	750 – 793	4	Mastery
725 – 749	725 – 749	3	Basic
700 – 724	700 – 724	2	Approaching Basic
650 – 699	650 – 699	1	Unsatisfactory

Preliminary Student Results on These Cut Scores

In most grade levels, in both subjects, typically 30 to 40 percent of Louisiana students show "mastery" command of skills needed in community college and universities.

English Language Arts

Mathematics

Grade	% at 5	% at 4	% at 3	% at 2	% at 1	Grade	% at 5	% at 4	% at 3	% at 2	% at 1
3	2	35	26	21	16	3	6	31	30	22	11
4	4	36	34	19	8	4	2	31	31	26	10
5	1	32	34	24	9	5	3	25	31	32	9
6	3	35	36	19	7	6	3	23	33	32	10
7	5	29	32	22	12	7	2	20	36	29	13
8	4	36	30	19	11	8	4	28	23	25	20

Preliminary Student Results on These Cut Scores

In most grade levels, in both subjects, typically 30 to 40 percent of Louisiana students show "mastery" command of skills needed in community college and universities.

English Language Arts

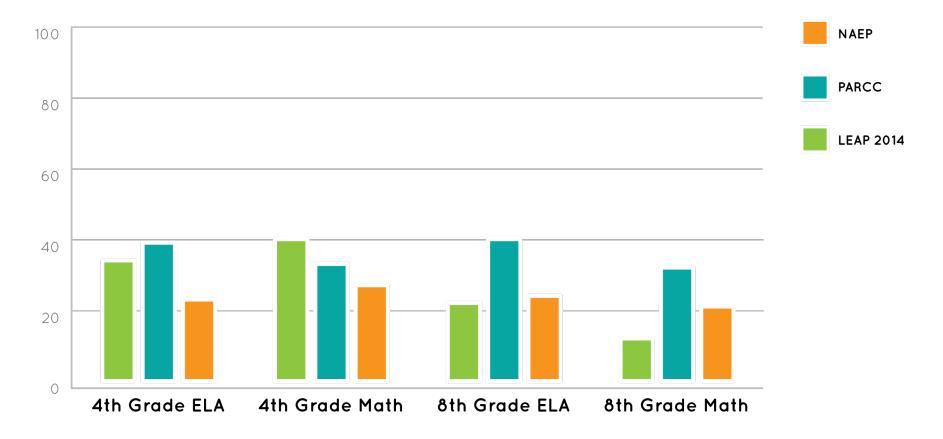
Basic	% at Basic and Above	% at Mastery and Above
3	64	37
4	73	39
5	67	33
6	74	38
7	67	35
8	70	40

Mathematics

Grade	% at Basic and Above	% at Mastery and Above
3	67	37
4	64	33
5	59	28
6	59	26
7	58	22
8	55	32

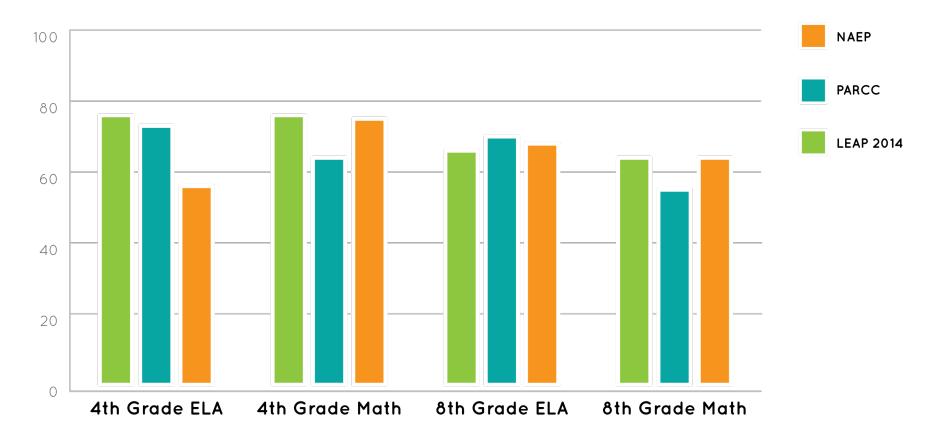
Student Results on These Cut Scores

The percentage of Louisiana students demonstrating at least "mastery" command of skills needed in community colleges and universities is generally consistent with evidence from other tests.



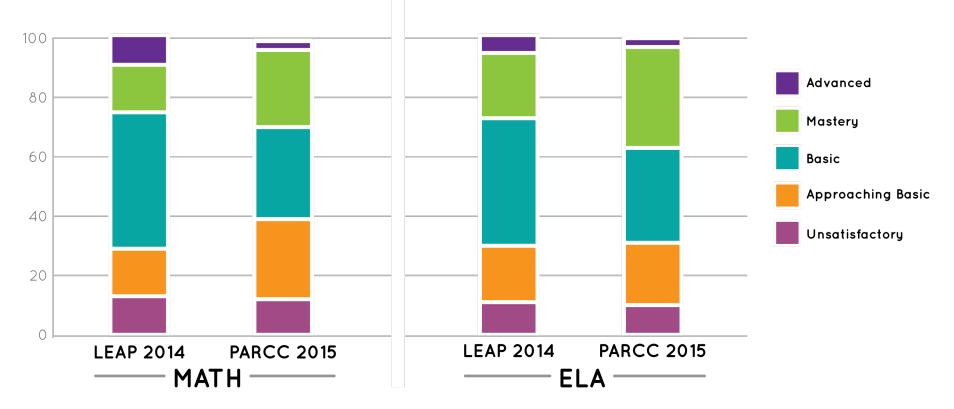
Student Results on These Cut Scores

The percentage of Louisiana students demonstrating at least "basic" command of skills needed in community colleges and universities is generally consistent with evidence from other tests.



Student Results on These Cut Scores

PARCC tasks were more challenging than LEAP questions, collecting more evidence across more standards. Higher performing students tended to show more evidence of mastery than in the past, while lower achieving students tended to show less evidence of even basic skills. Whereas nearly half of students performed at "basic" on the LEAP, PARCC has distributed scores to a greater degree across the spectrum.



Raising Expectations through 2025

Results from 2015 and 2016 will be comparable to one another and to results in other states. These results will combine to form a "baseline" measurement of Louisiana performance on new standards, in comparison with other states.

Having established this baseline, BESE will create a steady transition toward 2025, when an A-rated school in Louisiana will have an average performance of "mastery" rather than "basic," as is the case today.

This means that each year between 2017 and 2025, the state's accountability system will increasingly reward "mastery" results more and "basic" results less.

Achieving Improved Comparability

During the spring of 2015, 5,002,000 students across 12 jurisdictions took the PARCC assessment.

- Arkansas
- Colorado
- District of Columbia
- Illinois
- Louisiana
- Maryland
- Massachusetts
- Mississippi
- New Jersey
- New Mexico
- Ohio
- Rhode Island

Achieving Improved Comparability

Assuming that BESE approves the proposed cut scores, Louisiana's results will be significantly and reasonably comparable to those of all other states using PARCC content. This allows our state to analyze results using comparisons, as do with ACT or AP results.

The Center for Assessment, Louisiana's longstanding technical advisor, is performing an external audit to validate the significant comparability of PARCC scores in Louisiana with those in other PARCC states. The study will evaluate the extent to which it is appropriate to claim that a student's performance on PARCC in Louisiana would have been the same regardless of where she or he took the PARCC test.

Comparability is determined by examining processes, procedures, and materials in three key areas:

- •The content of the test
- The administration of the test
- •The scoring of the test and reporting of results

Achieving Improved Comparability

- ✓ Phase 1 The *Content* of the tests
 - ✓ Compare the test forms administered in Louisiana with those administered in other PARCC states to ensure that the tests were the same.
- ✓ Phase 2 The *Administration* of the tests
 - ✓ Examine test administration manuals, memos, and related materials to ensure that the administration policies and procedures followed in Louisiana were consistent with PARCC policies and procedures.
- Phase 3 Scoring and the Reporting of results
 - ✓ Phase 3a Evaluate the processes and procedures used to score individual items to ensure that all machine-scored and hand-scored items are being scored the same way for Louisiana as they are for other PARCC states.
 - Phase 3b Determine that individual item scores have been accurately combined to produce student raw scores and accurately converted to PARCC scaled scores, performance levels, and sub-category scores.
 - Phase 3c Examine the Louisiana policies regarding the inclusion of students in the reporting of school, district, and state results to ensure that those are consistent with those in other states administering PARCC.

Louisiana Believes

Raising Expectations and Improving Comparability

