INSTRUCTIONS

Section 200.105(a)(d)(3) of the regulations for the Innovative Assessment Demonstration Authority provide that State(s) receiving the authority must report the following annually to the Secretary, at such time and in such manner as the Secretary may reasonably require:

(i) An update on implementation of the innovative assessment demonstration authority, including--
   (A) The SEA’s progress against its timeline under 34 CFR 200.106(c) and any outcomes or results from its evaluation and continuous improvement process under 34 CFR 200.106(e); and
   (B) If the innovative assessment system is not yet implemented statewide consistent with 34 CFR 200.104(a)(2), a description of the SEA’s progress in scaling up the system to additional LEAs or schools consistent with its strategies under 34 CFR 200.106(a)(3)(i), including updated assurances from participating LEAs consistent with paragraph (e)(2) of this section.

(ii) The performance of students in participating schools at the State, LEA, and school level, for all students and disaggregated for each subgroup of students described in section 1111(c)(2) of the Act, on the innovative assessment, including academic achievement and participation data required to be reported consistent with section 1111(h) of the Act, except that such data may not reveal any personally identifiable information.

(iii) If the innovative assessment system is not yet implemented statewide, school demographic information, including enrollment and student achievement information, for the subgroups of students described in section 1111(c)(2) of the Act, among participating schools and LEAs and for any schools or LEAs that will participate for the first time in the following year, and a description of how the participation of any additional schools or LEAs in that year contributed to progress toward achieving high-quality and consistent implementation across demographically diverse LEAs in the State consistent with the SEA's benchmarks described in 34 CFR 200.106(a)(3)(iii).

(iv) Feedback from teachers, principals and other school leaders, and other stakeholders consulted under paragraph (a)(2) of this section, including parents and students, from participating schools and LEAs about their satisfaction with the innovative assessment system.
In addition, Title I, Part B, section 1204(c)(2) of the Act requires that progress shall be reported based on the annual information submitted by participating States described in subsection (e)(2)(B)(ix) and examine the extent to which—

(A) with respect to each innovative assessment system—

(i) the State educational agency has solicited feedback from teachers, principals, other school leaders, and parents about their satisfaction with the innovative assessment system;
(ii) teachers, principals, and other school leaders have demonstrated a commitment and capacity to implement or continue to implement the innovative assessment system; and
(iii) substantial evidence exists demonstrating that the innovative assessment system has been developed in accordance with the requirements of subsection (e)

(B) each State with demonstration authority has demonstrated that—

(i) the same innovative assessment system was used to measure the achievement of all students that participated in the innovative assessment system; and
(ii) of the total number of students, and the total number of each of the subgroups of students defined in section 1111(c)(2), eligible to participate in the innovative assessment system in a given year, the State assessed in that year an equal or greater percentage of such eligible students, as measured under section 1111(c)(4)(E), as were assessed in the State in such year using the assessment system under section 1111(b)(2).

Definitions:

- Participating LEA means an LEA in the State with at least one school participating in the innovative assessment demonstration authority.

- Participating school means a public school in the State in which the innovative assessment system is administered under the innovative assessment demonstration authority instead of, or in addition to, the statewide assessment under section 1111(b)(2) of the Act and where the results of the school’s students on the innovative assessment system are used by its State and LEA for purposes of accountability and reporting under section 1111(c) and 1111(h) of the Act.

*To meet the requirements for this annual performance report, please provide the requested information in each of the sections that follow. The U.S. Department of Education understands that coronavirus may have affected the development and implementation of innovative assessment systems during the reporting year (2022-23). To the extent your SEA would like to provide more context or details related to these impacts, please incorporate them into your responses where relevant.*
Note that in place of extensive appendices, each piece of evidence is provided as a separate file and is indexed in a separate spreadsheet file, Artifact Index LA IADA 2023.xlsx. This report references that evidence list. The evidence is indexed by section, for example 2_001 School System Participation for 23-24 is the first piece of evidence relevant to section II: Student Performance.

SY22-23 Annual Performance Report Table of Contents

I: Progress Toward Plan and Timeline.................................................................................................................................................................................................................... 4
   I.A. Current Progress................................................................................................................................................................................................................. 5
   I.B. Scaling Up................................................................................................................................................................................................................... 6
   I.C. Participating IAP School Systems.................................................................................................................................................................................. 29
II: Student Performance.................................................................................................................................................................................................. 32
   II.A. Performance of Students in Participating Schools Disaggregated for Each Subgroup........................................................................................................... 32
   II.B. State-level Participation Rate of Students Disaggregated for Each Subgroup........................................................................................................... 32
III: School Demographic Information.................................................................................................................................................................................................. 33
   III.A. School Demographic Information including Enrollment......................................................................................................................................... 33
   III.B. Schools Participating for the First Time in the 2023-24 School Year........................................................................................................................................... 33
IV: Consultation and Feedback.................................................................................................................................................................................................. 34
V: Requirements for the Innovative Assessment System........................................................................................................................................ 39
   V.A. Developing a Valid, Reliable, and Comparable System........................................................................................................................................... 39
   V.B. Update on Meeting Requirements................................................................................................................................................................................. 44
VI: Training on and Familiarization with the Innovative Assessment System.................................................................................................................................. 60
   VI.A. Statewide Training for Standard Administration of the Innovative Assessment System.......................................................................................... 60
   VI.B. Describe All Innovative Assessment Training Opportunities.................................................................................................................................. 63
   VI.C. Familiarization of Innovative Assessment to Students, Parents, and LEA Staff...................................................................................................... 68
VII: Use of Innovative Assessment Data.................................................................................................................................................................................................. 70
VIII: Changes in Consortium Governance or Membership (if applicable)................................................................................................................................ 70
IX: Parental Notification...................................................................................................................................................................................................... 71
X: Assurances.............................................................................................................................................................................................................. 71
XI: Budget.............................................................................................................................................................................................................. 72
XII: Certification........................................................................................................................................................................................................... 72
1: Progress Toward Plan and Timeline

Provide a description of the SEA’s (or Consortium’s) progress towards its plan and timeline in its approved application:

Louisiana’s IADA pilot assessment program, known as the Innovative Assessment Program (IAP), is a curriculum aligned, through-year English language arts (ELA) assessment. Students take three unit-based assessments throughout the year (Fall, Winter and Spring), which are then combined to produce single summative scores to support annual determinations for use in Louisiana’s Every Student Succeeds Act compliant system of school identification and support.

Currently, the IAP is designed for two specific curricula, the (1) ELA Guidebooks and (2) Wit and Wisdom. For these curricula, IAP End-of-Unit (EOU) assessments are developed to assess the knowledge students gain within each curricular unit, as well as how they can build on and apply that knowledge. There are, therefore, two sets of EOU assessments, those for ELA Guidebooks and those for Wit and Wisdom, in addition to the current statewide assessment, known as the LEAP (Louisiana Educational Assessment Program). The unit assessments are linked to the LEAP scale via a non-item nonequivalent groups design, in which students taking each IAP EOU assessment also take a set of items from the LEAP assessment. This means that each curricula option within the IAP is linked to, and reported out on, the current statewide scale. The ELA Guidebooks IAP has now reported operationally for two school years, for grade seven in 2021-2022 and for grades six, seven, and eight in 2022-2023. In addition, during 2022-23 unit based assessments for grade five were field tested for the ELA Guidebooks IAP with the intent to administer operationally during the 2023-24 school year. Pilot and field testing activities are ongoing for the ELA Guidebooks IAP. During 2022-23 unit based assessments for grade five were also field tested for the Wit and Wisdom IAP with the intent to administer operationally during the 2023-24 school year. However, as of late August, LDOE suspended the operational administration of the grade five Wit and Wisdom IAP, as no schools or districts volunteered to participate. The primary reason was that schools and districts felt that having a different test in a single grade, grade five, created a disconnect in student experience. These schools and districts want a complete set of assessments for grades three to five.

In addition to the ELA Guidebooks and Wit and Wisdom curriculum specific unit assessments, LDOE and its partners have been developing prototype unit assessments that are meant to span multiple curricula. These units are meant to be “curriculum-relevant,” in that they connect to curricula, but do so less directly than the current curriculum specific unit assessments. This curriculum-relevant option has been the source of much iteration, with LDOE and its content partners, particularly Odell education, developing numerous designs, one of which was the subject of a series of focus groups in early Spring 2023. Ultimately, this work resulted in a single grade five unit assessment that contained content that was shared across two curricula - ELA Guidebooks and Wit and Wisdom - as well as unique content. This grade five unit assessment was piloted with a small number of students during the Spring 2023 administration window¹. However, as explained below, this assessment is just a step towards a

¹ Psychometric examination of this unit assessment is now underway, and can be made available at a future date.
system that will scale across all curricula used in Louisiana. The work during the 2023-24 is meant to further develop options for scaling statewide across all curricula.

Note that the full set of administered units during the 2022-23 school year can be found within [1_001] SY2023 Administration Unit Plan – W1, W2, W3.
### I.A. Current Progress

#### Table 1-1. Implementation Plan with Counts

<table>
<thead>
<tr>
<th>Grade</th>
<th>Curriculum</th>
<th>SY 21-22</th>
<th>SY 22-23</th>
<th>SY 23-24</th>
<th>SY 24-25&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 7</td>
<td>ELA Guidebooks</td>
<td>4,447 (Operational&lt;sup&gt;1&lt;/sup&gt;)</td>
<td>6,437 (Operational)</td>
<td>9,000 (Operational)</td>
<td>12,000 + (Operational)</td>
</tr>
<tr>
<td>Grade 8</td>
<td>ELA Guidebooks</td>
<td>6,031 (Field Test&lt;sup&gt;3&lt;/sup&gt;)</td>
<td>4,343 (Operational)</td>
<td>9,000 (Operational)</td>
<td>12,000 (Operational)</td>
</tr>
<tr>
<td>Grades 6</td>
<td>ELA Guidebooks</td>
<td>7,338 (Field Test)</td>
<td>4,903 (Operational)</td>
<td>9,000 (Operational)</td>
<td>12,000 (Operational)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>ELA Guidebooks</td>
<td>2,737 (Pilot)</td>
<td>4,670 (Field Test)</td>
<td>5,000 (Operational)</td>
<td>6,000 + (Operational)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>Wit and Wisdom</td>
<td>958 (Pilot)</td>
<td>1,609 (Field Test)</td>
<td>Paused</td>
<td>3,000 + (Operational)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>ELA Guidebooks</td>
<td>--</td>
<td>Development</td>
<td>6,000 (Field Test)</td>
<td>5,000 (Operational)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>ELA Guidebooks</td>
<td>--</td>
<td>Development</td>
<td>6,000 (Field Test)</td>
<td>5,000 (Operational)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>Curriculum Relevant Third Option (also referred to as “Crawfish”)</td>
<td>Development (Including extensive stakeholder engagement)</td>
<td>350 (Pilot)</td>
<td>2,000 (Pilot)</td>
<td>To Be Determined</td>
</tr>
</tbody>
</table>

Notes: Green shading indicates operational testing, yellow shading indicates field test/pilot/prototype testing phases, and blue shading indicates research and development years. The counts above combine the unique patterns of unit assessments for ELA Guidebooks. Counts for 23-24 and beyond are estimates. <sup>1</sup>Operational counts are based on the number of students with scores potentially valid for use in Louisiana’s system of school identification and support. Final counts may differ slightly due to LDOE’s application of rules for reporting. <sup>2</sup>For comparison, the total enrollment for any given grade is approximately 42,000 students. <sup>3</sup>Field test counts reflect all tested students, not just those with potentially valid cases. The grade five Wit and Wisdom IAP unit assessments were originally meant to be administered operationally in 2023-24, however the state was unable to secure participation of any school or district.
I.B. Scaling Up
Extending this design to work statewide, i.e., scaling up, involves a significant amount of complexity. Whereas scaling up in the context of other programs involves increasing (a) the number of participating students, potentially involving more grades if not implemented in grades 3-8 and high school, and (b) items within the item bank. Scaling up may mean the IAP involves ensuring that all high quality curriculum in use within Louisiaian schools are accounted for within the IAP.

This means addressing:

1. **The number of curricula in use within the state.** To date, the Louisiana Department of Education (LDOE) has not yet identified any substantial number of schools and districts using curricula other than ELA Guidebooks or Wit and Wisdom. LDOE was able to identify one school district using EL Education, but was unable to pilot curriculum based assessments with that district until this 2023-2024 school year. Given this, a current design challenge is: how (or if) to provide an End of Unit assessment for less frequently used curricula.

2. **The variation in the implementation of curricula, within the state.** There are often substantial differences in the ways in which the same curriculum is implemented from classroom to classroom, in terms of pacing, as well as unit selection. For example, at any given point during the year, some classrooms may be moving on from one unit whereas other classrooms are only half way through the same unit, or a particular school system may choose a different unit text. The IAP, with three fixed windows, has provided choice in the order and sequence of the first two of three instructional units\(^2\). Beyond these straightforward ways to understand differences in curricular implementation, which are related directly to the IAP, there are also varying ways in which teachers emphasize, adapt, modify and supplement the curriculum, often as a result of varying pedagogical perspectives. These aspects of curricular implementation are important, but do not directly interact with the design of the IAP.

3. **Updates to curriculum.** Curricula are not static. Instead, they are updated in fairly regular intervals. In the case of a program connected to curricula like the IAP, updates to curricula may mean updates to the assessments. This is particularly true of a curriculum based assessment which relies on items that are not only thematically, but also specific to particular books which may not be in the curriculum after revisions. Additionally the timeline on which schools adopt the new curricula or choose to remain with the older curricula vary, which may mean that those implementing a program like the IAP may need to consider whether or not to support multiple versions of a particular curriculum, or just one version, like the latest. Currently, the IAP is designed around the most current version of the ELA Guidebooks and Wit and Wisdom curricula, but both are expected to be updated in the near future.

\(^2\) For ELA Guidebooks only, as this curriculum has units that are designed to be interchangeable. The Wit & Wisdom curriculum is meant to be sequential, so the IAP does not provide choice in the units administered.
Throughout the 2022-23 school year, the IAP teams worked to address point (1) above. Point (2) has largely been addressed through the administration structure, whereas (3) has yet to become an issue, although Wit & Wisdom is currently undergoing a redesign and ELA Guidebooks is slated to in the near future. In terms of point (1) the department has worked with its partners, specifically Odell Education, to consider approaches to assessment that span multiple curricula. As explained previously, this work culminated in the pilot of a unit assessment that had a section with content unique to two curriculum, Guidebooks and Wit & Wisdom, as well as content that is shared across the curriculum. Essentially, this unit assessment had two forms that are parallel in structure, but differ in terms of the curriculum specific sections. This unit assessment, referred to as the IAP Curriculum-Relevant model, was pilot tested in the Spring of 2023 with a limited number of students. However, this curriculum-relevant assessment was only able to be constructed due to fortunate intersections of the curriculum content. An analysis of the curriculum conducted by John Hopkins indicates that such intersections are rare (see [1_002] John Hopkins Curricula Survey). Upcoming work in 2023-24 will continue to investigate this issue.

These points above also illustrate some of the challenges in implementing this kind of model. A high level list of these challenges involves:

- **Connection to Curricula.** While the vast majority of classrooms in the state use Guidebooks or Wit and Wisdom, developing assessments that work for every curriculum is challenging. Essentially, there are just a few options for program design to address these challenges, each with tradeoffs: (1) developing as may assessments as there are curricula, (2) developing curricula-specific assessments for the most widely used curricula, with a “generic” or curriculum agnostic option for less popular curricula, (3) developing a model that spans, but still connects to curriculum, potentially alongside a generic option, or (4) disconnecting the content of assessment from curricula.
  
  - Option (1) is not viable due to the sheer number of possible curricula within the state.
  - Option (2) would require the state to determine what the most popular curricula are within the state, then determine which curricula will have a specific assessment.
  - Option (3) has been the focus development this year, which has produced a form that spans two curricula.
  - Option (4) would reflect a profound pivot from the current design, as well as the work done on curricula within the state.

- **Understanding Curricula within the State.** Understanding who is teaching what, at the state-level, is extremely challenging. Systematic data collection is needed to understand what curriculum is used within each classroom, district, and school. This kind of data collection, even if it is coupled directly with the state assessment program (e.g., by connecting the collection of curricula data to rostering early in the school year) involves the creation of new processes that require the investment of time and resources.

- **Need for Extended Engagement.** Given that this kind of model involves more interactions with students within the classroom, there is a need for significant engagement with students, parents, teachers, leaders and those who represent the interest of multiple student groups, including those students with disabilities and English language learners.
All of the above also point to increased complexity and cost.

I.C. Summary of Activities
The following list provides a high-level overview of key activities from the Table I-3, Detailed, High Level Summary of Activities. Key steps in this stream of work included:

- Continued development activities designed for the innovative assessments in ELA Guidebooks grades 3 and 4, which included:
  - Review of field test forms with the department.
  - Supported detailed discussions around Grades 3 and 4 development and defined the passage criteria.
  - Conducted passage finding and developer training started over the summer months.
  - Held Teacher Committee Reviews for Units 1-5.
  - Supported the Field Test Development for Grades 3 and 4
  - Conducted the passage finding for grades 3 and 4

- Administered assessments for Fall, Winter, and Spring Test Windows for the grades 5 field test - ELA Guidebooks, Wit and Wisdom, and the Curriculum Relevant Option - and grades 6, 7 and 8 operational - for ELA Guidebooks only, which included:
  - Reviewed and analyzed test data from 2021-22 to create field test forms.
  - Assembled and finalized test forms for grades 5-8, to include Text-To-Speech, Braille, and Accommodated Print Forms.
  - Prepared webinars, help desk materials, help desk readiness for the field.
  - Supported IAP school systems during the test windows via Help Desk.
  - Conducted pre-test validation test cases and user acceptance requirements of the state.
  - Conducted operational scoring throughout the test windows for End-of-Unit reporting purposes.
  - Finalized End-of-Unit report mock-ups as directed by the TAC.
  - Conducted post-test validation of scored data and approval of data file and reports.
  - Supported LDOE Reports webinars.
  - Prepared End of Unit reports for release to school districts.

- Conducted Rangefinding Activities, which included:
  - Conducted rangefinding activities to prepare for spring windows.
  - Rangefinding activities begin on Grade 5 field tests and summer workshop activities.

- Conducted Meaningful Stakeholder Consultation Activities, which included:
  - Led focus groups to support the creation of new “enhanced” classroom and student End of Unit reports for SY23-24.
2023 IADA Annual Performance Report

- Conducted a revised achievement level descriptor alignment workshop, in which Louisiana educators evaluated the alignment of IAP items to the statewide achievement level descriptors, and concurrent Reporting Category Standards Setting, in which Louisiana educators set cuts on each of the IAP’s reporting categories.
- Prepared for the Grades 6-8 alignment and reporting category workshop in May 2023 through conducting a “dry run” of the scaling and scoring approach to develop the methods used to produce the final scale scores and achievement levels based on complete data from Fall and Winter (Windows 1 and 2).

- Provided Program and Project Oversight, which included:
  - Delivered the IADA Annual Report for the 2021-2022 school year.
  - Presented the Innovative Assessment Program plans to Louisiana’s Technical Advisory Committee (TAC) at the October, 2022 meeting.
  - Initiated the Weekly Psychometric Calls focused on scaling efforts to prepare for ALD Alignment and final end of year scaling.
  - Supported scope and change management requirements as directed by the needs of the state.
  - Facilitated meetings across the department and the partner organizations through meeting agendas and minutes
  - Finalized the 2022-23 operational program schedule in grades 5-8.
  - Created the new schedule for 2022-2023 development and field tests in Grades 3 and 4.
  - Documented decisions through a decision log.
  - Refined plans for scaling, linking and comparability as recommended by Louisiana TAC which were presented at the NCME in April 2022.
  - Presented the Innovative Assessment plans to Louisiana’s TAC at the February, 2023 meeting.
  - Implemented a final scaling and linking analysis to produce the final scoring tables for the end of year (EOY) operational run.
  - In June 2023, summarized the Innovative Assessment Program to Louisiana Board of Elementary and Secondary Education to continue the contract for another year in 2023-2024.
  - Conducted a number of “comparability analyses,” aimed at providing additional evidence of score-level comparability between the IAP and the statewide assessment.
Table I-2. Multi-Year Summary of Progress Towards Timeline and Plan

<table>
<thead>
<tr>
<th>School Year</th>
<th>Progress Toward Plan and Timeline: the Louisiana Roadmap and Key Activities</th>
</tr>
</thead>
</table>
| 2017-2018   | Louisiana’s pilot assessment program, known as the Innovative Assessment Program (IAP), is a curriculum aligned, through-year assessment. Students take three unit-based assessments throughout the year (Fall, Winter and Spring) which are then combined to produce single summative scores to support annual determinations for use in Louisiana’s Every Student Succeeds Act (ESSA) compliant system of school identification and support. With the vision for improving assessment systems and improving outcomes for all students, the Louisiana Department of Education (LDOE) submitted its application to the US Department of Education (USDOE) for the Innovative Assessment Demonstration Authority (IADA) waiver on March 30, 2018. 

The LEAP (Louisiana Educational Assessment Program) ELA Guidebooks 2.0 assessment is aligned to the ELA curriculum content that the students have studied during each assessment window. This assessment design was put forward because it measures both the content knowledge and skills students derive from studying specific texts and upholds the State’s vision for (1) accessibility, (2) opportunity, (3) alignment, and (4) knowledge.
- Accessibility: Assessments should allow all students the opportunity to demonstrate what they know and can do.
- Opportunity: All students deserve access to complex texts and tasks.
- Alignment: Assessments should support high-quality instruction.
- Knowledge: To become literate, students must build knowledge through coherent units of study.

The IAP uses an ELA curriculum-integrated through-year approach consisting of several assessments administered throughout the school year to inform teaching and learning. The development for the IAP started in Grade 7 for the ELA Guidebooks 2.0 with a focus first on the middle grades. The IAP has historically used Tier I approved curricula for its content design and development plans. ELA Guidebooks 2.0 are available for free nationally through the Louisiana Curriculum Hub and over 80 percent of Louisiana school systems use this curriculum. This alignment to curriculum is important as the Louisiana legislature established an instructional materials process in 2015 which requires school systems to adopt approved textbooks and other instructional materials. LDOE conducts online review of instructional materials to determine the degree of alignment with state content standards. Each local school system determines if their use is appropriate to meet the educational needs of their students.

The original design for the through-year assessment was called the Humanities Assessment. The plan for this assessment was to assess both ELA and social studies standards. On July 24, 2018, LDOE was granted the authority to pilot the curriculum-integrated through-year assessments starting in Grade 7. |
### 2018-2019
LDOE partnered with NWEA, Odell Education, Johns Hopkins University, The Center for Assessment, MZ Development, and Strategic Measurement and Evaluation to develop and pilot an innovative, joint ELA and social studies assessment. LDOE worked with these partners to develop and pilot items in Grade 7. This school year marked the first year in operation under the IADA waiver.

During November and December 2018, an initial design pattern, an Evidence Centered Design tool, was developed to articulate the design of the End-of-Unit assessments. The end-of-unit field test items reflected LDOE’s core beliefs by focusing on 1) texts students read in class and 2) topically-related texts, referred to colloquially as “hot reads” and “warm reads,” respectively. Though students had not been exposed to the second category of texts prior to taking the assessment, the topically-related texts were aligned to the content of the curriculum unit, thereby providing a more equitable assessment experience for students and valuing the knowledge built in class.

The planning teams established the following working definitions:

- **COLD read texts**: random selection of grade-level texts often not connected to anything students have studied in class
- **HOT read texts**: unit texts studied in ELA class
- **WARM read texts**: texts students have not read previously but that are topically related to the information and knowledge they encountered in class

Participating school districts administered the first assessment pilots in Grade 7 and this test delivery was a partial model. Technical documentation for the IAP started as a requirement of the waiver to provide USDOE with annual performance reports. Subsequent validity studies were conducted to deploy an operational assessment for the next school year. Ultimately, the IAP was intended to demonstrate the viability of a through-year assessment as an alternative to the State’s accountability assessment, by closely aligning assessment to curriculum and research, providing actionable information to teachers in a timely manner, and supporting the state’s current systems of accountability.

### 2019-2020
This school year was the second year in operation under the IADA waiver. Grade 7 assessments were administered in the fall and winter testing windows. Using grade 7 as the model, teams began developing tryout forms for grades 6 and 8. Concurrently with the development of the forms design in late 2018 and early 2019, the achievement level descriptors (ALDs) were also developed and informed development. The IAP was designed around the current ALDs for the LEAP 2025 ELA assessment as required under Louisiana’s Board of Elementary and Secondary Education (BESE) rules for nationally recognized and consortium performance standards. These ELA ALDs articulate five levels of performance - Advanced, Mastery, Basic, Approaching Basic, or Unsatisfactory. The scaling plans for this school year were impacted by the COVID-19 pandemic, and LDOE made the decision that no operational testing would occur during this school year.
In June 2020, LDOE submitted for the Competitive Grant for State Assessments (CGSA) for the continued development in elementary Grades 3-5. The USDOE awarded the nearly $3.0 million grant to LDOE to continue through-year assessment development in Grades 3-5 under Absolute Priority 1, that is, "intended to encourage the use of multiple measures of academic achievement."

This school year was the third year in operation under the IADA waiver, but the IAP remained paused due to school disruptions with the pandemic and major hurricanes. Hurricane Ida greatly impacted many school systems in the state. USDOE granted states waivers for accountability, school improvement, and assessments during this school year.

On June 17, 2021, LDOE asked USDOE to grant a two-year extension of the IADA waiver in order to successfully implement its IAP program.

With a new grant and the pandemic effects subsiding, this school year was the fourth year in operation under the IADA waiver and the first time LDOE could fully implement the operational assessment in Grade 7. LDOE successfully administered three assessments in the fall, winter, and spring for Grade 7.

In this year, the IAP expanded with twenty-five school systems administering assessments for grades 5-8. With an exciting opportunity to preserve local control by partnering with Great Minds for the Wit and Wisdom curriculum, the LDOE recruited five school systems to join the Wit and Wisdom pilot. This expanded the curriculum-integrated approach and provided an opportunity for schools that had selected the Great Mind curriculum to participate in the curriculum-aligned model. The Great Minds option became the LEAP ELA Wit and Wisdom for IAP systems with students in Grade 5. The focus of the IADA Waiver turned to scaling in Grades 3-5.

In addition to operationalizing Grade 7, concerted efforts to conduct meaningful consultation with stakeholders were implemented. Surveys to students and teachers were collected to determine the preference for and engagement with the IAP in comparison to the traditional end-of-year LEAP 2025. LDOE conducted an alignment study and reporting category workshop with educators to create IAP ALDs for five levels of performance (i.e., Advanced, Mastery, Basic, Approaching Basic, or Unsatisfactory). These findings were presented to BESE in the summer of 2022. Promising results were found in these initial student achievement data results and stakeholder preferences as measured through the student and teacher surveys.

Another exciting evolution of the IAP was the Education First Grant which allowed exploration into a “curriculum relevant” option as a pilot for the upcoming school year. In the fall of 2021, the Bill & Melinda Gates Foundation, Walton Family Foundation, and the Chan Zuckerberg Initiative seeded funding for Education First to work with assessment developers and state education agencies to pilot and test out through-year models.
Ed First invested in the research and design efforts to prototype three different through-year model approaches. Ed First provided a grant to NWEA and LDOE to pilot a “curriculum-relevant” option. Many school systems adopt more than one high-quality curriculum since curriculum is a local choice. In addition, there are logistical challenges of tying assessment to curriculum: movement of students from one school system to another school system, and new students from outside the state coming in at different times.

The IAP development teams defined the curriculum-relevant approach as follows:

- **Curriculum-Relevant:** meaning it’s not aligned to a specific curriculum but is built on a review of common themes, texts and domains found among the most commonly used curricula in Louisiana. The goal is to provide the same level of actionable, instructionally-oriented student information to educators as the ELA Guidebooks through-year assessment and to incentivize deep engagement in materials and texts throughout the year.

| 2022-2023 | In September 2022, USDOE announced the CGSA award to LDOE for "Curriculum Relevant" assessments under Absolute Priority 2 "focused on the development of comprehensive academic assessments that emphasize the mastery of standards and aligned competencies in a competency-based education model." This school year was the fifth year in operation under the IADA waiver and under the State’s innovative goals for knowledge building throughout the academic year. It was the first year that Wit and Wisdom field test and Curriculum Relevant pilot were delivered to students in school systems that have curriculum selected other than ELA Guidebooks. This implementation plan aligns with the State’s vision to preserve local control. Tests in Grades 6-8 were administered operationally for all three administrations (fall, winter, and spring) and field tests in Grade 5 for the ELA Guidebook and Wit and Wisdom curriculum. Under the IADA, grades 6-8 students from participating IAP school districts received valid scores from each testing window in order to receive the end-of-year report that LDOE uses for accountability purposes in lieu of a LEAP 2025 score. Extensive evidence for persons most proximate to the problem was gathered throughout the academic year as a requirement of the grants and LDOE’s practice of responsiveness to its stakeholders. In addition to this evidence, the State engaged stakeholders to learn more about interests and characteristics of Louisiana’s next generation of assessments. |
| 2023-2024 | All of LDOE’s innovative assessments under the IAP are at varying stages of development which are described further in Table 1. This school year marks the sixth year in operation under the IADA waiver and continues the IAP’s expansion to grades 3-8. In the 2023-2024 school year, the IAP includes the curriculum-relevant pilot for grade 5, field testing for grades 3-4, and operational assessments for grades 5-8. Twenty IAP school systems will administer assessments to over 27,000 students. |
The Grades 3-4 field test unit progression for the 23-24 school year will allow the IAP school systems to have two options in the Fall, two options in the Winter, and one option in the Spring. In Grades 5-8 the LEAP ELA Guidebooks will continue to be operational for school systems. However, LDOE had to make the tough decision to discontinue operational testing in the LEAP Wit and Wisdom program as LDOE was unable to recruit any school systems to participate. Instead, Wit and Wisdom will be included in the Crawfish Pilot for Grade 5.

The current and projected development is shown in the table below. This table is ordered in terms of operational implementation, with the operationally reported Grade 7 ELA Guidebooks in the first row, to the currently under development third option in the final row. As explained below, single summative scores were produced for the first time in 2021-2022 for the seventh grade ELA Guidebooks aligned assessment and used operationally with Louisiana’s compliant system of school identification and support (i.e., as part of the school performance score). This operational administration and subsequent reporting represents a major milestone for the IAP.

Table I-3. Detailed, High Level Summary of Activities

To find the details necessary for the Table I-3. Detailed, High Level Summary of Activities requirement, the reader should reference exhibit [1_003] Table I-3. Detailed, High Level Summary of Activities.
I.C. Participating IAP School Systems
If the innovative assessment system is not yet implemented statewide, provide a description of the SEA’s progress in scaling up the system to additional LEAs or schools.

See prior responses to section 1 above. In particular, the state was able to bring on two more grades, six and seven, during the 2022-23 school year, increasing operational participation by about 200%. In terms of the increases in enrollment, this involved two school systems joining the IAP. In addition, a number of schools and school systems who were participating in pilots or field tests moved those students into the operational tests. The difference from 2021-2022 to the 2022-2023 school year, in terms of schools and systems operationally testing is:

- **2021-2022 (gr. 7)**
  - 9 systems
  - 4760 students
- **2022-2023 (gr. 6-8)**
  - 16 systems
  - 17,000 students

In addition, to better inform the progress of scaling up the system, please provide:
- The list of LEAs that participated in the 2022-23 school year.
- For each participating LEA, the list of participating schools in 2022-23.
- For each participating school, the grade(s) and subject(s) in which the innovative assessment system was administered in 2022-23.
- The list of LEAs that will participate in the 2023-24 school year.
- For each participating LEA, the list of participating schools in 2023-24.
- For each participating school, the grade(s) and subject(s) in which the innovative assessment system will be administered in 2023-24 (a sample of the data structure is provided below; if the list of participating LEAs and schools is long, it may be submitted as an attachment).

### Table I.4 School System Participation for 22-23

The list of participating schools for the 2022-23 school year is provided within evidence [1_004] School System Participation for 22-23.
Table I.5 School System Participation for 23-24

The list of anticipated participating schools for the 2023-24 school year is provided within evidence [1_005] School System Participation for 23-24.

Provide any outcomes or results from its evaluation and continuous improvement process regarding the SEA’s progress in scaling up the system. This information may come from the State’s annual evaluation of its IADA assessment system. The information should include how data, feedback, evaluation results, and other information are used to improve the quality of the IADA assessment system (e.g., summary report of recommended changes from teachers/principals/school leaders, summary feedback from test administrator or scorer training, summary feedback from parent meetings). Please attach a copy of the annual evaluation.

LDOE is committed to the continuous evaluation and improvement of the IAP. As in prior years of implementation, LDOE and its partners collected data and feedback relevant to program improvement through a variety of methods, including:

- A student survey (see [4_013] Winter 2023 Student Survey Results)
- Nine focus groups aimed at soliciting feedback on report design, assessment design, as well as open feedback on implementation (see [4_014] October 2022 Initial Score Report Focus Group Slide Deck to [4_022] February 2023 Open Feedback Focus Group Notes)
- Regular “office hours” for participating teachers and leaders (see [4_026] Combined Office Hour Agendas and Notes)
- Teacher interviews at Louisiana's Teacher Leader Summit during the Summer of 2022 (see [4_027] Summary of Summit Teacher Interviews)
- Regular meetings of the IAP “Advisory Committee” composed of Louisiana teacher and leaders (see [4_025] Combined IAP Advisory Meeting Agendas and Participants)

These collections built on themes from prior years, especially those in 2019-2020 and 2021-22. Emerging themes included: (1) students wanting feedback about their learning while also having that information posed in a way that facilitated productive conversations with the adults in their lives (e.g., one student commented that “I just want my parents to be proud of me”), (2) teachers want to have better supports and clearer guidance for instructional next steps, and (3) teachers and leaders wanting to have as much information from score reports as possible. In addition, themes from the prior years were also present this year as well, including: (A) the need for more clear connections between curriculum and assessment, (B) a continuous test program across grades 3 to 8 (e.g., having the system just in middle school created a “disconnected system” between elementary and middle school, and (C) that the way in which other curriculum would be supported is unclear (i.e., while the
ELA Guidebook 2.0 curriculum is widely adopted, with the majority of districts and schools using the curriculum in middle school, not all schools and districts have adopted it).

To address these themes, LDOE and its partners:

- Developed new “enhanced” classroom score reports and student reports, that are connected to an updated version of the Teacher Guidance. **This updated guidance, and corresponding classroom report, is meant to help serve a critical intended use of the assessment results: to support instructional decision making going into the next unit.** As such, it has been the focus of extended development during the 2022-23 school year. This collaboration has included a partnership with LDOE, Great Minds, and the Center for Assessment. These materials will be used operationally in 2023-24.
- Continued to develop unit based assessments in elementary grades, specifically the field test of grade 5 content and the development of grades 3 and 4 content.
- Continued research and development on the curriculum-relevant design, which culminated in the pilot administration of a single curriculum-relevant form in Spring 2023. Analysis of the resulting data from this form is ongoing and is anticipated to be completed late 2023. Results are available on request.
- Successfully applied to a Carnegie Mellon University grant program, which will allocate a postdoctoral fellow to investigate the IAP during the 2023-24 and 2024-25 school years. In particular, this fellow will conduct summative evaluation activities.
Do you plan to administer the operational versions of the innovative assessments for some schools in the state, provide individual student reports, and use the results in state and local report cards and in the State’s federal accountability system in place of the regular state assessment for at least one grade and one subject area in 2023-2024?

Yes.

Do you plan to administer the operational versions of the innovative assessments for some schools in the state, provide individual student reports, and use the results in state and local report cards and in the State’s federal accountability system in place of the regular state assessment for at least one grade and one subject area in 2024-2025?

Yes.
II: Student Performance

II.A. Performance of Students in Participating Schools Disaggregated for Each Subgroup

A. Attach a report on the performance of students in participating schools at the State, LEA, and school level, for all students and disaggregated for each subgroup of students described in section 1111(c)(2) of the Act, on the innovative assessment, including academic achievement and participation data required to be reported consistent with section 1111(h) of the Act, except that such data may not reveal any personally identifiable information. Please be sure to include the subject area, the grade level(s), the number of students participating, the number of enrolled students, and % of students at each level of achievement for each school and LEA participating in the innovative assessment pilot.

Table II-A.5. School System Participation for 22-23

| The table summarizing student performance by school and grade, at the overall and subgroup level, are provided within [2_001] Student Performance by School and Subgroups within Schools 22-23. |

II.B. State-level Participation Rate of Students Disaggregated for Each Subgroup

B. Also provide the state-level participation rate of students, for all students and disaggregated for each subgroup of students described in section 1111(c)(2) of the Act, on the assessments required under section 1111(b)(2) of the Act for the grades and subjects that correspond to the operational innovative assessment administered in 2022-23.

Table II-B.6 Student Performance by School and Subgroups within Schools 22-23

| The table summarizing student performance by school and grade, at the overall and subgroup level, are provided within [2_001] Student Performance by School and Subgroups within Schools 22-23. |
III: School Demographic Information

III.A. School Demographic Information including Enrollment

III.A. If the innovative assessment system is not yet implemented statewide, attach school demographic information, including enrollment and student achievement information, for the subgroups of students described in section 1111(c)(2) of the Act, among participating schools and LEAs in the reporting year (2022-23).

Table III-A.7 Student Performance by School and Subgroups within Schools 22-23

Since the IAP is implemented at the school and grade level, the tables are identical to those provided above in II-A.5. As such, the table summarizing student performance by school and grade, at the overall and subgroup level, are provided within [2_001] Student Performance by School and Subgroups within Schools 22-23 provides the relevant information.

III.B. Schools Participating for the First Time in the 2023-24 School Year

III.B. For any schools or LEAs that will participate for the first time in the following year (2023-24), attach school demographic information, including enrollment information, for the subgroups of students described in section 1111(c)(2) of the Act, and describe how the participation of any additional schools or LEAs in that year contributed to progress toward achieving high-quality and consistent implementation across demographically diverse LEAs in the State consistent with the SEA’s benchmarks described in 34 CFR 200.106(a)(3)(ii).

Table III-B.8 School System Participation for 23-24

The table summarizing enrollment, at the overall and subgroup level, are provided within [2_002] School System Participation for 23-24. As shown in the 2019-2020 Annual Performance Report, the participating schools matched the distribution of students by subgroups for the state overall. The current sample simply expands the numbers of participating students, with the eventual goal of statewide implementation. In addition, much of the work on the IAP during the 2022-23 school year was focused both on increasing the overall participation, but also in increasing the number of grades and curricula covered.
IV: Consultation and Feedback

Describe feedback obtained during the reporting year (2022-23) from teachers, principals and other school leaders, and other stakeholders consulted, including parents and students, from participating schools and LEAs about their satisfaction with the innovative assessment system. Include a description of the method used to solicit the feedback (e.g., through surveys, focus groups, meetings) and the extent to which the feedback was solicited from each participating school and LEA.

Table IV-9. Description of Consultation Activities

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column).</th>
<th>Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary).</th>
</tr>
</thead>
</table>
| **Consultation.** Evidence that the SEA or consortium has developed an innovative assessment system in collaboration with-- (1) Experts in the planning, development, implementation, and evaluation of innovative assessment systems, which may include external partners; and | As noted in prior Annual Performance Reports, LDOE has continued to partner with a number of national organizations to develop and implement the innovative assessment program. Technical Partners:  
  ● The National Center for the Improvement of Educational Assessment, Inc (the Center)  
  ● NWEA  
  ● Progressive Measurement  

Content Partners:  
  ● Great Minds  
  ● Odell Education  
  ● John Hopkins University  

Platform and Scoring  
  ● MZ Development  
  ● Strategic Measurement and Evaluation |
<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column).</th>
<th>Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary).</th>
</tr>
</thead>
</table>
| TAC. LDOE maintains a technical advisory panel of recognized assessment experts. Presentations on key topics have included reporting, comparability and psychometric analysis. | [4_001] October 2022 TAC Presentation Slides  
[4_002] October 2022 TAC Notes  
[4_003] February 2023 TAC Presentation  
[4_004] February 2023 TAC Notes  
Note: The supporting materials that were provided to the TAC are not provided here, but are available on request. |
| Annual Meeting. In August 2023, LDOE convened its partners to consider the status of the program on a number of multiple aspects. During this two day long series of meetings, topics on scaling, assessment design and implementation were addressed | [4_005] August 2023 Annual Meeting Combined Slide Decks  
Note: meeting notes are available on request. The total set of notes totaled 75 pages, and are thus not included here. |
| Ed First Memos. As part of participation in a grant program, LDOE and its partners attended a number of learning sessions facilitated by Ed First, conducted research on the Curriculum-Relevant model, and authored a series of three memos. | [4_006] Ed First Memo #1  
[4_007] Ed First Memo #2  
[4_008] Ed First Memo #3  
Note: Materials from Ed First learning sessions available on request. |
| Collaboration on Reporting and Supports. An important part of the work in 2022-23 was to develop (1) “enhanced” classroom reports for teachers (2) materials that support teachers in using these materials. These materials were | This work involved close collaboration and feedback with Great Minds, who provided numerous points of feedback on the reports and supports, as well as feedback from focus groups with Louisiana teachers and leaders. |
## Requirement | Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column). | Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary). |
|---|---|---|
| | developed in partnership with curriculum experts at LDOE and Great Minds, as well as through focus groups with Louisiana teachers and leaders. | ● [4_009] Draft Enhanced Classroom Score Report with Comments from Great Minds  
● [4_010] Summary of Louisiana Teacher and Leader Focus Group Feedback on Draft Enhanced Classroom Score Reports |
| | As part of dissemination efforts, the scaling and comparability approach taken to support the IAP was documented in two conference papers and provided to Dr. Stephen Sireci. Dr Sireci provided potential recommendations and next steps for the work. | ● [4_011] 2023 NCME Conference Papers  
● [4_012] 2023 NCME Discussant Slide Deck |
| (2) Affected stakeholders in the State, or in each State in the consortium, including--  
(i) Those representing the interests of children with disabilities, English learners, and other subgroups of students described in section 1111(c)(2) of the Act;  
(ii) Teachers, principals, and other school leaders;  
(iii) Local educational agencies (LEAs);  
(iv) Representatives of Indian tribes located in the State;  
(v) Students and parents, including parents of children described in | During the 2022-23 school year, LDOE and its partners built on the extensive engagement work conducted in prior years, especially 2021-22, by examining more closely (1) potential designs for the Curriculum-Relevant model, (2) student and teacher preferences, ideas and feedback on score reports, and (3) overall feedback, value and satisfaction with the program.  
In 2022-23 LDOE and its partners chose to engage deeply on the above topics, generally with a core group of partner educators and leaders. This core group of partner educators and leaders are made up of multiple district groups, which are referred to as “IAP Collaboration Groups”. In addition to these | Student Survey  
● [4_013] Winter 2023 Student Survey Results  
● [7_001] Directions for Administering Window 2 Student Survey-Grades 6-8  
● [7_002] IAP System Grade 5 Survey Instructions  
Initial Score Report Focus Group  
This was a single focus group that examined a potential student score report, which was eventually phased out based on direction from the TAC.  
● [4_014] October 2022 Initial Score Report Focus Group Slide Deck  
● [4_015] October 2022 Initial Score Report Focus Group Notes |
<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column).</th>
<th>Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary).</th>
</tr>
</thead>
</table>
| paragraph (a)(2)(i) of this section; and (vi) Civil rights organizations.   | newly convened groups, LDOE and its partners continued to engage with the field through monthly office hours and the IAP Advisory Group, both of which were instituted during the 2021-2022 school year. | IAP Collaboration Groups: Score Report Focus Groups  
This was a series of four focus groups with teachers and leaders from the IAP Collaboration groups were provided with real data mockups of score reports and invited to give feedback.  
- [4.010] Summary of Louisiana Teacher and Leader Focus Group Feedback on Draft Enhanced Classroom Score Reports  
Note: focus group notes are available on request.  
IAP Collaboration Groups: Curriculum-Relevant Design Focus Groups  
This was a series of three focus groups with teachers and leaders from the IAP Collaboration groups. This focus group examined an early possible design for the curriculum-relevant model. The design focused on using topics studied through science and social studies as the key knowledge in the assessment. The groups greatly preferred the current IAP to the proposed curriculum-relevant model. As a result, the team moved away from this design.  
<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column).</th>
<th>Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary).</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAP Collaboration Groups: Open Feedback Focus Groups</td>
<td>This was a series of four focus groups with teachers and leaders from the IAP Collaboration groups which served as a listening session in which participants were asked to share whatever feedback they would like LDOE to hear.</td>
<td>Note: focus group notes are available on request.</td>
</tr>
<tr>
<td></td>
<td>● [4_022] February 2023 Open Feedback Focus Group Notes</td>
<td></td>
</tr>
<tr>
<td>In addition to these ongoing engagements, LDOE and its partners also engaged with those who represent the interests of parents by presenting at an annual conference of LOuisiana Parent and Family Engagement (PAFE) Coordinators, conducted interviews with implementing teachers during Lousianna’s annual teacher leader summit, and had a number of meetings with district leaders.</td>
<td>Additional Supporting Engagements</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● [4_023] Summit Teacher Interviews</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● [4_024] Parent and Family Engagement (PAFE) Conference Meeting Presentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regularly Occurring Engagements: Office Hours and Advisory Group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● [4_025] Combined IAP Advisory Meeting Agendas and Participants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● [4_026] Combined Office Hour Agendas and Notes</td>
</tr>
</tbody>
</table>
### Requirement

| Description of Consultation and Feedback Methods (be sure to describe the extent of consultation and method of obtaining feedback for each of the listed entities in the left-hand column). |
| Summary of Feedback of Stakeholders (note: you may attach artifacts of the actual feedback received in lieu of providing a summary). |

| A number of teachers presented at the Louisiana Teacher Leader Summit in the Summer of 2023. These teachers were interviewed on their perspectives and opinions on the IAP. | Summit Teacher Interviews  
- [4_027] Summary of Summit Teacher Interviews |

| Finally, in a separate process LDOE engaged with stakeholders at a broad scale to envision what the future of assessment might look like within the state. This work culminated in a report that summarized the contributions of numerous stakeholders. | The Future of Louisiana's State Assessment System Report  
- [4_028] The Future of Louisiana’s Assessment System |
V: Requirements for the Innovative Assessment System

V.A. Developing a Valid, Reliable, and Comparable System

Describe the process, procedures, or steps followed to develop a valid, reliable, and comparable innovative assessment system.

Table V-A.10. Evidence Innovative Assessment System is Valid, Reliable, and Comparable

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Information, Summary, Process, Procedures, or Steps (be sure to describe each activity listed in the left-hand column. You may attach artifacts in lieu of providing a description.)</th>
</tr>
</thead>
</table>
| Evidence that the SEA or consortium developed a valid, reliable, and comparable innovative assessment system. | As noted in the 2021-22 Annual Performance Report, the design of all of the innovative assessments in grades 5 to 8 all follow the same test specifications and blueprints. The test specifications and blueprints are detailed within chapter 2 of the [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21 (see pages 19-20, section 2.3). This technical documentation captures a variety of aspects of the program. It includes information typically found in a technical report, as well as additional information, like the program theory of action.  

- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially section 2.3 Assessment Design (pages 19-20)

Note that technical reports are typically released well after the operational administration, and this program is no exception. Thus the technical report provided here is for the 2021-2022 year, not the current 2022-22 year. Relevant technical information for the 2022-23 school year are contained in a number of additional supporting documents that are addressed later in this document - most importantly the [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary, [5_003] ALD Alignment, Scaling Methodology, and Outcomes Technical Summary, and [5_004] Comparability Analysis Memo.

As noted in the 2021-22 Annual Performance Report, two additional sets of specifications were developed during the 2021-22 school year to formalize the design approach taken in grades six to eight and ensure consistency across elementary and middle school development.  

- [5_005] Content Development Guide Specifications, especially the Item Pool Development Process and Strategy and Item Development Priorities sections
  - The Standards Guidance for Content Developers and Cognitive Complexity Considerations with the latter section may be particularly helpful
- [5_008] Passage Selection Specifications
2. IADA assessment development is guided by test specifications (e.g., purpose and intended uses; test format and length; info about content, psychometric characteristics of items and test; software and hardware requirements);

3. Descriptive information (e.g., feedback from item development reviews) and empirical evidence (e.g., item difficulty, item discrimination) that IADA item selection supports item specifications/blueprint;

The intended uses of the IAP are at a high level, to (1) inform instructional decision making as teachers transition from their first instructional unit to their section, and then from their second instructional unit to their third, (2) support year-to-year reflection and adjustments by teachers to their instruction, and (3) to support Louisiana’s Every Student Succeeds Act compliant system of school identification and support. These intended uses are documented with chapter 1 of the [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21 (see pages 7-12, section 1.3). These purposes are reflected within the unit based design. Each End-of-Unit assessment is designed to inform instruction, and then be combined with all other administered end-of-unit assessments to produce summative annual determinations (i.e., scale scores and achievement levels).

The development of the IAP unit assessments is guided by the test design outlined in section 2.3 of the [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21 (see pages 19-20). As explained within the 2021-22 Annual Performance Report, the IADA assessments reflect the same content emphasis as the current statewide assessment, in terms of the major content domains - Reading Informational, Reading Literary and Writing, albeit distributed across multiple unit assessments. In addition, the number of items and score points are slightly greater on the set of IADA pilot assessments. Additional information on the properties of the grades 6-8 forms during the 2023-23 school year can be found within the [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary.

- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially:
  - Section 2.3 Assessment Design (pages 19-20)
  - Section 2.4 Assessment Development (page 23)
  - Section 6.3.4 Descriptive Item Statistics (page 57)
- [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary
- [5_007] Grade 5 Educator Data Review and Range Finding Committee Materials. (The materials include the training for data review and the satisfaction exit surveys from educators from the rangefinding and data review workshop. Note: the rangefinding training materials are not included due to the secure nature of the teacher packets.)

In addition, materials previously provided within the 2021-22 Annual Performance Report are relevant here:
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.</td>
<td>Procedures to develop IADA item pool to support test specifications/blueprint (e.g., summary of crosswalk of item pool and test blueprint, algorithm used to select IADA items and how algorithm covers blueprint); As noted in the 2019-20 and 2021-22 Annual Performance Reports, the form assembly and item review processes are meant to ensure that the assessments reflect the standards emphasis embodied across the curriculum as well as reflective of the emphasis placed on the content domains by the statewide summative assessment. All of the currently developed items are on the operational, field test and try out forms, so examining the forms in terms of support of the specifications and blueprints is an examination of the item pool. Details on form construction and assembly are found in the Updated Content Brief and the Content Development Guide Specifications.</td>
</tr>
<tr>
<td>5.</td>
<td>Summary of IADA item specifications, by subject and grade (e.g., standards or targets to be assessed; item types, response format, and scoring; cognitive complexity; level of difficulty; accessibility tools and features); [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially: Section 2.3 Assessment Design (pages 19-20) [5_005] Content Development Guide Specifications, especially the Item Pool Development Process and Strategy and Item Development Priorities sections [5_008] Passage Selection Specifications These documents pertain to the development of the pilot and field test content in grades 3, 4 and 5. Information on the development of the content used within the grades 6, 7 and 8 operational administrations can be found within the 021-22 Annual Performance Report.</td>
</tr>
<tr>
<td>6.</td>
<td>Qualifications of item writers and reviewers (e.g., content expertise, experience); As noted in the 2019-20 and 2021-22 Annual Performance Reports, assessment content was developed by content experts at Odell Education and John Hopkins University in close partnership with content experts at LDOE. [5_009] Expert Curriculum Vitaes</td>
</tr>
<tr>
<td>7.</td>
<td>Instructions provided to develop and review IADA items, including instruction to align items to content standards, steps to ensure accessibility to students, and information about accessibility tools All of the content developed for the pilot during the 2022-23 operational administrations took place during the prior 2021-22 school year, and extensive documentation of that process can be found within the 2021-22 Annual Performance Report, especially for grade 5. The development process across all grades follows the process outlined within the [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, Section 2.4 Assessment</td>
</tr>
</tbody>
</table>
| and features; 8. Procedures to ensure IADA items adhere to IADA item specifications/blueprint; 9. Procedures to ensure content accuracy of IADA items; | Development (page 23), and provided in additional detail in the Content Development Guide Specifications, as well as the Passage Selection Specifications.  
- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially: Section 2.4 Assessment Development (page 23)  
- [5_005] Content Development Guide Specifications, especially the Item Pool Development Process and Strategy and Item Development Priorities sections  
- [5_008] Passage Selection Specifications |
| --- | --- |
| 10. Procedures to ensure the technical adequacy of IADA items (e.g., field and operational testing, thresholds for eliminating items, differential item functioning (DIF) analysis, statements that flagged items are appropriate for student subgroups); | The items with the operational administrations were field tested in the prior year, 2021-22. or in the case of grade 7 in the Winter and Fall of 2019. Items were reviewed in terms of content and data by teacher committees, and revised or rejected, as needed. The data review consisted of classical test theory statistics and flags based on these statistics, including p-value, item total and item distractor-total correlations, non-modal-key and differential item functioning flags. A high level overview of this process can be seen within the materials used within the [5_007] Grade 5 Educator Data Review and Range Finding Committees. After optional administration, items were removed based on both classical and item response theory statistics, if needed. Details on the analysis and removal of operational administrations is provided within [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary.  
- [5_007] Grade 5 Educator Data Review and Range Finding Committee Materials  
- [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary |
| 11. Procedures to ensure IADA items elicit intended response processes (e.g., cognitive labs, think-aloud sessions); | As explained within the 2021-22 Annual Performance Report, evidence related to intended response processes rests primarily on the specifications, which articulate the processes the items are meant to elicit, as well as the application and accuracy of the text and cognitive complexity frameworks.  
- [5_005] Content Development Guide Specifications  
- [5_010] Text Complexity Guide  
- [5_011] Achieve's Framework to Evaluate Cognitive Complexity in Reading Assessments |
| 12. Steps taken to consider potential bias in IADA items; 13. Steps taken to review IADA items for sensitivity and potential offensiveness (e.g., criteria for sensitivity, specifications and rules | As noted in the 2019-20 and 2021-22 Annual Performance Reports, issues of bias and sensitivity are addressed during (1) initial content development, (2) during rounds of review by LDOE content experts of the initial development, and (3) through bias and sensitivity review committees made up of Louisiana educators. |
14. Procedures to ensure all major content domains or strands assessed by IADA assessment are aligned to the IADA test specifications/blueprint

As noted in the 2019-20 and 2021-22 Annual Performance Reports, LDOE and Odell content experts reviewed and maintained a tracker during the development process with metadata to ensure alignment to the design and alignment to [ELA Louisiana Student Standards](https://www.louisiana.gov/cms/266485). This tracker helped ensure that the combinations of assessments taken by students was reflective of both the standards emphasis embodied across the curriculum and the emphasis placed on the content domains by the statewide summative assessment.

15. Process to reduce construct irrelevance (e.g., reduce inappropriate reading load, avoid use of idioms or culturally specific words).

As noted in the 2019-20 and 2021-22 Annual Performance Reports, the development and review process addresses construct irrelevant variance from the first stages of development by ensuring that the texts used for the warm-read section are appropriate for the grade level, based on the [Guide for Determining Text Complexity: Kindergarten through Grade 12](https://www.achieve.org/determining-text-complexity-kindergarten-through-grade-12), as well as the use of [Achieve's Framework for Cognitive Complexity](https://www.achieve.org/framework-for-cognitive-complexity). The use of the framework is called out with the [Cognitive Complexity Considerations](https://www.achieve.org/framework-for-cognitive-complexity) subsection of the [Content Development Guide Specifications](https://www.achieve.org/framework-for-cognitive-complexity).

Construct irrelevant variance is also reduced throughout the item design through the use of an Item Review Checklist through the development processes, as well as the specific steps in the item review process detailed with the [Item Development Priorities](https://www.achieve.org/framework-for-cognitive-complexity) section of the [Content Development Guide Specifications](https://www.achieve.org/framework-for-cognitive-complexity).

- [5_005] Content Development Guide Specifications
- [5_010] Text Complexity Guide
- [5_011] Achieve's Framework to Evaluate Cognitive Complexity in Reading Assessments
- [5_007] Grade 5 Eductor Data Review and Range Finding Committee Materials
### V.B. Update on Meeting Requirements

Please provide a brief report on the required elements of the Innovative Assessment System. This brief report is intended to update the State’s demonstration that the innovative assessment system does or will meet the requirements of section 1111(b)(2)(B).

#### Table V-B.11. Brief Report that the Innovative Assessment System Meets the Requirements of Section 1111(b)(2)(B)

<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
</table>
| **Innovative assessment system. A demonstration that the innovative assessment system does or will**--  
(2)(i) Align with the challenging State academic content standards under section 1111(b)(1) of the Act, including the depth and breadth of such standards, for the grade in which a student is enrolled; and | Following the same process used in 2021-22, LDOE’s partners, specifically the Center for Assessment and NWEA, conducted an Achievement Level Descriptor (ALD) alignment workshop during the summer of 2023. This workshop examined the degree of alignment between the IAP items and the ALDs of the statewide program. Whereas in the summer of 2022 only grade 7 was investigated, in 2023 grades 6 to 8 were investigated by parallel groups of educators facilitated by LDOE and its partners.  
- [5.003] Achievement Level Descriptor (ALD) Alignment and Reporting Category Workshop for Grades 6, 7, 8 |                                                                                                                                                                                                                                           |
<p>| (ii) May measure a student’s academic proficiency and growth using items above or below the student’s grade level so long as, for purposes of meeting the requirements for reporting and school accountability under sections 1111(c) and 1111(h) of the Act and | N/A - IAP does not use off grade items                                                                                                                                                                                                       |                                                                                                                                                                                                                                           |</p>
<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>paragraphs (b)(3) and (b)(7)-(9) of this section, the State measures each student’s academic proficiency based on the challenging State academic standards for the grade in which the student is enrolled;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Express student results or competencies consistent with the challenging State academic achievement standards under section 1111(b)(1) of the Act and identify which students are not making sufficient progress toward, and attaining, grade-level proficiency on such standards;</td>
<td>As noted in the 2021-22 Annual Performance Report, the results of the IAP program are reported in terms of scale scores and achievement levels from the statewide assessment. That is, the IAP items were placed onto the scale of the current statewide assessment, known as LEAP, via a common item non-equivalent group design in which statewide item sets were embedded into the IAP unit assessments.</td>
<td></td>
</tr>
<tr>
<td>(4)(i) Generate results, including annual summative determinations as defined in paragraph (b)(7) of this section, that are valid, reliable, and comparable for all students and for each subgroup of students described in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, to the results generated by the State academic assessments</td>
<td>As mentioned above, the results of the IAP program are reported in terms of scale scores and achievement levels from the statewide assessment.</td>
<td></td>
</tr>
</tbody>
</table>

- [5.001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6, pages 49-62
- [5.002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary
- [5.014] Example End-of-Year Student Score Report
<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
</table>
| described in 34 CFR 200.2(a)(1) and section 1111(b)(2) of the Act for such students. Include:  
1. Objective nature of IADA items machine scoring (e.g., scoring rule limits for number of errors, scoring rules for technology-enhanced score capture and validity checking, how artificial intelligence (AI) scoring engine is trained and its accuracy);  
2. Procedures to transform raw IADA scores to scale scores (overall and by subtest); | Item scoring is subject to extensive pre- and post-test validation to ensure the accuracy of scoring. Documentation of this process is available on request.  
Following the approach developed in 2021-22, within each grade, 6-8, the operational ELA Guidebooks unit assessments were combined into a single item response matrix, then scaled using the two parameter logistic and generalized partial credit item response theory models as if they were from a single test. This “pooling” approach essentially weights each test equally, which matches LDOE’s values in terms of equal emphasis on student performance on each unit. This equal emphasis is similar to typical classroom grading practices wherein students are given end of unit quizzes.  
- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6, pages 49-62  
- [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary |
<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Description of IADA equating process (overall and, if appropriate, by subtest), including equating study design, statistical methods used and person parameters, overall information functions, size and relevant characteristics of examinee samples, characteristics of anchor items/test, and accuracy of equating functions;</td>
<td>Following the approach developed in 2021-22, within each grade, 6-8, the combined item response matrix was placed on the statewide scale using a non-equivalent groups common item design that used passage sets and associated items as an external anchor, one per window. This is explained within section chapter 6 of the [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6. To do so, the item parameters for the anchor items were fixed to the values from the statewide program and then the item parameters were estimated for the IAP items (i.e., fixed calibration item parameter approach). Details for the current year can be found within the [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary. Note that resulting IADA scores are equated across years based solely on the link to the anchor items. That is, there is no unique IAP scale, instead the IAP is completely reliant on the statewide scale for year-to-year equating.</td>
<td></td>
</tr>
<tr>
<td>4. Process to equate IADA scores across academic years;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6, pages 49-62
## 2023 IADA Annual Performance Report

<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23)</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. IADA assessment form equivalence, by grade and subject (e.g., raw scores and p-values, standard error of measurement (SEM), dimensionality, test characteristic curve (TCC), test information function (TIF), conditional standard error of measurement (CSEM), score distributions);</td>
<td>Within the IAP, unique patterns of end of unit assessments constitute forms. These forms are considered within the [5.002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary.</td>
<td>Additional detail on the measurement precision can be found within the [5.001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 5, Scoring (pages 43-48) and Chapter 7, Reliability and Precision (pages 63-68).</td>
</tr>
<tr>
<td>6. Indication that the TCC or TIF for all IADA tested grades and subjects is reasonable (overall and, if appropriate, by subtest);</td>
<td>[5.002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary</td>
<td></td>
</tr>
<tr>
<td>7. Indication that CSEM or SEM for all IADA tested grades and subjects is reasonable (overall and, if appropriate, by subtest) (e.g., CSEM for each IADA interim assessment and final assessment for the entire scale or at cut scores, overall estimate of test error);</td>
<td>Additional detail on the measurement precision can be found within the [5.001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 5, Scoring (pages 43-48) and Chapter 7, Reliability and Precision (pages 63-68).</td>
<td></td>
</tr>
<tr>
<td>8. xReliability estimates, including, as appropriate:</td>
<td>[5.001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 5, Scoring (pages 43-48) and Chapter 7, Reliability and Precision (pages 63-68).</td>
<td></td>
</tr>
<tr>
<td>a. Reliability estimate for entire IADA student population (e.g., alpha coefficient)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Reliability estimate for each reported IADA subgroup (e.g., alpha coefficient)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Reliability estimate for summative assessment for all pilot students and each reported subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Reliability estimate for interim assessments for all pilot students and each reported subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Interrater reliability estimate for each reported dimension for all pilot students and each reported subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulatory Requirement</td>
<td>Accomplishments in the Reporting Year (2022-23).</td>
<td>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>f. Cohen’s Kappa for all pilot students and each reported subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Decision consistency and accuracy reliability estimates of student classifications based on IADA cut scores, classification accuracy conditioned on achievement level, and classification consistency conditioned on achievement cut points,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Reliability estimates of correctly classified and incorrectly classified students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Procedures to ensure use of simple language and uniform format in IADA score reports;</td>
<td>The End-of-Unit score reports for the 2022-23 school year were the same as those used during the 2021-22 school year. These End-of-Unit score reports underwent numerous revisions for clarity of communications, which included multiple rounds of feedback from LDOE and its partners.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_014] End-of-Unit Score Report Example</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The End-of-Year score reports were largely unchanged from the 2021-22 school year, however, the reporting categories classifications (i.e., subscores in the form of classifications based on the Knowledge, Application and Synthesis categories) were revised to provide three levels of differentiation with richer descriptive language drawn from the ALD workshop.</td>
<td></td>
</tr>
</tbody>
</table>
## Regulatory Requirement

### Accomplishments in the Reporting Year (2022-23).

- [5.015] End-of-Year Score Report Example
- [5.003] Achievement Level Descriptor (ALD) Alignment and Reporting Category Workshop for Grades 6, 7, 8

The End-of-Unit and End-of-Year score reports underwent numerous revisions for clarity of communications, which included multiple rounds of feedback from LDOE and its partners.

### Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).

#### 10. Availability of and access to translations who require accommodations to interpret IADA scores/results;

As noted in the 2022-22 Annual Performance Report, following the approach used within the statewide assessment program, accommodated reports are provided to families who request them during one on one meetings with school staff. Parent Guides to the student reports are translated into French, Spanish, Arabic, and Vietnamese.

- [5.015] Translated Parent Guides

#### 11. State generates annual State, district, and school IADA assessment reports;

LDOE and its partners generated individual student reports as well as summary files for use by schools and districts. In addition, the results were used within the state’s system of reporting (i.e., reported in an identical fashion to scores from the statewide assessment program) and the state’s system of school identification and support.

These results were disseminated to the field through the individual End-of-Unit and
2023 IADA Annual Performance Report

<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. Annual IADA assessment reports include student performance related to content and knowledge of assessed standards (e.g., scale scores); academic content descriptions of what students can and cannot do using achievement level descriptors (ALDs), performance level descriptors (PLDs), content knowledge learning maps or networks (e.g., subscores); and information to facilitate interpreting results and addressing specific academic needs of students (e.g., itemized score analyses);</td>
<td>As noted within the 2021-22 Annual Performance Report, the End-of-Unit score report is meant to provide information in a way that is transparent to teachers three times a year. This report is supported with guidance through the [5_019] IAP Guidance for Score Reports, which attempts to bridge the gap between the information found within the reports and instructional decision making. The End-of-Year score reports provide information based on the IRT scaling, and thus provides a student scale score and achievement levels school and IAP average scale scores and achievement levels, reporting category classifications with information about what students know and can do,</td>
<td></td>
</tr>
</tbody>
</table>

End-of-Year score reports, within the state’s public facing reports as well as within summary spreadsheet files provided by LDOE directly to teachers, schools and districts. Accompanying the release of each set of reports, End-of-Unit or End-of-Year, LDOE conducted a score report webinar to provide notice to the field.

- [5_016] IAP Guidebooks Window 1 Reports Webinar
- [5_017] IAP Guidebooks Window 2 Reports Webinar
- [5_018] IAP Guidebooks Window 3 Reports Webinar
<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>and percent of students at each achievement level for the school, school system and overall IAP program.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_020] End-of-Unit Score Report Example</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_021] End-of-Year Score Report Example</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_019] IAP Guidance for Score Reports</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_022] IAP Interpretive Guide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_023] LEAP ELA Guidebooks IAP Operational Assessment Guide - Grades 5, 6, 7, 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [7_003] LEAP ELA Assessment Guide - Grade 5 Innovative Assessment Program Field Test</td>
<td></td>
</tr>
<tr>
<td>13. State documents that IADA assessments in each relevant grade and subject were used to inform the annual determination of achievement for all participating students;</td>
<td>As noted within the 2021-22 Annual Score Report, the Board of Elementary and Secondary Education (BESE) bulletin 111 states that IAP scores are used to inform the annual determination of achievement for all students.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• [5_024] Bulletin 111 Section 307</td>
<td></td>
</tr>
<tr>
<td>14. Annual IADA student assessment reports include indicator of annual IADA proficiency or summative achievement determination; indicators of annual student progress (e.g., subscores, ALDs or PLDs, learning maps); and indicators for identifying</td>
<td>See V-B(12) above, i.e., item 12 within this section.</td>
<td></td>
</tr>
</tbody>
</table>
## Regulatory Requirement

<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>students not making progress (e.g., subscores on student report);</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Annual IADA school report includes summative achievement results disaggregated by</td>
<td>As noted for point 11 above, the results were designated within the state’s public facing reports as well as within summary spreadsheet files provided by LDOE directly to teachers, schools and districts. Both of these approaches to reporting allow for disaggregated reporting.</td>
<td></td>
</tr>
<tr>
<td>important subgroups;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Annual IADA district and State reports, with both including summative achievement of</td>
<td>Expectations for release of results were communicated within each of the score reporting webinars.</td>
<td></td>
</tr>
<tr>
<td>annual progress for all IADA pilot students and for important IADA pilot student</td>
<td>- [5_016] IAP Guidebooks Window 1 Reports Webinar</td>
<td></td>
</tr>
<tr>
<td>subgroups;</td>
<td>- [5_017] IAP Guidebooks Window 2 Reports Webinar</td>
<td></td>
</tr>
<tr>
<td>17. Expectations from State of timeline for releasing individual student IADA reports to</td>
<td>- [5_018] IAP Guidebooks Window 3 Reports Webinar</td>
<td></td>
</tr>
<tr>
<td>schools and districts;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Expectations from State and district for delivering student IADA score reports to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parents;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Procedures to protect security of IADA assessment personally identifiable information</td>
<td>As noted within the 2021-22 Annual Score Report, procedures to protect personally identifiable information (PII) are articulated within the Test Coordinator and Test Administration Manuals. Only the manuals from Window 3 have been provided here, but manuals from the other windows are available on request. In addition, LDOE provides support to school systems to ensure that they are protecting personally identifiable information (PII).</td>
<td></td>
</tr>
<tr>
<td>(e.g., staff procedures, letter to parents, scoring manual).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Regulatory Requirement

<table>
<thead>
<tr>
<th>Accomplishments in the Reporting Year (2022-23).</th>
</tr>
</thead>
</table>
| **Grades 6, 7 and 8 Operational Manuals**
  - [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3
  - [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3
| **Grade 5 Field Test Manuals**
  - [5_027] LEAP ELA Guidebooks Grade 5 Field Test Coordinator Manual - Window 3
  - [5_028] LEAP ELA Guidebooks Grade 5 Field Test Administration Manual - Window 3
| **Data Governance & Student Privacy Guidebook**
  - [5_029] Louisiana's Data Governance & Student Privacy Guidebook

### Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).

Consistent with the SEA’s or consortium’s evaluation plan under 34 CFR 200.106(e), the SEA must plan to annually determine comparability during each year of its demonstration authority period in **one of the following ways:**

(C) Including, as a significant portion of the innovative assessment system in each required grade and subject in which both an innovative and statewide

During the 2022-23 school year, the approach to scaling, scoring and comparability that was operationally implemented in 2021-22 for grade 7 was expanded to grades 6, 7 and 8. As explained within the 2021-22 Annual Performance Report, evidence of comparability is based on three key sources of evidence:
<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>assessment are administered, items or performance tasks from the statewide assessment system that, at a minimum, have been previously pilot tested or field tested for use in the statewide assessment system. (ii) Generate results, including annual summative determinations as defined in paragraph (b)(7) of this section, that are valid, reliable, and comparable, for all students and for each subgroup of students described in 34 CFR 200.2(b)(11)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, among participating schools and LEAs in the innovative assessment demonstration authority. Consistent with the SEA’s or consortium’s evaluation plan under 34 CFR 200.106(e), the SEA must plan to annually determine comparability during each year of its demonstration authority period;</td>
<td>1. Similar functioning, i.e., invariance of the anchor items from the statewide test are embedded within the IAP (see [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary). 2. Empirical comparisons of student performance on the IAP to that of the statewide assessment. Specifically, these comparisons looked at multiple aspects of the resulting scale scores and achievement levels for students participating in the IAP, relative to historical and contemporaneous groups of students who did not participate in the IAP (see [5_004] Comparability Analysis Memo). 3. Expert judgment about the alignment between IAP items and the achievement levels on the statewide assessments, captured during the Achievement Level Descriptor Alignment workshop that was conducted during the summer of 2023 (see [5_003] ALD Alignment, Scaling Methodology, and Outcomes Technical Summary). Much of this logic is also documented within the Chapter 6 Scaling and Linking of the [5_001] Progress Report.</td>
<td></td>
</tr>
</tbody>
</table>
### Regulatory Requirement

In addition to providing the information noted above, be sure to include the following information:

1. Evidence that IADA test results are comparable to those from the non-IADA system (e.g., provide within-grade IADA and non-IADA results for participating districts are comparable, student proficiency classification for IADA and non-IADA districts are comparable in terms of complexity included in each achievement level, comparability results align with expectations outlined in State’s theory of action);
2. Description of across-years scaling procedures to transform IADA raw scores to scale scores; and
3. Description of across-years IADA equating process that includes design of equating study; statistical methods used and person parameter, and overall information functions; size and relevant characteristics of examinee samples; characteristics of anchor items/test; and accuracy of equating functions.

### Accomplishments in the Reporting Year (2022-23).

- [5_002] IAP Grades 6-8 ELA Guidebooks 22-23 Scaling Summary
- [5_004] Comparability Analysis Memo
- [5_003] ALD Alignment, Scaling Methodology, and Outcomes Technical Summary
- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6, Linking and Scaling (pages 49-62)

### Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).

(5)(i) Provide for the participation of all students, including children with disabilities and English learners;

(ii) Be accessible to all students by incorporating the principles of universal design for learning, to the extent practicable, consistent with 34 CFR 200.2(b)(2)(ii); and

As noted in the 2020-21 and 2021-22 Annual Performance Reports, the IAP’s approach to maximizing participation and accessibility is rooted in a Universal Design for Learning approach, which is summarized in Chapter 4, Accessibility and Accommodations, of the [5_001] Progress Report (pages 29-42).

- [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 6, Linking and Scaling (pages 49-62)
### Regulatory Requirement

(iii) Provide appropriate accommodations consistent with 34 CFR 200.6(b) and (f)(1)(i) and section 1111(b)(2)(B)(vii) of the Act;

### Accomplishments in the Reporting Year (2022-23).

Language Arts Grade 7, School Year 2020-21, especially *Chapter 4, Accessibility and Accommodations*, (pages 29-42)

Test accommodations are provided to minimize the effects of a disability to ensure students can demonstrate the degree of achievement they actually possess. Accommodations are listed in the testing manuals. See, for example, the operational manuals:

- [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Field Test Coordinator Manual - Window 3
- [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Field Test Administration Manual - Window 3

### Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).

Scores from the IAP are treated as interchangeable with scores from the statewide assessment, and are therefore combined with the results of the statewide assessment for the purposes of Louisiana’s Every Student Succeeds Act compliant system of school identification and support. Therefore, schools participating in the IAP are subject to the same participation requirements as all other schools. Consistent with our ESSA plan Louisiana incentivizes the highest participation rate.

<table>
<thead>
<tr>
<th>Regulatory Requirement</th>
<th>Accomplishments in the Reporting Year (2022-23).</th>
<th>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</th>
</tr>
</thead>
<tbody>
<tr>
<td>(iii) Provide appropriate accommodations consistent with 34 CFR 200.6(b) and (f)(1)(i) and section 1111(b)(2)(B)(vii) of the Act;</td>
<td>Language Arts Grade 7, School Year 2020-21, especially <em>Chapter 4, Accessibility and Accommodations</em>, (pages 29-42)</td>
<td>Test accommodations are provided to minimize the effects of a disability to ensure students can demonstrate the degree of achievement they actually possess. Accommodations are listed in the testing manuals. See, for example, the operational manuals:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Field Test Coordinator Manual - Window 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Field Test Administration Manual - Window 3</td>
</tr>
<tr>
<td>(6) For purposes of the State accountability system consistent with section 1111(c)(4)(E) of the Act, annually measure in each participating school progress on the Academic Achievement indicator under section 1111(c)(4)(B) of the Act of at least 95 percent of all students, and 95 percent of students in each subgroup of students described in section 1111(c)(2) of the Act, who are required to take such assessments consistent with paragraph (b)(1)(ii) of this section;</td>
<td>Scores from the IAP are treated as interchangeable with scores from the statewide assessment, and are therefore combined with the results of the statewide assessment for the purposes of Louisiana’s Every Student Succeeds Act compliant system of school identification and support. Therefore, schools participating in the IAP are subject to the same participation requirements as all other schools. Consistent with our ESSA plan Louisiana incentivizes the highest participation rate</td>
<td></td>
</tr>
<tr>
<td>Regulatory Requirement</td>
<td>Accomplishments in the Reporting Year (2022-23).</td>
<td>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(7) Generate an annual summative determination of achievement, using the annual data from the innovative assessment, for each student in a participating school in the demonstration authority that describes—</td>
<td>possible by forcing a zero in accountability for non-testers that otherwise were expected to test.</td>
<td></td>
</tr>
<tr>
<td>(i) The student’s mastery of the challenging State academic standards under section 1111(b)(1) of the Act for the grade in which the student is enrolled; or</td>
<td>Annual summative determinations were created following the process noted previously (e.g., see V-B(2) above, which explains the pooling approach used to create an “overall” estimate of student performance on the theta metric, which was then transformed into scale scores and achievement levels on the statewide scale.</td>
<td></td>
</tr>
<tr>
<td>(ii) In the case of a student with the most significant cognitive disabilities assessed with an alternate assessment aligned with alternate academic achievement standards under section 1111(b)(1)(E) of the Act, the student’s mastery of those standards;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) Provide disaggregated results by each subgroup of students described in 34 CFR 200.2(b)(1)(i)(A)-(I) and sections 1111(b)(2)(B)(xi) and 1111(h)(1)(C)(ii) of the Act, including timely data for teachers, principals and other school leaders, students, and parents consistent with 34 CFR 200.8 and section 1111(b)(2)(B)(x) and (xii) and section 1111(h) of the Act, and provide results to parents in a manner consistent with paragraph (b)(4)(i) of this section and part 200.2(e);</td>
<td>As explained in V-B(4)(i)(11) above, LDOE and its partners provided individual student reports, data spreadsheets and state-level reporting to support disaggregation and subsequent follow up by key stakeholders.</td>
<td></td>
</tr>
<tr>
<td>(9) Provide an unbiased, rational, and consistent determination of progress toward the State’s long-term goals for academic achievement under section</td>
<td>As explained within the 2021-22 Annual Performance Report, the IAP scale scores and classifications are used within the state’s system in</td>
<td></td>
</tr>
<tr>
<td>Regulatory Requirement</td>
<td>Accomplishments in the Reporting Year (2022-23).</td>
<td>Explanation of Delays or Concerns, with a description of a plan to resolve the concern (if applicable).</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------------------</td>
<td>---------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 1111(c)(4)(A) of the Act for all students and each subgroup of students described in section 1111(c)(2) of the Act and a comparable measure of student performance on the Academic Achievement indicator under section 1111(c)(4)(B) of the Act for participating schools relative to non-participating schools so that the SEA may validly and reliably aggregate data from the system for purposes of meeting requirements for--

(i) Accountability under sections 1003 and 1111(c) and (d) of the Act, including how the SEA will identify participating and non-participating schools in a consistent manner for comprehensive and targeted support and improvement under section 1111(c)(4)(D) of the Act; and

(ii) Reporting on State and LEA report cards under section 1111(h) of the Act. | the exact same fashion as scores and classifications from the current statewide summative assessment. The IAP scores therefore provide the same degree of (a) unbiased, rational, and consistent determination of progress toward the State’s long-term and (b) support to meet the requirements of the state’s system of school identification and support. |
VI: Training on and Familiarization with the Innovative Assessment System

VI.A. Statewide Training for Standard Administration of the Innovative Assessment System

Describe training provided to teachers, principals and other school leaders, and other stakeholders during the reporting year (2022-23) to implement the innovative assessment system, including the standard administration of the innovative assessments.

Table VI-A.12. Training Evidence for Standard Administration of the Innovative Assessment System

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Training (be sure to describe the training provided for each activity listed in the left-hand column. You may attach artifacts of the training in lieu of providing a description).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training. Evidence that the SEA or consortium provided training or instructions for standard administration of the innovative assessment system on each of the following activities:</td>
<td>As noted within the 2021-22 Annual Performance Report, procedures for administering the IADA assessments were presented to the participating districts prior to each administration. Changes to any administration procedures, if made, were highlighted. Slide decks were shared with participating districts. These procedures are documented in Test Coordinator Manuals and Test Administration Manuals including accommodations for students with disabilities and English learners. Further support is provided through an online help desk that provides on demand support, as well as questions and answers on a number of topics.</td>
</tr>
<tr>
<td>● Standard procedures for administering the IADA assessments (e.g., manual, slides);</td>
<td></td>
</tr>
<tr>
<td>● Administering IADA assessment supports and accommodations to students with disabilities;</td>
<td></td>
</tr>
<tr>
<td>● Administering IADA assessment supports and accommodations to English learners;</td>
<td></td>
</tr>
<tr>
<td>Grades 6, 7 and 8 Operational Manuals</td>
<td>[5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3</td>
</tr>
<tr>
<td></td>
<td>[5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3</td>
</tr>
<tr>
<td>Grade 5 Field Test Manuals</td>
<td>[5_027] LEAP ELA Guidebooks Grade 5 Pilot Test Coordinator Manual - Window 3</td>
</tr>
<tr>
<td></td>
<td>[5_028] LEAP ELA Guidebooks Grade 5 Pilot Administration Manual - Window 3</td>
</tr>
<tr>
<td>Training Webinars</td>
<td>[6_001] Window 1 Adam Training Webinar</td>
</tr>
<tr>
<td>Requirement</td>
<td>Description of Training (be sure to describe the training provided for each activity listed in the left-hand column. You may attach artifacts of the training in lieu of providing a description).</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>● Hand-scoring constructed responses or essays (e.g., results of exact, adjacent, and discrepant agreement; validity check results; number of read-behind flags);</td>
<td><strong>Help Desk Materials</strong></td>
</tr>
<tr>
<td></td>
<td>● [6_004] Zen Desk Test Administration Articles</td>
</tr>
<tr>
<td></td>
<td>Similar to the evidence provided in the 2021-22 Annual Performance Report, the process for the hand scoring of the constructed response and essay items, i.e., range finding, anchor set construction, training and monitoring, can be found in [6_005] Hand Scoring Guidelines and Process and [6_005] Hand Scoring Guidelines and Process. Additional detail is provided in Chapter 5, Scoring, of the [5_001] Progress Report (pages 43-48), but again, this detail is from the prior school year.</td>
</tr>
<tr>
<td></td>
<td>● [6_005] Hand Scoring Guidelines and Process</td>
</tr>
<tr>
<td></td>
<td>● [6_014] Hand Scoring Agreements</td>
</tr>
<tr>
<td></td>
<td>● [5_001] Progress Report: Louisiana Innovative Assessment Program English Language Arts Grade 7, School Year 2020-21, especially Chapter 5, Scoring, (pages 43-48)</td>
</tr>
<tr>
<td>● Handling test irregularities during IADA assessment administrations (e.g., test security handbook, test security plan, reports of internal or independent monitoring procedures);</td>
<td><strong>Grades 6, 7 and 8 Operational Manuals</strong></td>
</tr>
<tr>
<td></td>
<td>● [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3</td>
</tr>
</tbody>
</table>
### Requirement | Description of Training (be sure to describe the training provided for each activity listed in the left-hand column. You may attach artifacts of the training in lieu of providing a description).
---|---
- Conducting external reviewing of IADA items for potential bias (e.g., criteria for review, steps where potential bias is considered, review by external review committee); | During the 2022-23 school year, LDOE conducted a review of grade 5 field tests with a group of educators. Review of other unit assessments was conducted during the Summer of 2021 and 2022 and detail on these reviews can be found in Louisiana's 2019-20 and 2021-22 Annual Performance Reports, respectively.
  - [5_007] Grade 5 Educator Data Review and Range Finding Committee Materials
  - [5_012] IAP Guidebooks Grades 3 and 4 Passage Review Training
  - [5_013] Level 1 ELA Guidebooks Grades 3 and 4 2023 Item Review Presentation
- Reviewing IADA items for sensitivity and potential offensiveness (e.g., criteria for review, specifications and rules followed, list of reviewers and expertise); | As explained within the 2021-22 Annual Performance Report, procedures to protect personally identifiable information (PII) are articulated within the Test Administration Manuals and Test Administration Manuals. Only the manuals from Window 3 have been provided here, but manuals from the other windows are available on request. In addition, LDOE provides support to school systems to ensure that they are protecting personally identifiable information (PII).
- Protecting IADA-related personally identifiable information (PII). | Grades 6, 7 and 8 Operational Manuals
  - [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3
  - [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3

#### Grade 5 Field Test Manuals
- [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3

#### Grades 6, 7 and 8 Operational Manuals
- [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3
- [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3

#### Grade 5 Field Test Manuals
<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description of Training (be sure to describe the training provided for each activity listed in the left-hand column. You may attach artifacts of the training in lieu of providing a description).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Governance &amp; Student Privacy Guidebook</td>
<td>- [5_029] Louisiana's Data Governance &amp; Student Privacy Guidebook</td>
</tr>
</tbody>
</table>
**VI.B. Describe All Innovative Assessment Training Opportunities**

For each of the training topics below, briefly describe all training opportunities that your State provided for teachers, principals, and other school leaders during the reporting year (2022-23). For each training opportunity, report the number of individuals eligible to participate and the number of individuals who actually participated.

Table VI-B.13. Describe All Training Opportunities in the State

<table>
<thead>
<tr>
<th>Training Topic</th>
<th>Brief Description of Training Opportunity, Including How Eligibility for the Training was Defined. (You may attach artifacts of the training in lieu of providing a description, such as training slides, sections, or an entire manual).</th>
<th>Number of Eligible Participants by Type</th>
<th>Number of Actual Participants by Type</th>
</tr>
</thead>
</table>
| (1) Training to familiarize teachers or school staff with the innovative assessment system (e.g., training on goals of innovative assessment system design including alignment to State standards for student learning, highlights of the key differences between the new and existing assessment systems, format, timeline for administration, and reporting) | LDOE provided training for teachers and school staff in a variety of ways throughout the 2022-23 school year. These trainings included:  
  - Administration Webinars  
    LDOE hosted administration webinars prior to each window to inform the field on appropriate administration methods, which also included information on the purpose and goals of the program.  
    - [6_001] Window 1 Adam Training Webinar Slides (see slide 5 for an overview of purpose, slides 7 and 66 for key dates for administration; similar slides exist within the other two training webinar slide decks)  
    - [6_002] Window 2 Adam Training Webinar Slides  
    - [6_003] Window 3 Adam Training Webinar Slides  
    See also the training materials summarized in VI-A.12, in particular the administration manuals and help desk materials. | All participating IAP systems-23 school systems; any school system leader could attend | 21-35 school system leaders (Primary attendance: District and School Test Coordinators) |
| Office Hours                                                                 | On the third Wednesday of each month, LDOE continued open office hours this year, which were introduced during the 2021-22 | All participating IAP systems as well as other | Attendance ranged from 15 to 25 teachers |
### Training Topic

<table>
<thead>
<tr>
<th>Brief Description of Training Opportunity, Including How Eligibility for the Training was Defined. (You may attach artifacts of the training in lieu of providing a description, such as training slides, sections, or an entire manual).</th>
<th>Number of Eligible Participants by Type</th>
<th>Number of Actual Participants by Type</th>
</tr>
</thead>
</table>
| School year. Teachers and leaders from participating districts could attend to learn about various topics related to the innovative assessment, as well as ask questions.  
- [4_026] Combined Office Hour Agendas and Notes | Interested systems–20-25 school systems | and leaders per session |
| **LDOE Led Professional Development for Districts by Request**  
At the request of two districts, LDOE facilitated professional development sessions on the purposes of the IAP and how to connect their instruction to the IAP.  
- [6_006] Assumption Parish Professional Development  
- [6_007] Central Community Schools Professional Development  
- [6_008] Central Community Schools Professional Development - Writing Focused | 30-40 Educators and School Leaders | 30-40 Educators and School Leaders |
| **Reporting Webinars**  
LDOE hosted reporting webinars in between the close of an administration and the release of score reports to inform the field on how to appropriately interpret the results of the report as well as trends. This presentation also included information on the purpose and goals of the program.  
- [5_016] IAP Guidebooks Window 1 Reports Webinar Slides  
- [5_017] IAP Guidebooks Window 2 Reports Webinar Slides | All participating IAP systems-23 school systems; any school system leader could attend | 21-35 school system leaders (Primary attendance: District and School Test Coordinators) |
<table>
<thead>
<tr>
<th>Training Topic</th>
<th>Brief Description of Training Opportunity, Including How Eligibility for the Training was Defined. (You may attach artifacts of the training in lieu of providing a description, such as training slides, sections, or an entire manual).</th>
<th>Number of Eligible Participants by Type</th>
<th>Number of Actual Participants by Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>● [5_018] IAP Guidebooks Window 3 Reports Webinar Slides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>● [6_009] IAP Guidebooks End-of-Year Report Webinar Slides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teacher Leader Summit Sessions</strong></td>
<td>LDOE led sessions at the 2023 Teacher Leader Summit, which focused on both familiarizing participants about the innovative assessment and providing guidance on scoring writing.</td>
<td>Two sessions for each presentation – 175 participants in each session</td>
<td>Two sessions of 175 Educators and School Leaders for each presentation</td>
</tr>
<tr>
<td>● [6_010] Are We There Yet: The Brave New World of Through-Year Assessments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>● [6_011] Sharing Best Practices for the ELA Innovative Assessment Program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Training on test security for the innovative assessment system (e.g., training on handling and distribution of innovative assessment materials, monitoring administration of innovative assessments)</td>
<td>The Administration Webinars mentioned above in VI-B.13(1), i.e., point 2 above, address security and accommodations in detail. In addition, security and accommodations are also addressed within the test coordinator and administration manuals (see [5_025], [5_026], [5_027] and [5_028]). Finally, [5_029] <em>Louisiana's Data Governance &amp; Student Privacy Guidebook</em> provides additional guidance on security.</td>
<td>Same as VI-B.13(1), i.e., point 2 above.</td>
<td>Same as VI-B.13(1), i.e., point 2 above.</td>
</tr>
<tr>
<td>(3) Training on providing accommodations for students with disabilities in the innovative assessment system (e.g., training on specific types of accommodations that can be made in the presentation, response, timing and/or setting)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>● [6_001] Window 1 Adam Training Webinar (see slides 51-57 on security, slides 45-49 on accommodations; similar slides exist within the other two training webinar slide decks)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>● [6_002] Window 2 Adam Training Webinar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training Topic</td>
<td>Brief Description of Training Opportunity, Including How Eligibility for the Training was Defined. (You may attach artifacts of the training in lieu of providing a description, such as training slides, sections, or an entire manual).</td>
<td>Number of Eligible Participants by Type</td>
<td>Number of Actual Participants by Type</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>of the innovative assessment to support participation of students with disabilities)</td>
<td>(4) Training on providing accommodations for English learner (EL) students in the innovative system (e.g., training on specific types of accommodations that can be made in the presentation, response, timing and/or setting of the innovative assessment to support participation of EL students)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Training on using innovative assessment data to inform instruction (e.g., training on analysis and interpretation of individual, subgroup, and/or class-level data for the purposes of identifying struggling students; checking student mastery; adapting instructional resources and/or pacing; differentiating instruction; changing instructional strategies)</td>
<td>Training on data use is addressed within the evidence provided in VI-B.13(1), i.e., point 2 above, specifically the LDOE Led Professional Development for Volunteer Districts, Reporting Webinars and Teacher Leader Summit Sessions. One key resource leveraged in all of these trainings is the [5_019] IAP Guidance for Score Reports, which connects performance on the IAP to potential next instructional next steps. Finally, LDOE provides spreadsheet files at the classroom, school and district level with guidance on how to use the data in their own investigations.</td>
<td>Same as VI-B.13(1), i.e., point 2 above.</td>
<td>Same as VI-B.13(1), i.e., point 2 above.</td>
</tr>
<tr>
<td><img src="image_url" alt="Table Image" /></td>
<td><img src="image_url" alt="Table Image" /></td>
<td><img src="image_url" alt="Table Image" /></td>
<td><img src="image_url" alt="Table Image" /></td>
</tr>
<tr>
<td>Training Topic</td>
<td>Brief Description of Training Opportunity, Including How Eligibility for the Training was Defined. (You may attach artifacts of the training in lieu of providing a description, such as training slides, sections, or an entire manual).</td>
<td>Number of Eligible Participants by Type</td>
<td>Number of Actual Participants by Type</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>(6) Training on using innovative assessments for accountability (e.g., training on analysis and interpretation of class and grade-level data for the purposes of informing curricular decisions and allocation of resources to support instruction at the school)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Training on using innovative assessments for accountability across student subgroups (e.g., training on analysis and interpretation of subgroup, class, and grade-level data for the purposes of identifying and addressing any gaps between student subgroups)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
VI.C. Familiarization of Innovative Assessment to Students, Parents, and LEA Staff

Describe how the SEA or consortium familiarized students, parents, and LEA staff with the innovative assessment system during the reporting year (2022-23). Familiarization may include sharing a description of the new innovative assessment system, highlights of the key differences between the innovative and existing assessment systems, initial challenges associated with implementing the new system, and benefits of the innovative assessment system. Examples of familiarizing students and parents include materials that were sent to parents describing the innovative assessment system, agendas of meetings with parents and students to describe the innovative assessment system, and postings about the innovative assessment system on schools’/districts’ websites. Examples of familiarizing LEA staff include materials from meetings to describe the innovative assessment system, agendas and materials from trainings for staff on implementing the innovative assessment system. The focus of this section is twofold: (a) information the State or consortium provided to students and parents to familiarize them with and acclimate them to the innovative assessment system and (b) support and training the State or consortium provided to LEA staff to familiarize and enable them to implement the innovative assessment system. Familiarizing students, parents, and LEA staff goes beyond the basic parental notification requirement in Section IX.

Table VI-B.14. Familiarization of Innovative Assessment to Students, Parents, and LEA Staff

<table>
<thead>
<tr>
<th>SEA or Consortium Takes Action to Familiarize the Following Individuals with the Innovative Assessment System</th>
<th>Description of (a) the Process the State or Consortium used to Familiarize and Acclimate Students and Parents to the Innovative Assessment System and (b) the Support and Training the State or Consortium Provided to LEA Staff to Implement the Innovative Assessment System (be sure to describe the process for each group listed in the left-hand column. You may attach artifacts [e.g., letter to parents, practice IADA items, meeting or training agenda, training session manual/materials] of the actual process in lieu of providing a description).</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Familiarize and acclimate students and parents to the IADA assessment system</td>
<td>LDOE has provided a number of interpretive materials to support students and parents to familiarize them with the innovative assessment system. These materials include:</td>
</tr>
<tr>
<td></td>
<td>● [6_012] End of Unit Parent Score Report Guide</td>
</tr>
<tr>
<td></td>
<td>● [5_015] Translated Parent Guides</td>
</tr>
<tr>
<td></td>
<td>● [5_023] LEAP ELA Guidebooks Assessment Guide Grades 6, 7, 8</td>
</tr>
<tr>
<td></td>
<td>● [5_022] IAP Interpretive Guide</td>
</tr>
<tr>
<td></td>
<td>● [5_021] Louisiana Key Initiatives Flyer for Parent Open House</td>
</tr>
</tbody>
</table>
Support and train LEA and school staff to implement the IADA assessment system and administer the IADA assessments.

As explained previously, LDOE hosted administration webinars prior to each window to inform the field on appropriate administration and security methods. All of the materials referenced within VI.B are relevant here.

Again, within each administration webinar, as well as publicly available online, LDOE has provided materials on the appropriate and secure administration of the IAP. These materials include:

**Grades 6, 7 and 8 Operational Manuals**
- [5_025] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Test Coordinator Manual - Window 3
- [5_026] LEAP ELA Guidebooks Grades 6, 7 and 8 Operational Administration Manual - Window 3

**Grade 5 Field Test Manuals**
- [5_027] LEAP ELA Guidebooks Grade 5 Pilot Test Coordinator Manual - Window 3
- [5_028] LEAP ELA Guidebooks Grade 5 Pilot Administration Manual - Window 3
VII: Use of Innovative Assessment Data

Please describe how teachers, principals, and other school leaders are using the innovative assessment data during the reporting year (2022-23). You may attach artifacts in lieu of providing a description.

In particular:

To the extent the SEA has tracked teacher participation in activities that involve using innovative assessment data to inform instruction, report the percentage of participating teachers who have engaged in these activities. Examples of activities include using the data to identify struggling students, check student mastery, group students to deliver differentiated instruction, or change the pacing of lessons. Note that teachers may participate in activities using assessment data to inform instruction either individually or in teams.

To the extent the SEA has tracked principal and other school leader participation in activities that involve using innovative assessment data to improve accountability, report the percentage of participating principals and other school leaders who have engaged in these activities. Examples of activities include monitoring students’ participation rates, evaluation of interim progress against long-term school improvement goals, root cause analysis, action planning, or identifying and addressing gaps between student subgroups.

As explained within VI-B.14, LDOE led professional development within two districts oriented around using the information to support instructional decision making. In addition, numerous teachers reported using the [5_019] IAP Guidance for Score Reports, but such feedback was given on an ad hoc basis.
VIII: Changes in Consortium Governance or Membership (if applicable).

Describe any changes in the Consortium governance structure, roles and responsibilities, or membership, during the reporting year (2022-23), or any changes anticipated in the future.

N/A

IX: Parental Notification

Describe how the SEA or Consortium is ensuring that each participating LEA informs parents of all students in participating schools about the innovative assessment, including the grades and subjects in which the innovative assessment will be administered, and, consistent with section 1112(e)(2)(B) of the Act, at the beginning of each school year during which an innovative assessment will be implemented. Such information must be--

(i) In an understandable and uniform format;
(ii) To the extent practicable, written in a language that parents can understand or, if it is not practicable to provide written translations to a parent with limited English proficiency, be orally translated for such parent; and
(iii) Upon request by a parent who is an individual with a disability as defined by the Americans with Disabilities Act, provided in an alternative format accessible to that parent.

LDOE follows the same procedure taken with all statewide summative assessments. School systems are tasked with the responsibility for informing parents of their students’ testing experiences. This expectation is communicated through a variety of webinars and LDOE disseminated materials.
X: Assurances

If the innovative assessment system will initially be administered in a subset of LEAs or schools in a State, please attach an assurance from the SEA that affirms it has collected assurances from each participating LEA that the LEA will comply with all requirements of this section.

The superintendent of each participating school system signed an assurance. An example assurance is provided in [10_001] Example Assurance. Signed assurances can be provided securely upon request.

XI: Budget

Please describe any changes to the budget that vary from the approved application budget.

The Louisiana Department of Education budget for the program described for the 2023-2024 school year in this report was over $4.3M. At this time, the project is within the budgeted amount allocated for the tasks completed. Additional funding sources that made this work possible include the partial allocation of the CGSA 2020 grant to support the grades 3-5 development in the 2022-2023 school year. In addition, there was a grant from Ed First to LDOE for the curriculum relevant pilot program. USDOE also awarded LDOE with the CGSA 2022 grant for scaling the program in the 2023-2024 school year with the curriculum relevant pilot program.

XII: Certification

To the best of my knowledge and belief, all data in this annual performance report are true and correct and the report fully discloses all known weaknesses concerning the accuracy, reliability, and completeness of the data.

<table>
<thead>
<tr>
<th>Name of Authorized Representative:</th>
<th>Title:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomas Y. Lambert</td>
<td>Assistant Superintendent for Assessments, Accountability, and Analytics</td>
</tr>
</tbody>
</table>

Signature: Date:

09/29/2023
<table>
<thead>
<tr>
<th>Grade</th>
<th>CurriculumToken</th>
<th>Test Status</th>
<th>Form</th>
<th>ADAM Form Name</th>
<th>Units</th>
<th>Fall 2022</th>
<th>Winter 2023</th>
<th>Spring 2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>The War Between Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>The War Between Us</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>Birch bark</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>Steve Jobs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>Stones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Birch bark</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Underground</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Happiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Everyday Use / Dream Catcher</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Summit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>A Christmas Carol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>Behind the Scenes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>Cultures in Conflict</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Cultures in Conflict</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>LWW Lion, Witch, Wardrobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>LWW Lion, Witch, Wardrobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Spruce</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Bird</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Why Practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Home</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>Call of the Wild</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>Scientist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Scientist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>SOTS Shutting Out the Sky</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>SOTS Shutting Out the Sky</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>Wonderstruck</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 GB</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Wonderstruck</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form A</td>
<td></td>
<td>Word Play</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 5 WW</td>
<td>FT</td>
<td>Form B</td>
<td></td>
<td>Word Play</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Cranes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>DUST-A Dust</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>HATCHET-A Hatchet</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>JOBS-A Steve Jobs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>STONES-A Stones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Woman</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Underground</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Happiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Summit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>CC-B A Christmas Carol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>SCENES-B Behind the Scenes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Crossing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Magic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>GIVER-C The Giver</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>WRITTEN IN BONE-C Written in Bone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Bird</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Why</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 7 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Home</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>COTW-A Call of the Wild</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Chronicles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>FLOWERS-A Flowers for Algernon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>SUGAR-A Sugar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 GB</td>
<td>OP</td>
<td></td>
<td></td>
<td>YTH-A Tell Tale Heart</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Genius</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Archeology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Alyssa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Courage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 8 LEAP</td>
<td>OP</td>
<td></td>
<td></td>
<td>Fun</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Teachers are the single most important in-school determinant of students’ success, and the curricular materials teachers use to shape what and how teachers teach.\(^1\) Research shows that implementing a high-quality, knowledge-building curriculum can improve student learning and equity across classrooms\(^2\) while supporting teachers.

Given the importance of curriculum use, the Johns Hopkins Institute for Education Policy (Institute) and Louisiana Department of Education (LDOE) surveyed all English language arts (ELA) teachers statewide from Grades 3 through 8 — to gain a better understanding of curricular use, fidelity of implementation, and support for the schools’ current curriculum. Researchers at the Institute administered the survey, which adopts validated questions from the RAND Corporation’s American Teacher Panel.

The survey was administered district-wide from 1/10/2023 to 4/30/2023. The total number of expected teacher responses for this survey was approximately 7,077 regular classroom ELA teachers. For this survey, 411 ELA teachers responded, resulting in a 5.8% response rate. *Due to the low response rate, findings should be interpreted with caution and are not generalizable in any way.* Table 1 below provides characteristics of the teachers who responded to the survey.

<table>
<thead>
<tr>
<th>Table 1: Descriptive statistics of the teacher respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average years of teaching at current school</td>
</tr>
<tr>
<td>Average years of teaching experience</td>
</tr>
<tr>
<td>Average number of students</td>
</tr>
<tr>
<td>Average number of ELL students</td>
</tr>
<tr>
<td>Average number of students receiving special education services</td>
</tr>
<tr>
<td>Average number of students identified as gifted</td>
</tr>
</tbody>
</table>

The following research summary analyzes the survey responses, highlights significant findings, and offers suggestions to promote the greater use of high-quality instructional materials. The research questions and the Institute’s high-level findings are as follows:

**Which curricular materials do teachers use?**

- **Most elementary** ELA teachers report using instructional materials from a single source. The ELA Guidebooks curriculum is the most widely used curriculum by survey respondents.
When asked which instructional materials ELA teachers use, 54% of teachers report drawing upon a single source from the curricular options provided in the survey for their ELA classroom lessons. 26.2% report using two sources, and 19.7% report using three or more for their ELA instructional materials.

Of those ELA teachers who responded to the survey, 45% stated they used *ELA Guidebooks*, while 25.5% stated they used *i-Ready*. Figure 1 below shows the full set of instructional materials the respondents draw upon for their lessons. Note that teachers can list all materials they use, so the percentages do not equal 100%.

*Figure 1: Instructional resources elementary ELA teachers draw upon for their ELA classroom lessons*

- ELA Guidebooks: 45.0%
- Other published materials: 29.2%
- i-Ready: 25.5%
- Materials created by the district/school: 23.8%
- Materials I developed myself: 19.5%
- Wit & Wisdom: 8.0%
- Springboard: 0.5%
- Reading Street: 0.6%
- EL Education: 1.2%
- Engage NY (CKLA): 1.0%
- Core 5 Reading: 0.5%

When asked about the curriculum they use *most* frequently, 47.5% of teachers who responded to the survey stated that *ELA Guidebooks* were their most frequently used curriculum. The complete list of responses to this question is shown in Figure 2 below.

*Figure 2: Elementary ELA teacher responses to the question “Overall, which of the following published materials do you use most frequently?”*

- ELA Guidebooks: 47.5%
- Other published materials: 23.1%
- Materials created by the district/school: 11.9%
- Wit & Wisdom: 6.8%
- i-Ready: 3.6%
- EL Education: 1.3%
How do teachers use their curricular materials?

- **Most ELA teachers consistently use ELA Guidebooks, the most frequently used curriculum, for various lesson elements (e.g., for lesson objectives and texts to refresh content knowledge).** For example, 93% report using *ELA Guidebooks* for lesson objectives “always” or “most of the time” compared to 4.9% who state they “never” or “sometimes” use the curriculum for lesson objectives, as shown in Figure 3.

Likewise, 92.5% of teachers use *ELA Guidebooks* to introduce or teach an ELA lesson, compared to just 4.3% who state they “never” or “sometimes” use *ELA Guidebooks* for this purpose. In fact, most *ELA Guidebooks* users employ the materials for nearly every purpose outlined in the survey. One area in which teachers seem to lack consistent use of *ELA Guidebooks* is as a source for students to complete homework, with 40.5% using it “always” or “most of the time.” That said, these findings suggest that many teachers use *ELA Guidebooks* consistently and as intended.

*Figure 3: Responses to the question “In a typical month, how often do you use ELA Guidebooks as a source for lesson objectives?”*

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>79.5%</td>
</tr>
<tr>
<td>Most of the time</td>
<td>13.5%</td>
</tr>
<tr>
<td>About half the time</td>
<td>2.2%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>3.8%</td>
</tr>
<tr>
<td>Never</td>
<td>1.1%</td>
</tr>
</tbody>
</table>

What are teachers’ opinions of the curricular materials they use?

- **ELA teachers who use ELA Guidebooks generally have favorable opinions of the curriculum.** Of the 47.3% of ELA teachers who use *ELA Guidebooks* most frequently, the most commonly used, published ELA curriculum, overall opinions are generally favorable. For example, 59.9% said they have a favorable impression of the curriculum (i.e., strongly agree and agree, combined), compared to 29.4% who did not. Similarly, 59.9% agreed with the statement “The curriculum is easy to use,” whereas only 27.1% disagreed, as shown in Figure 4. Additionally, ELA teachers who responded to the survey felt that *ELA Guidebooks* met the needs of students above grade level (75.7%) and at grade level (59.4%).
Figure 4: Responses to the statement "The curriculum (ELA Guidebooks) is easy to use (e.g., for lesson planning, student work, etc.)."

<table>
<thead>
<tr>
<th>Response</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>24.9%</td>
</tr>
<tr>
<td>Somewhat agree</td>
<td>35.0%</td>
</tr>
<tr>
<td>Neutral</td>
<td>13.0%</td>
</tr>
<tr>
<td>Somewhat disagree</td>
<td>19.2%</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>7.9%</td>
</tr>
</tbody>
</table>

Although opinions were largely positive, ELA teachers also agreed on a few key areas in which they felt ELA Guidebooks fell short. For example, 65% did not believe the curriculum meets the needs of students below grade level; 51.5% did not feel the curriculum meets the needs of English Language Learners (ELL); and 57.6% did not believe the curriculum meets the needs of students with an IEP. These responses suggest that these resources provide less support for struggling students and students learning English.

Which online sources do teachers consult when planning their instruction?

- The vast majority of ELA teachers report that they use a few online resources frequently. For example, 71.3% of ELA teachers who responded to the survey use the Louisiana Department of Education website on a near-weekly basis (i.e., 2-3 times per month or more). Other frequently used resources include Google (73%), Teachers Pay Teachers (46.7%), and Readworks (40%). Figure 5 highlights the online resources ELA teachers use most frequently. On the other end of the spectrum, we found that almost no teachers use National Council for Teachers of English (NCTE), Sharemylesson, or Teaching Channel online resources for their ELA instruction.

Figure 5: The most frequently used online resources by ELA teachers.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>73.0%</td>
</tr>
<tr>
<td>LDOE Website</td>
<td>71.3%</td>
</tr>
<tr>
<td>Teachers Pay Teachers</td>
<td>46.7%</td>
</tr>
<tr>
<td>Readworks.org</td>
<td>40.0%</td>
</tr>
</tbody>
</table>
How much time do teachers spend planning?

- **ELA teachers who responded to the survey spend nearly equal amounts of time selecting materials and planning lessons, irrespective of the materials they use.** Specifically, ELA teachers report spending an average of seven and a half hours (7.5) per week planning lessons using either district-purchased materials or instructional resources. These teachers also report spending an average of three hours selecting their own materials. Of particular note, 6% of teachers state that they spend one hour or less planning with district-purchased materials and instructional resources, and 13.9% do not spend any time selecting their own materials.

How much professional development support do teachers receive? What did teachers think about this support?

- **ELA teachers indicate they received a beneficial amount of professional development (PD) to support using their ELA instructional materials.** Over the past year, elementary ELA teachers received an average of 19.4 hours of PD to support using ELA instructional materials. However, 6.8% of teachers indicated they received no PD to support using ELA instructional materials over the past year.

- **ELA teachers held fairly positive opinions about the PD received.** For example, 59.8% agree (i.e., strongly and somewhat agree, combined) with the statement, “The professional development enhanced my content knowledge of the curriculum,” and 68.9% agreed with the statement, “The professional development I received is aligned with district/state standards.” Also of note, 60.2% felt that the professional development they received helped them teach their most frequently used curriculum, as shown in Figure 6.

![Figure 6: Responses to the statement “The professional development has helped me teach my most frequently used curriculum.”](image)

LDOE supplemental questions on writing modes, text genres, and tasks and assessments

- In addition to our analysis of teachers’ thoughts and opinions of the curriculum they used, we asked teachers to identify at what time of the year to introduce specific writing modes, text genres, tasks, and assessments based on LDOE standards. Figures 7, 8, and 9 (including 9a, 9b, and 9c) outline teacher responses to those questions:
Figure 7: Responses to the statement “Please indicate which of the following writing modes are the focus of your unit instruction, and when in the school year they are taught.”

<table>
<thead>
<tr>
<th>Writing Tasks</th>
<th>BOY</th>
<th>MOY</th>
<th>EOY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literary Analysis (W.5.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informative/Explanatory (W.5.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrative (W.5.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opinion (W.5.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| BOY= Beginning of Year; MOY=Middle of Year; EOY=Middle of Year

Figure 8: Responses to the statement “Please indicate which of the following text genres are the focus of your unit instruction, and when in the school year they are taught.”

<table>
<thead>
<tr>
<th>Literary Texts</th>
<th>BOY</th>
<th>MOY</th>
<th>EOY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poetry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drama</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novel/Short Story</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| BOY= Beginning of Year; MOY=Middle of Year; EOY=Middle of Year

©2023 by Johns Hopkins University
Figure 9: Please indicate which of the following types of culminating tasks/assessments are included in your units of instruction, and when in the school year they are administered.

**Writing Tasks**

- **Literary Analysis (W.5.9)**
  - BOY: 50.4%
  - MOY: 50.6%
  - EOY: 43.1%
- **Narrative (W.5.3)**
  - BOY: 29.7%
  - MOY: 41.6%
  - EOY: 29.4%
- **Informative/Explanatory (W.5.2)**
  - BOY: 45.3%
  - MOY: 38.4%
  - EOY: 48.2%
- **Opinion (W.5.1)**
  - BOY: 32.8%
  - MOY: 28.0%
  - EOY: 26.0%

**Formal Discussions**

- **Debate**
  - BOY: 22.1%
  - MOY: 32.1%
  - EOY: 28.0%
- **Socratic Seminar**
  - BOY: 47.9%
  - MOY: 51.3%
  - EOY: 41.4%
Based on these findings, the Institute would like to provide LDOE with the following suggestions:

- **Suggestion 1 – Continue to promote consistent use of high-quality instructional materials.** Research is clear that using a high-quality, knowledge-building curriculum is important for both student learning and teacher support. Survey responses suggest that teachers use district-endorsed curricular materials. However, motivating teachers to use high-quality materials with fidelity is often challenging for many districts. Based on the analysis above, LDOE leadership is largely meeting that challenge. We suggest that instructional and school leaders work with the small group of teachers who are not consistently using high-quality materials in order to support their specific instructional needs better.

- **Suggestion 2 – Continue to provide professional development that supports teacher knowledge and use of high-quality instructional materials.** Survey results indicate that nearly 7 in 10 teachers believed that professional development enhanced their content knowledge and helped them teach their most frequently used curriculum. We suggest you continue providing professional development that directly supports teachers’ knowledge and ability to leverage instructional materials effectively.

- **Suggestion 3 – Ensure that the district-endorsed materials meet the needs of struggling English language learners.** Teachers who primarily use *ELA Guidebooks*, the most widely-used curricula by grade band, state that they had favorable impressions of the materials and that they mostly meet the needs of students above and at grade level. However, few teachers agree that these materials meet the needs of students below grade level, with IEPs, or those learning English. We recommend that these materials be reviewed and, if necessary, augmented to support students with these diverse learning needs better.

### Table I-3. Detailed, High Level Summary of Activities

The following table outlines the high level summary of activities for the 2022-2023 academic year. The LDOE collaborated with national partners and school systems to continue its work with the Innovative Assessment Program. These national partners are included in the Parties Responsible column of the table and include the following organizations:

- Odell Education [Odell] - content design and development
- Johns Hopkins University [JHU] - content design and development
- Great Minds [Great Minds] - Wit and Wisdom content design and development
- Progressive Measurement [PM] - psychometrics, research, data analysis, and reporting
- MZ Development [MZD] - test delivery, data management, and reporting
- Strategic Measurement and Evaluation [SME] – rangefinding and hand-scoring
- Center for Assessment [Center] - technical assistance, research/psychometrics, and reporting
- EducationFirst (EdFirst)--philanthropic funding organization; strategy and policy support

<table>
<thead>
<tr>
<th>Dates</th>
<th>Activities</th>
<th>Status</th>
<th>Parties Responsible</th>
</tr>
</thead>
</table>
| July 1, 2022 – June 30, 2023 (Wednesday) | Ed First Weekly Team Meeting  
  - Develop a model for ELA curriculum--relevant through-year assessments that connects with multiple high-quality curricula, an assessment. Create a blueprint and then administer the assessment in the spring of 2023 in a small number of schools at grade 5. Develop innovative reports and supports for improved understanding of results with direct instructional applications.  
  - The assessments will be designed to connect flexibly with high-quality curriculum across the state, based on analyses of commonalities in ELA curricula.  
  - In consultation with stakeholders, the Department will design innovative score reports connected to corresponding curriculum-relevant supports that provide more timely, actionable, and understandable information on student performance. | Completed | NWEA, Center, LDOE, PM, Odell, JHU, MZD, SME, Great Minds |
|------------------------------| IAP Office Hours with LDOE  
  - Held the third Wednesday of every month  
  - Provides support for assessment design, administration, and reports  
  - Allows IAP school systems to interact and learn from each other | Completed | LDOE |
|------------------------------| Content Call  
  - Designed to focus on content-related assessment design/item and forms development topics  
  - Provide status on current tasks  
  - Discuss topics related to the project, as needed  
  - Discuss questions and concerns around upcoming deliverables and milestones | Completed | NWEA, Center, LDOE, Odell, JHU, Great Minds |
|------------------------------| Innovative Assessment Rangefinding Thursday  
  - Weekly series for internal LDOE/SME team rangefinding  
  - SME selects "rangefinding/annotate" materials  
  - SME reviews student responses and pulls representative samples for LDOE’s review  
  - Meeting is used to discuss patterns, trends, and observations  
  - Set score ranges and scorer training materials for all prompts | Completed | LDOE, SME |
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Description</th>
<th>Completion</th>
<th>Responsible Entities</th>
</tr>
</thead>
</table>
| January 1 – June 30, 2023 (Thursday)          | Weekly Psychometric Call  
- Collaborate and partner on the research/psychometric work plan and future need  
- Discuss the key areas of work and the priorities for the current school year.  
- Discuss questions and concerns around upcoming deliverables and milestones  
- Discuss research-related topics, as needed                                                                                             | Completed    | NWEA, Center, LDOE, PM                 |
| **May 2022**                                  |                                                                                                                                            |              |                                        |
| May 1, 2022                                   | LDOE releases the Assessment Calendar for the 22-23 school year.  
- Window 1 - Fall Test Window (11 days)  
- Window 2 - Winter Test Window (14 days)  
- Window 3 - Spring Test Window (31 days)                                                                                     | Completed    | LDOE                                   |
| **May 2022**                                  | Student Engagement Survey  
- Focused on engagement, demonstration of learning, assessment preferences, experience, empowerment, preparation, and results  
- Purpose was to gain an understanding of student's general thoughts regarding the innovative assessment                                                                                     | Completed    | LDOE, NWEA, Center, Ed First           |
| May 26 – September 15, 2022                   | LDOE prepared the IAP forms for Window 1 in partnership with Odell, JHU, and Great Minds. This includes forms that are fully checked for accessibility with accompanying accommodated print forms, CAS, braille, and final test maps.  
- Grade 5 (GB)-Birchbark  
- Grade 5 (WW)-Cultures in Conflict  
- Grade 5 (GB)-Scientist  
- Grade 6 (GB)-Dust  
- Grade 6 (GB)-Hatchet  
- Grade 7 (GB)-The Giver  
- Grade 7 (GB)-Written in Bone  
- Grade 8 (GB)-Flowers for Algernon                                                                                                   | Completed    | LDOE, Odell, JHU, Great Minds          |
| **July 2022**                                 | LDOE assembled Fall and Winter Window 1/Window 2 Linking (LEAP) Sets to support comparability with LEAP 2025 and to place the IAP on the same scale. This process was repeated for the Spring Linking Sets.  
- LDOE requests items and other data for needed passage sets  
- DRC provides requested items and data  
- Review test information functions for linking passage sets  
- Items imported into ADAM testing platform by MZD  
- Items reviewed by MZD for completeness; scoring properties  
- Grades 6-8 Linking items TTS approved  
- Content ready for pre-test validation (LDOE)                                                                                       | Completed    | LDOE, NWEA, DRC, Center, MZD           |
| **August 2022**                               | SME prepared for the Fall Window 1 OP/FT Scoring and Rangefinding. This process was repeated for the Winter and Spring test administrations.  
- Fall operational scoring  
  - SME builds scoring guides  
  - LDOE reviews scoring guides                                                                                                         | Completed    | SME, LDOE                              |
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Description</th>
<th>Complete Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 29, 2022 – October 14, 2022</td>
<td>LDOE conducted the Fall Window 1 End of Unit (EOU) student data file layout review and process for establishing business rules. This process was repeated for the Winter and Spring test administrations.</td>
<td>Completed</td>
</tr>
</tbody>
</table>
| August 26 – December 12, 2022 | LDOE prepared the IAP forms for Window 2 in partnership with Odell, JHU, and Great Minds. This includes forms that are fully checked for accessibility with accompanying accommodated print forms, CAS, braille, and final test maps.  
- Grade 5 (GB)-Lion, Witch, Wardrobe  
- Grade 5 (GB)-Wonderstruck  
- Grade 5 (WW)-Word Play  
- Grade 6 (GB)-Dust  
- Grade 6 (GB)-Hatchet  
- Grade 6 (GB)-Steve Jobs  
- Grade 7 (GB)-A Christmas Carol  
- Grade 7 (GB)-The Giver  
- Grade 7 (GB)-Written in Bone  
- Grade 8 (GB)-Call of the Wild  
- Grade 8 (GB)-Tell Tale Heart | Completed LDOE, Odell, JHU, Great Minds |
| August 1-September 20, 2022 | LDOE creates assessment guides.  
- Develop assessment guides for all IAP grades (operational and field tests) |                |
| **September 2022**            |                                                                                                                                           |                 |
| September 6 – October 10, 2022 | LDOE creates the test delivery platform setup procedures for Fall, Winter, and Spring.  
- Fall org file submitted to MZD  
- Updated District Test Coordinator list provided to MZD  
- Provided district unit assignment to MZD  
- Student roster data delivered from Escholar to MZD  
- Submitted student accommodations information to MZD  
- System opened for districts to review school rosters - Grades 5-8  
- System opened for creation of Fall Test Administrations - Grades 5-8 | Completed LDOE, MZD |
| September 6 – October 10, 2022 | LDOE creates the test administration manuals and test coordinator manuals for Fall, Winter, and Spring. This process was repeated for the Winter and Spring.  
- Fall Window 1 Manuals (Operational/Field Test)  
- Grade 6-8 OP Fall Test Administration Manual  
- Grade 5 FT Fall Test Administration Manual  
- Grade 6-8 OP Fall Test Coordinators Manual  
- Grade 5 FT Fall Test Coordinators Manual | Completed LDOE, SME, MZD |
| September 6 – October 11, 2022 | LDOE conducted Fall Window 1 Pre-Test Validation quality assurance activities with the vendors as part of the end-to-end testing. This process was repeated for the Winter and Spring test administrations.  
- Create administrations with fake student profiles  
- Walk-Through on Test Case Process with LDOE  
- Build out new reports (District/School unit reports) and update class report layout  
- MZD conducts full walk through / test case input to validate content, scoring  
- LDOE to take UAT test cases with fake student profiles | Completed NWEA, Odell, JHU, MZD, LDOE |
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Event Description</th>
<th>Status</th>
<th>Participants</th>
</tr>
</thead>
</table>
| September 16 – December 5, 2022| MZD submits data file to LDOE  
LDOE reviews/verifies Pre-Test Validation File test results  
LDOE signs off on UAT data file  
Create live fall test administration in ADAM | Completed       | Center, NWEA, LDOE               |
| October 2022                    |                                                                                   |                 |                                   |
| October 18, 2022                | Student Survey                                                                    | Completed       | LDOE, Center, NWEA, Ed First      |
| October 20, 2022                | Curriculum-Relevant Teacher Engagement  
- Focus group with 6 teachers  
- Discuss end-of-unit reporting  
- Initial information to create new way to organize information in Class Report | Completed       | LDOE, Center, NWEA, Ed First      |
| October 26 – 29, 2022           | Louisiana TAC Meeting  
- TAC Meeting  
- TAC Meeting Debrief  
- TAC Debrief with LDOE  
- TAC Debrief with LDOE  
- TAC Debrief with LDOE | Completed       | LDOE, Center, NWEA               |
| November 2022                   | LDOE conducted the Fall Window 1 Post Window Validation in the operational Grades 6-8. This process was repeated for the Winter and Spring test administrations.  
- Fall Post-window validation file and Blue Dot reports delivered to LDOE  
- LDOE reviews Fall post-window data file and Blue Dot reports  
- Fall 2022 data file delivered to LDOE  
- Fall 2022 reports delivered to LDOE for release (end of units class, student, and reports files)  
- EOU supplemental materials for teachers  
- IAP Window 1 Reports Webinar | Completed       | LDOE, MZD, NWEA                  |
| December 2022                   |                                                                                   |                 |                                   |
| December 8, 2022                | Ed First Curriculum-Relevant Engagement Session 1  
Ed First Curriculum-Relevant Engagement Session 2  
Ed First Curriculum-Relevant Engagement Session 3 | Completed       | LDOE, Center, NWEA, Ed First     |
| December 1, 2022 – March 16, 2023| February 2023 Spring TAC Meeting  
- Identify list of potential topics  
- Meet with vendors/presenters to confirm topics  
- Draft high-level agenda and share with presenters  
- LDOE receives advanced meeting materials from presenters  
- LDOE request revisions to materials  
- Meeting-ready TAC agenda  
- Share meeting-ready agenda and materials with TAC  
- Attend Spring 2023 TAC meeting  
- TAC debrief with LDOE | Completed       | Center, NWEA, LDOE               |
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Event Description</th>
<th>Status</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 9 – January 11, 2023</td>
<td>Committee Meetings/Passage Review (Virtual)</td>
<td>Completed</td>
<td>LDOE, NWEA</td>
</tr>
<tr>
<td></td>
<td>- Prepare materials (training deck, passage review trackers)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grades 3 and 4 passage review meeting</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Obtain educator list from LDOE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Provide list to procurement for stipends</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- DocuSign sent to educators with instructions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>December 12 – March 10, 2023</td>
<td>LDOE prepared the IAP forms for Window 1 in partnership with Odell, JHU, and Great Minds. This includes forms that are fully checked for accessibility with accompanying accommodated print forms, CAS, braille, and final test maps.</td>
<td>Completed</td>
<td>LDOE, Odell, JHU, Great Minds</td>
</tr>
<tr>
<td></td>
<td>- Grade 5 (WW)-A War Between Us</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grade 5 (GB)-Shutting Out the Sky</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grade 6 (GB)-Stones</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grade 7 (GB)-Behind the Scenes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grade 8 (GB)-Sugar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>December 2022 – February 2023</td>
<td>Survey Teacher Engagement</td>
<td></td>
<td>LDOE, Center, NWEA, Ed First</td>
</tr>
<tr>
<td></td>
<td>- Analyze teachers’ classroom practices and utilization of instructional materials. The information gathered in this survey will demonstrate how ELA teachers in grades 3-8 understand and utilize their curricula.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Feedback on reports and results; how they engage in learning more about reading and writing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Better understand curriculum use in the state.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 2023</td>
<td>SY22 Technical Report - Grade 7</td>
<td>Completed</td>
<td>Center, NWEA, LDOE</td>
</tr>
<tr>
<td></td>
<td>- Makes detailed Grade 7 technical report outline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- LDOE reviews Grade 7 tech report outline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Revise Grade 7 tech report outline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- LDOE signs off on final tech report outline</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Finalize Grade 7 technical report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 3, 2023</td>
<td>Curriculum-Relevant Engagement Planning Meeting</td>
<td>Completed</td>
<td>LDOE, Center, NWEA, Ed First</td>
</tr>
<tr>
<td>January 4, 2023</td>
<td>Innovative Assessment Focus Group Feedback (Plaquemines Parish)</td>
<td>Completed</td>
<td>LDOE, Center, NWEA, Ed First</td>
</tr>
<tr>
<td>January 10, 2023</td>
<td>IAP Partners Prepared Student Survey Instructions</td>
<td>Completed</td>
<td>LDOE, Center, NWEA, Ed First</td>
</tr>
<tr>
<td></td>
<td>- Adherence to Louisiana’s requirements for student surveys</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Compliance with the Protection of Pupil Rights Act</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Notice to families of purpose, use, and intent of survey questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 12 – February 1, 2023</td>
<td>Scaling Specifications</td>
<td>Completed</td>
<td>LDOE, MZD, NWEA, Center</td>
</tr>
<tr>
<td></td>
<td>- NWEA sends LDOE draft scaling specifications for review</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- LDOE/NWEA Meeting - Weekly Psychometric Call</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- LDOE reviews draft scaling specifications</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- NWEA incorporates feedback and edits</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- NWEA sends revised scaling specifications to LDOE for psychometric implementation for the program</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Scaling specifications finalized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 27, 2023</td>
<td>Window Debrief Process</td>
<td>Completed</td>
<td>NWEA, LDOE</td>
</tr>
<tr>
<td></td>
<td>- The intention for the Window Program Debrief was to ask the partners to come together and think about their</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
“reactions” and “perspectives” from the recent test windows [Window 1 and Window 2].
- The goal for these debriefs was to support continuous process improvement and needs to streamline workflow in order to prepare for Window 3 [Spring 2023] activities and beyond.
- The feedback sessions highlighted areas to celebrate and maintain energy and efforts on established procedures, and in some areas, where improvement is needed, such as, communication, monitoring, and streamlining workflow.

### January 2023
- **District Engagement Survey**
  - Collection of information regarding curriculum usage at the district level; gauging general interest level in participating in pilot.
  - Supported the creation of “Collaboration Groups” that supported the work on a report redesign.
- **February 2023**
  - **Plaquemines Parish Focus Group**
    - Open Feedback Session
  - **Committee Meetings**
    - Window 3 Rangefinding (Virtual)
    - Window 3 Grade 7 Rangefinding Meeting
    - Window 3 Grade 6 and 8 Rangefinding Meeting
    - Obtain educator list from LDOE
    - Provide list to procurement for stipends
    - DocuSign sent to educators with instructions
- **February – March 2023**
  - **Teacher Engagement/IAP Collaboration Groups**
    - To support the work of a report redesign at the class level.
    - Impact was to redefine the class level report.
- **February 2023**
  - **Desoto IAP Collaboration Group**
    - Open Feedback Session
  - **Ouachita IAP Collaboration Group**
    - Open Feedback Session
  - **Lincoln IAP Collaboration Group**
    - Open Feedback Session
- **January – February 2023**
  - **Student Engagement Survey Grades 6-8**
    - Feedback on reports and results; how they engage in learning more about reading and writing.
    - Will support the creation of an updated Individual Student Report; will also be the basis for a Student Focus Group.
- **February 2023**
  - **LDOE TAC Meeting**
- **March 2023**
  - **Progressive Measurement** followed the “2022-23 Louisiana IAP Scaling Analysis Specifications” for IRT analysis part 1 prior to the Stocking-Lord transformation.
    - LDOE posts linking item parameters
    - Psychometric team converts linking item parameters to nominal response Model [GPCM]
    - Nominal parameter comparison
    - FlexMIRT output files prepared

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Completion Status</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2023</td>
<td>District Engagement Survey</td>
<td>Completed</td>
<td>Center, NWEA, LDOE, Ed First</td>
</tr>
<tr>
<td>February 3, 2023</td>
<td>Plaquemines Parish Focus Group</td>
<td>Completed</td>
<td>LDOE, Center, NWEA, Ed First</td>
</tr>
</tbody>
</table>
| February 4, 2023 – March 6, 2023 | Committee Meetings  
  - Window 3 Rangefinding (Virtual)  
  - Window 3 Grade 7 Rangefinding Meeting  
  - Window 3 Grade 6 and 8 Rangefinding Meeting  
  - Obtain educator list from LDOE  
  - Provide list to procurement for stipends  
  - DocuSign sent to educators with instructions | Completed         | NWEA, LDOE               |
| February – March 2023 | Teacher Engagement/IAP Collaboration Groups  
  - To support the work of a report redesign at the class level.  
  - Impact was to redefine the class level report. | Completed         | LDOE, Center, NWEA, Ed First          |
| February 13, 2023     | Desoto IAP Collaboration Group                                       | Completed         | LDOE, Center, NWEA, Ed First          |
| February 13, 2023     | Ouachita IAP Collaboration Group                                      | Completed         | LDOE, Center, NWEA, Ed First          |
| February 14, 2023     | Lincoln IAP Collaboration Group                                       | Completed         | LDOE, Center, NWEA, Ed First          |
| January – February 2023 | Student Engagement Survey Grades 6-8  
  - Feedback on reports and results; how they engage in learning more about reading and writing.  
  - Will support the creation of an updated Individual Student Report; will also be the basis for a Student Focus Group. | Completed         | Center, NWEA, LDOE, Ed First          |
| February 16-18, 2023  | LDOE TAC Meeting                                                      | Completed         | Center, NWEA, LDOE                    |
| March 2023            | Progressive Measurement followed the “2022-23 Louisiana IAP Scaling Analysis Specifications” for IRT analysis part 1 prior to the Stocking-Lord transformation.  
  - LDOE posts linking item parameters  
  - Psychometric team converts linking item parameters to nominal response Model [GPCM]  
  - Nominal parameter comparison  
  - FlexMIRT output files prepared | Completed         | LDOE, PM, Center               |
<table>
<thead>
<tr>
<th>Date</th>
<th>Activity Description</th>
<th>Status</th>
<th>Participants</th>
</tr>
</thead>
</table>
| March 13, 2023 | EOU Reporting Next Steps  
- Focus groups/interviews of ~20 teachers Teacher Engagement  
- High level feedback on test design sketches and reporting considerations.  
- Supported the decision in using the Funnel model for the spring administration | Completed | Center, NWEA, LDOE, Ed First |
| March 17, 2023 | Continue Work on Score Report Focus Groups | Completed | Center, NWEA, LDOE, Ed First |
| March 29, 2023 | Desoto Collaborative Group | Completed | Center, NWEA, LDOE, Ed First |
| March 2023 | Parent Engagement  
- Presentation to the Parent and Family Engagement Coordinators (PAFE).  
- To introduce the PAFE team to state assessment in Louisiana, and discuss possible opportunities to partner and next steps.  
- Impact was to develop new relationships created; intend to utilize interested coordinators in future planning | Completed | Center, NWEA, LDOE, Ed First |
| April 2023 | Committee Meetings  
- April Teacher Item Review Committee (Virtual)  
- Recruit educators  
- Prepare materials (training deck, passage review trackers)  
- Teacher Meeting [Unit 1: Stories Julian Tells; Whipping Boy / Unit 2: Cajun Folk Tales; Hurricanes]  
- Obtain educator list from LDOE  
- Provide list to procurement for stipends  
- DocuSign sent to educators with instructions | Completed | LDOE, NWEA, Educators |
| April 17 – July 20, 2023 | ALD Alignment Workshop Preparation (Operational - Grades 6-8)  
- Make/Review Workshop Materials (Agenda, OIBs, Presentations, etc.) - Includes LDOE Review  
- Recruit educators  
- Grades 6-8 ALD Alignment Meeting  
- Grades 6-8 Reporting Category Meeting  
- Write-up ALD alignment findings  
- LDOE reviews ALD write-up  
- Revise ALD write-up  
- LDOE signs off on ALD write-up | Completed | LDOE, MZD, NWEA, Center, Educators |
| May 2023 | Grades 6-8 Alignment for Achievement Level Descriptor and reporting Category Workshop  
- Day 1 - ALD Alignment/Reporting Category Workshop  
- Day 2 - ALD Alignment/Reporting Category Workshop  
- Day 3 - ALD Alignment/Reporting Category Workshop  
- Day 4 - ALD Alignment/Reporting Category Workshop | Completed | LDOE, NWEA, Center, Educators |
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Event Description</th>
<th>Status</th>
<th>Collaborators</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 1 – May 17, 2023</td>
<td>Committee Meetings</td>
<td>Completed</td>
<td>LDOE, NWEA, Center, Educators</td>
</tr>
<tr>
<td></td>
<td>- Grades 6-8 ALD Alignment/Reporting Category Workshop (Virtual)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grades 6-8 ALD Alignment Meeting</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Grades 6-8 Reporting Category Meeting</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Obtain educator list from LDOE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Provide list to procurement for stipends</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- DocuSign sent to educators with instructions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May – July 2023</td>
<td>Teacher Engagement Collaboration</td>
<td>Completed</td>
<td>Center, NWEA, LDOE, Ed First</td>
</tr>
<tr>
<td></td>
<td>- Groups receive feedback on reports</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Redefined class level report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Create final adjustments for use in 2023-24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 2023</td>
<td>Parent Engagement /Assessment Collaboration</td>
<td>Completed</td>
<td>Center, NWEA, LDOE, Ed First</td>
</tr>
<tr>
<td></td>
<td>- Meeting with PAFE Coordinators</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- To discuss with the coordinators: What conversations are you having with parents?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Continue to develop relationships with the PAFE coordinators:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Understand better how to engage with parents in the established framework</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 2023</td>
<td>Student Engagement Student Survey (Optional 5th grade survey)</td>
<td>Completed</td>
<td>Center, NWEA, LDOE, Ed First</td>
</tr>
<tr>
<td></td>
<td>- The focus will be on their experience with the assessment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our proposed analysis is aimed at comparing and contrasting student experiences across the W&amp;W and Guidebook forms.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Evidence to determine if the assessments related to the different curricula are providing similar student experiences.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 2022 – May 2023</td>
<td>John Hopkins Teacher Survey</td>
<td>Completed</td>
<td>LDOE, JHU, Ed First</td>
</tr>
<tr>
<td></td>
<td>- Grade 3-8 ELA Teacher Survey Summary Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Johns Hopkins Institute for Education Policy Integrated Education Resources platform.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- The Institute will work with LDOE to field the national Teacher Survey on Curriculum Use to all elementary ELA teachers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- The survey, based in part on the RAND American Teacher Panel , is designed to explore teachers’ attitudes towards their curriculum, their use of additional materials (and the sources of those materials), and their classroom practices.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 17, 2023</td>
<td>ALD Strong-Moderate-Weak Descriptor Initial Review</td>
<td>Completed</td>
<td>LDOE, Center</td>
</tr>
<tr>
<td>May 24, 2023</td>
<td>Grade 5 Writing Task and Item Analysis and Needs</td>
<td>Completed</td>
<td>Odell, SME, JHU, LDOE</td>
</tr>
<tr>
<td>May 10 – June 29, 2023</td>
<td>LDOE conducted the Spring Window 3 - Post-Test Window Validation Activities which included some new procedures from the Fall and Winter processes to account for the End of Year (EOY) requirements.</td>
<td>Completed</td>
<td>LDOE, MZD, NWEA, PM</td>
</tr>
<tr>
<td></td>
<td>- Spring Window 3 post-window validation early data pull</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- MZD posts the early data file on SFTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- LDOE notifies MZD of voids and accountability code decisions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
- LDOE provides MZD with the student demographics data for overlay file from Escholar
- MZD applies the overlay file for EOU Window 3 data file
- Spring Window 3 post-window validation data file and Blue Dot Reports submitted to LDOE
- MZD posts the Window 3 final data file on SFTP
- LDOE reviews Spring Window 3 data file and submits findings in key check tracker
- MZD/NWEA/LDOE hold stand-up meetings during post-test validation period [15-30 mins] (if needed)
- EOU reports available for LDOE’s release to IAP systems
- Date LDOE identified for EOU reports - LDOE checks report assets in ADAM for release approval

### June 2023

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity Description</th>
<th>Status</th>
<th>Responsible Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 12, 2023</td>
<td>Progressive Measurement followed the 2022-23 Louisiana IAP Scaling Analysis Specifications in order to generate the incomplete data matrices (IDMs) for Grades 6-8.</td>
<td>Completed</td>
<td>LDOE, PM, Center</td>
</tr>
</tbody>
</table>
  - CTT analysis (IAP Operational and Linking Set ELA Passages)  
  - Unit reports include RCWE and LC runs  
  - Summary spreadsheet at each grade  
  - CTT statistic check/compare  
  - Psychometric team delivers CTT results to LDOE | Completed   | LDOE, PM, Center          |
| June 13 – June 29, 2023 | Timing Data  
  - Operational Analysis [Grades 6-8]  
  - Psychometric Team Runs the Item Timing Data [i.e., Average, standard deviation, min, max, 90th percentile, 95th percentile, 99th percentile for all items]  
  - Psychometric Team Runs the Session Level Timing Data [i.e., Session level average, standard deviation, min, max, 90th percentile, 95th percentile, 99th percentile]  
  - Psychometric Team Runs the Test Level Timing Data [i.e., Test level average, standard deviation, min, max, 90th percentile, 95th percentile, 99th percentile]  
  - Summarize Unit/Module Test [Gather data by unit/module test and across units/modules for grades 6-8, all forms/test] | Completed   | LDOE, PM, Center          |
| June 28 – July 6, 2023 | Progressive Measurement followed the 2022-23 Louisiana IAP Scaling Analysis Specifications for IRT analysis after LDOE item evaluations.  
  - Rerun LC and RCWE Runs (After Item Keep/Reject Decisions) and Do Stocking Lord Transformation  
  - FlexMIRT Output Files  
  - Linking Constants  
  - ItemParameter Files  
  - Psychometric Team Delivers Final Item Parameters to LDOE | Completed   | LDOE, PM, Center          |

### July 2023
<table>
<thead>
<tr>
<th>Date Range</th>
<th>Task Description</th>
<th>Status</th>
<th>Responsible Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 3 – July 25, 2023</td>
<td>Score Table Creation (Operational - Grades 6-8)</td>
<td>Completed</td>
<td>LDOE, PM, Center, MZD</td>
</tr>
<tr>
<td></td>
<td>• RSSS Tables</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Generate Raw Score-to-Scale Score Tables [RSSS] (Overall Scores and by Reporting Category)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• IAP Score Table (RSSS) - Psychometric Delivers Generates RSSS Tables (Overall and Reporting Category)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Score Table Compare</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• IAP Score Table (RSSS) - Psychometric Delivers RSSS Tables (Overall and Reporting Category) to LDOE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• NWEA Delivers Executive Leadership Debrief Materials [i.e., Scaling and Equating Summary, Comparability Analyses, ALD Summary]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Reviews Summary and Sends Questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Reviews Summary with Executive Leadership</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Signs Off on RSSS Tables</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• NWEA Delivers RSSS Tables (Overall and Reporting Category) to MZD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• MZD Validates the EOY File with RSSS Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Review of EOY Data File and Blue Dot Reports</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Signs Off on EOY File and Blue Dot Reports in ADAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• EOY Reports Available for Release</td>
<td></td>
<td></td>
</tr>
<tr>
<td>June – July 6, 2023</td>
<td>Impact Analysis</td>
<td>Completed</td>
<td>LDOE, PM, Center</td>
</tr>
<tr>
<td></td>
<td>• Modified IDMs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• IDM/Modified IDMs Compare</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Impact Data Table</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Impact Data Table Compare</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Summarize Number and Percent of Students in Each Performance Level (Overall and Reporting Category)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Psychometric Team Delivers Impact Data to LDOE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>June 5 – July 26, 2023</td>
<td>Progressive Measurement analyzed the field test data in Grade 5 including the pilot for Crawfish. This process was replicated for each test window (Window 1-3). Form/Item Analyses - Field Test (Grade 5)</td>
<td>Completed</td>
<td>LDOE, PM, Center</td>
</tr>
<tr>
<td></td>
<td>• Unit Reports [HTML]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Other IRT, CTT, and DIF Deliverables [TBD]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Enhanced Summary Worksheet (w/ IRT and DIF Results)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Timing Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>July 10 – July 21, 2023</td>
<td>July 21 Grade 5 FT Data Review Meeting Prep</td>
<td>Completed</td>
<td>Odell, Progressive Measurement, LDOE</td>
</tr>
<tr>
<td></td>
<td>• Psychometric Team Delivers Data to Content Team</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Content Team Prepares Recommendations and Materials for LDOE Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Content Team Delivers Materials to LDOE for Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Reviews Materials</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• LDOE Signs Off on Materials</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data Review Meeting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer and Fall 2023</td>
<td>Parent Engagement Survey</td>
<td>Completed</td>
<td>NWEA, Center, LDOE, Ed First</td>
</tr>
<tr>
<td></td>
<td>• Expected to support an understanding of parent perspective on reports.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Expected to support the redesign of an individual student report.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>July 2023</td>
<td>ALD Summary</td>
<td>Completed</td>
<td>NWEA, Center, PM, LDOE</td>
</tr>
</tbody>
</table>
- This summary provides the processes used and the outcomes achieved for the achievement level descriptor (ALD) alignment and reporting category meeting.

<table>
<thead>
<tr>
<th>Month</th>
<th>Activity</th>
<th>Status</th>
<th>Responsible Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>July – August 2023</td>
<td>Comparability Analyses</td>
<td>Completed</td>
<td>Center, PM</td>
</tr>
</tbody>
</table>
LEAP ELA Guidebooks Assessment (LDOE and NCIEA)

Louisiana’s Value-Added Model (VAM)

Laura Boudreaux from LDOE gave an overview of Louisiana’s VAM model. The VAM measures students’ success compared to similar peers year to year. The model predicts how well students will perform on the assessment in comparison to their peers with similar prior assessment scores and background. Once a student has taken state assessments, the model shows the extent to which his or her achievement was on target with what was expected (i.e., student expected score). The difference between a student’s actual achievement and his or her expected achievement is known as the “value added.”

Laura shared the data used in the VAM calculations. She noted that the student data elements used in the model have changed over the years. One example is the expansion of the economically disadvantaged combined indicator to different types of SES indicators. Results from the ELA Guidebooks Assessments will not be used in VAM for now but may be in the future. Science and social studies are used in the VAM for teacher evaluation, but not for accountability use. Laura provided technical details of the VAM calculations and shared an example of how it is for students, teachers (for effectiveness), and schools (for accountability). She also gave a history of how VAM was created.

TAC Discussion and Feedback

The TAC complimented Laura’s presentation. One TAC member pointed out that using norm referenced criteria, such as percentile ranking of teachers, for effectiveness can be potentially misleading. It may obfuscate the growth that teachers are making. The TAC suggested a more criterion-based approach and interpretation based on the VAM values, acknowledging that it may be more difficult to understand.

The TAC also recommended LDOE consider anchoring the VAM calculation to a base year so that growth over time can be captured. The TAC was pleased to hear that VAM is one of multiple measures used to evaluate teacher effectiveness.

Innovative Assessment Program (IAP) Overview

Nathan Dadey from the Center provided an overview of Louisiana’s ELA Guidebook assessment (also known as the Innovative Assessment Program, or IAP). He provided some context for the IAP, including how it will be used in the state’s school performance score. The analysis results in Nathan’s presentation were based on data collected from Windows 1 and 2 of the 7th grade ELA assessment aligned to the ELA Guidebooks 2.0 curriculum.

Nathan reminded the TAC of the operational design, which includes linking items to LEAP 2025. He reviewed the theory of action for the IAP. The key claim is at the summative level and that students typically work efficiently with a range of fiction and nonfiction texts from, and related to, the ELA Guidebooks 2.0 curriculum units. One key clarification based on this claim is that LDOE prefers a method of aggregation that treats each unit assessment equally, regardless of window.
**TAC Discussion and Feedback**

The TAC noted that it is important to make clear in the theory of action how we are departing from the traditional claim of end-of-year (EOY) achievement. It will be critical to provide validity evidence across years that scores from the IAP are comparable with LEAP. For example, consider looking at student trajectories over time for both programs – this could be evaluated with LEAP as the program of reference (e.g., how a student with 7th grade IAP score compares with one in LEAP when they take the 8th grade test) and in the VAM model.

**IAP Scale Creation**

Nathan described the data used in the Windows 1 & 2 scaling analysis. He shared the form sets along with some descriptive statistics for each set of unit assessments. The AB/BA forms were spiraled at the student level.

Nathan then shared the approach and results of the scaling analysis. Two approaches were considered based on the TAC’s previous feedback: separate calibration and pooled calibration. The 2PL and GPC IRT models were calibrated with the EM algorithm. Theta estimates were created using EAP. Parameters for linking items were fixed in the calibration to place the IAP items on the LEAP scale. Besides a few items with higher a-values, the item parameters estimates and their CSEMs from each window were similar to the pooled values.

**TAC Discussion and Feedback**

To evaluate whether differences in the linking item statistics between windows are meaningful, the TAC suggested asking content experts to review the flagged items and to not make too much of the differences until we see how it plays out on the overall test after all three windows.

One TAC member questioned the choice to fix the item parameters for linking items. Consider freely calibrating the linking items and apply chained equating across windows using the Stocking-Lord approach. If the outcomes come out similar, then the chained equating approach is recommended going forward.

Consider taking a closer look at the potential structures that underlie the item components or sets of items (e.g., linear logistic test models) across time points. If appropriately modeled, it may be possible to estimate thetas across time points to measure growth.

The TAC feels that it is premature to transition to pre-equating next year. If the double testing field test design (i.e., taking all three windows + LEAP 2025 at EOY) in grades 6 and 8 is stronger, then it should be used as a yardstick to compare results from the other designs.

One TAC member asked whether the end-of-unit (EOU) reporting is focused on where a student is at the end of the unit, or where they likely will be at the end of the year. Nathan clarified that it will be the former – unit-specific reporting. In that case, the TAC felt that the benefits of the initial linking would be minimal for EOU reporting.
Summative Score Creation

Nathan shared a comparison of the theta distributions from the pooled and separate calibrations. The distributions were similar and the correlations of thetas between approaches (separate vs. pooled) were high. The correlations between windows in the separate calibration is lower, but comparable when correlated to prior scores.

TAC Discussion and Feedback

The TAC asked which scaling and scoring approaches are the default. Nathan said the pooled approach to scaling, with a single overall theta. The TAC suggested revisiting or reconsidering how to articulate the claim that each unit assessment is treated equally in the score aggregation method given that the effective weighting of units under the pooled approach depends on the discrimination of items within each window. The TAC also recommended examining literature on composite scores that goes beyond the current compensatory approaches and may yield unique information. Such approaches do not need to be constrained by the comparability requirements.

The TAC asked what the next steps for scaling and scoring were. Nathan said:

- Perform separate linking with the Stocking-Lord approach
- Generate model fit statistics
- Draft analysis specifications
- Run scaling and scoring analysis with preliminary Window 3 data
- Develop achievement level descriptors (ALDs)

An ad-hoc TAC meeting was set for July 6, 2022, to potentially follow-up on these next steps. LDOE noted that the BESE meeting during which results from this year’s IAP will be reviewed is scheduled for August 23-24, 2022.

Nathan asked for the TAC’s thoughts about the potential of reporting expected raw scores. The TAC questioned whether report users can appropriately interpret expected raw scores given the already complex assessment model, especially with the intended classroom use. One TAC member wondered whether there is a way to convey scores that are comparable and formative. Think about a way to demystify scores – send the message that different scores for different purposes are appropriate. The TAC agreed that assessment literacy is especially important with the IAP, and it should start with the staff at LDOE.

Nathan also asked how we can better understand the measurement error under each scaling/scoring approach. The TAC acknowledged that the standard error for the pooled estimate is an underestimate because of violation of assumptions. The TAC recommended looking into literature for the measurement error of composite scores based on separate estimates.

LDOE noted that once the IAP launches, LEAP 2025 will no longer be available. Will having scores from two windows be sufficient? Nathan noted that while it may be supportable technically, this has potentially thorny policy implications. The TAC’s advice is to not plan for the exceptions but focus on...
what LDOE intends for the program first and address these exceptions and unintended consequences as a next step.

**ALD Alignment Approach & Reporting Category Classification**

Nathan described the ALD Alignment Method previously recommended by the TAC. The approach is being used for NAEP and adapted for the IAP context. One key difference is that the IAP approach does not modify the LEAP ALDs. Instead, the process is intended to inform future item development and to produce content-based summary statements that can be used for IAP reporting. An initial study based on this approach is planned for April with LDOE staff.

Eric Moyer from Pearson shared that in their experience implementing the ALD Alignment Method with NAEP. They grouped items by genre because of the different skill sets needed for each type of passage (literary vs. informational). Nathan suggested that we can try this with the IAP process.

**TAC Discussion and Feedback**

The TAC offered the following feedback on the proposed ALD alignment process.

- Consider organizing the item in the quasi-ordered item booklet by ‘chapters’, i.e., by passages, and with increasing difficulty within chapters.
- Considering an ID matching or body of work approach.
- Incorporate a step in which the panelists become more familiar with the ALDs and the appropriate level of granularity for the summary statement.
- Include educators and content experts in the process to make this process as independent as possible. The perception of independence could be important to the public.
- Get the panel’s endorsement that the LEAP ALDs can be used for reporting, as well as where alignment is weak. It could start earlier in the process where we asked panelists to judge the alignment of items to ALDs as they review the items. Summary questions at the end could then be something like: “How accurately do the LEAP ALDs capture items on the IAP?” “How confident are you that the ALDs could be used to support interpretation of performance on the IAP?”
- Whether the summary description should be included with the ALDs for reporting is ultimately a policy decision.

For the reporting category classifications, the TAC was not in favor of using the three sections of the IAP (i.e., knowledge of unit texts, application of unit knowledge, and synthesis and expression of knowledge across texts) as reporting categories. These sections do not align to the LEAP reporting categories, and it is not clear how teachers, parents, or users could use these in a formative way. The TAC encouraged LDOE to think more innovatively about what type of information could provide better guidance. It would be useful to give students more direct instruction or feedback about how they can improve for each reporting category. It seems like a lost opportunity to not leverage this information if it is available. Parents need to know what they should do next for their child to help them improve in each category.

Nathan asked the TAC what other types of reporting categories could be considered. The TAC felt that the current LEAP reporting categories are more actionable for parents and students. It may make sense
to design different reports for different users with custom reporting categories. Consider convening focus groups with various users to sharpen what information is most helpful. If the reporting categories are based on the cognitive processes (knowledge, application, etc.), then actions that can be taken to improve in each section should be offered.

One TAC member asked whether there are learning progressions underlying the design of the IAP that could be leveraged for the reporting categories. Traditional reporting categories may not be useful in supporting the instructional use intended by the IAP. This might be fine if the IAP is meant to be a mini-summative but is questionable for a through-year design.

There is nothing wrong with focusing on the consequences of the test as a start. That is, it is okay if our initial goal is to drive good instructional practices so that teachers are not teaching to the test, as they are with LEAP. This does not imply that reporting is not a critical issue. Practically speaking, it does not seem appropriate to be focusing deeply on this issue right now. Consider concentrating on it after the current operational administration – plan to do more comprehensive research, including psychometrics and user testing as soon as you can and before initiating any reporting.

One TAC member shared their experience with a writing pilot study that resulted in a map or continuum based on expert input of writing items that can be used to provide actionable feedback to teachers and parents. Another TAC member learned from working with teachers that “next steps” feedback is usually less helpful than grouping students based on the type of support they need. They cautioned that if these assessments have high stakes, it will lead to the field treating them as mini-assessment and promoting “test prep” behaviors.

**Action Items**

- Consider using a more criterion-based approach and interpretation based on the VAM values and anchoring the VAM calculation to a base year so that growth over time can be captured.
- Make clear in the theory of action how we are departing from the traditional claim of end-of-year (EOY) achievement.
- To provide validity evidence across years that scores from the IAP are comparable with LEAP, consider looking at student trajectories over time for both programs with LEAP as the program of reference.
- Ask content experts to do an initial review of the linking items with substantial differences in their statistics between the first two windows. However, do not make too much of the differences until we see how it plays out on the overall test after all three windows.
- Freely calibrate the linking items and apply chained equating across windows using the Stocking-Lord approach. If the outcomes come out similar to the fixed parameter item calibration approach, then use the chained equating approach going forward.
- Take a closer look at the potential structures that underlie the item components or sets of items (e.g., linear logistic test models) across time points.
 ➢ Revisit how to articulate the claim that each unit assessment is treated equally in the score aggregation method given that the effective weighting of units under the pooled approach depends on the discrimination of items within each window.

 ➢ Examine literature on composite scores that goes beyond the current compensatory approaches and may yield unique information. Such approaches do not need to be constrained by the comparability requirements.

 ➢ Follow up on the next steps for scaling and scoring:
  o Perform separate linking with the Stocking-Lord approach
  o Generate model fit statistics
  o Draft analysis specifications
  o Run scaling and scoring analysis with preliminary Window 3 data
  o Develop achievement level descriptors (ALDs)

 ➢ Think about how to demystify scores, sending the message that different scores for different purposes are appropriate. Assessment literacy is critical with the IAP, and it should start with the staff at LDOE.

 ➢ Look into literature for the measurement error of composite scores based on separate estimates.

 ➢ Consider the TAC’s suggestions (see page 9) on the proposed ALD alignment process.

 ➢ Convene focus groups with various report users to sharpen what information is most helpful. Consider designing different reports for different users with custom reporting categories.

**Upcoming TAC Meetings**
- Fall 2022: October 27-28, 2022
- Spring 2023: February 16-17, 2023
Louisiana Innovative Assessment Program (IAP) Update

February 16, 2023
Louisiana Department of Education
Technical Advisory Committee Meeting

Audra Kosh, Senior Research Scientist, NWEA
Nathan Dadey, Senior Associate, Center for Assessment
Agenda

1. IAP Implementation
2. Reports in 2023-2024 and beyond
3. IAP Research Agenda
Questions for the TAC, Continued

IAP Operational Psychometrics

- What does the TAC recommend for the linking design moving forward?
  - In addition to LEAP linking items, should we start including linking items from the previous year’s IAP as well?

- For the ALD alignment workshop and reporting category descriptors:
  - How many educators per grade should we aim to recruit?
  - Overall feedback on plans

- What considerations does the TAC recommend for removing or retaining poor performing items from unit reports, before we know how those items perform when the full year of data are aggregated?
Questions for the TAC, Continued

End-of-Unit Reporting Development
● Does the TAC have recommendations for a process or structure for the focus groups or individual interviews? E.g., specific resources from published studies?
● What questions or probes does the TAC recommend we use to investigate
  ○ Score interpretations?
  ○ Instructional relevance?
Questions for the TAC, Continued

IAP Research Agenda

- What analyses does the TAC recommend to evaluate whether achievement gaps between students of different socioeconomic backgrounds were less on IAP as compared to LEAP?
- What studies does the TAC recommend be conducted to support statewide transition to IAP in 2025?
- What considerations does the TAC have for maintaining the IAP assessment as part of the LEAP 2025 assessments vs. breaking trend as a new assessment?
- The IAP is intended to better connect state assessment to curricula, and in turn, encourage instructional shifts. What specific studies would help demonstrate the value, or not, of this program relative to LEAP 2025?
- What does the TAC recommend for the linking design moving forward?
IAP Implementation

2022-23

a. Operational psychometrics plans
b. ALD alignment workshop plans
SY 2022-23 Operational Psychometrics Workflow

**Data Cleaning and Item Evaluation**
- Data Preparation
- Classical Stats and Preliminary Calibration (2PL and GPCM)
- RCWE fixed run
- LC fixed run
- Evaluate Items (operational items and anchor item stability)

**Final Item Parameters**
- Final Calibration
- RCWE fixed run
- LC fixed run
- Stocking-Lord transformation of LC items to RCWE scale
- Retain parameters from RCWE run, and SL-transformed parameters from LC run

**Scoring**
- Raw Score to Theta and CSEM Tables
- Raw Score to Scale Score Tables
- Impact Data, ALD Alignment, Reporting Category ALDs
1. Each LEAP passage set contains 6-8 items with parameters from the LEAP scale.

2. All students take LEAP passage sets 1 and 2, alternating between window 1 and 2.

3. LEAP passage sets 1-4 are spiraled in window 3.
Across-year Linking/Equating

Currently, IAP forms include linking items to LEAP. SY 2023-24 is the third year of operational IAP forms in grade 7 and the first year of operational IAP forms in grades 6 and 8.

- What does the TAC recommend for the linking design moving forward? In addition to LEAP linking items, should we start including linking items from the previous year’s IAP as well?
Item Evaluation for Unit Report and EOY Reports

Students in IAP receive four reports a year: one unit report for each of the three windows, plus one end of year report that aggregates all three windows.

We plan to start evaluating item statistics within each window (i.e., data from a single window for unit reports) in addition to end of year (i.e., data from three windows aggregated for end of year reports).

- What considerations does the TAC have for removing or retaining poor performing items from unit reports, before we know how those items perform when the full year of data are aggregated?
IAP Implementation

2022-23

a. Operational psychometrics plans
b. ALD alignment workshop plans
ALD Alignment Refresher from SY 21-22

- 3-day workshop with 6 Louisiana educators in May 2022
- Output is ratings of IAP items to LEAP achievement levels as weak, moderate, or strong, and recommendations for improving alignment
- Half day allocated to reporting category standard setting and descriptors

**Item Statements**
- Individually: Develop item statements describe the knowledge and skills required to receive a specific score point.

**Summary Statements**
- Individually, then small group, then whole group: Develop statements that summarize the item statements for the items grouped by each LEAP 2025 ALD.

**Alignment Judgments**
- Individually: Judge the degree of alignment between each summary statement and the matching LEAP 2025 ALD.

**Alignment Recommendations**
- Whole Group: Make recommendations to improve alignment between the IAP and the ALDs.
### ALD Alignment and Reporting Category Plans for SY 22-23

<table>
<thead>
<tr>
<th></th>
<th>Implemented for SY 21-22</th>
<th>Proposed for SY 22-23</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALD Alignment Judgements</strong></td>
<td>Workshop style of educator ratings</td>
<td>Workshop style of educator ratings</td>
</tr>
</tbody>
</table>
| **Reporting category performance levels for knowledge, application, synthesis** | 2 levels:  
  - Working toward expectations  
  - Meets expectations           | 3 levels:  
  - Weak  
  - Moderate  
  - Strong  
  (same labels as LEAP)        |
| **Reporting category cut scores**   | Standard setting held with educators and LDOE policy committee | Utilize same cuts as LEAP                      |
| **Reporting category descriptors**  | Written by educators during ALD alignment workshop          | Written by educators during ALD alignment workshop |
| **Format**                          | Zoom, one facilitator for one grade (G7)                    | Zoom, running concurrently with one facilitator per grade (G6-8) |
| **Timing**                          | May 2022, with items from window 1 and 2                  | May 2023, with items from window 1 and 2         |
ALD Alignment and Reporting Category Descriptors

- How many educators per grade should we aim to recruit?
- What is the TAC’s feedback on the overall ALD alignment workshop and reporting categories plan?
  - Moving to the same reporting category levels and cut scores as LEAP
  - Running the ALD alignment workshop across grades concurrently with three different facilitators
  - Timing the workshop during May 2023, with window 1 and 2 items
End-of-Unit Reporting Development Plan

Spring & Summer 2023

a. Approach
b. Implementation
Current Reports and Supports

End-of-Unit:

- Individual Student Score Report
- Classroom Roster Report
- Interpretive Guide
- Teacher Score Report Guidance

End-of-Year:

- Individual Student Score Report
- Classroom Roster Report
- Interpretive Guide
Goals:
- Develop reports and supports that better connect to instruction.
- Better integrate supports into reports.

Unlike other statewide programs, we know what *was* taught and what *will be* taught. We hope to leverage this into more concrete guidance for stakeholders.
Instructional Context

Educators implement Guidebook instruction with greater fidelity, drawing on the guidebook practices, including those outlined in the diverse learners cycle.

Educators adjust Unit 2 instruction based on assessment results and information from their classroom practice.

Improved Student Learning Throughout Unit 2

- **Input**
- **Action Mechanism**
- **Effect**

|------|------|------|------|------|-------|-------|-----|
Framing and Prioritization

- Any score report can, at best, play a small role in supporting instructional decision making.
  - Our development is prioritized around two high level reports and corresponding supports.

<table>
<thead>
<tr>
<th>Who is the report about?</th>
<th>Who is the report for?</th>
<th>What are they supposed to do with the report?</th>
</tr>
</thead>
</table>
| Student                 | Parent                 | • Understand if their student is **doing well or not**.  
                          |                        | • **Signal if they need help**. |
| Classroom               | Teacher                | • Identify specific areas of **weaknesses at the classroom level to inform the next unit of instruction**. |
Student Report & Supports for Parents

● **Purpose.**
  ○ Inform parents and caregivers about student performance so that they can support their students as they move into the next unit of instruction.

● **Planned Materials.**
  ○ A parent friendly report
  ○ Revised guidance for teachers on:
    ■ Sending score reports with students home
    ■ A process or protocol on having conversations with parents on their students’ reports for those who want follow up, including considering classroom performance
  ○ Revised guidance for parents and caregivers on:
    ■ Talking to their students about their performance
    ■ What next steps can be done to support their child’s learning
Classroom Report & Supports for Teachers

● **Purpose.**
  ○ Capture classroom level performance in ways that (1) help educators better understand the variation in performance in their classrooms and and (2) connects student performance to instructional next steps

● **Planned Materials.**
  ○ A report that provides groupings of students tied to instructional next steps
  ○ Revised guidance for teachers:
    ■ That connects current classroom performance to the next unit of instruction, potentially including guidance on what should be prioritized in instruction moving forward
End-of-Unit Reporting Development Plan

Spring & Summer 2023

a. Approach
b. Implementation
Proposed Development Process

- Develop multiple “rough” mockups informed by:
  - A scan of reports in the field
  - Feedback from teachers and students from multiple rounds of surveys
  - Early visioning work with LDOE about uses of reports

- Internal review with LDOE

- Focus Groups with “IAP Collaborative Groups”
  - Currently have partnered with four groups of teachers and leaders from four districts (~22 teachers and leaders) who are implementing the IAP
  - Present a limited number of prioritized mockups for feedback (see TAC question on appropriate questions and protocols)
  - Ideally, provide mockups based on real data to better ground feedback.

- Revision based on Focus Group Feedback
Questions for the TAC

Our approach is to generate multiple possible score reports, then review with the field.

- Does the TAC have recommendations for a process or structure for the focus groups or individual interviews? E.g., specific resources from published studies?
- What questions or probes does the TAC recommend we use to investigate
  - Score interpretations?
  - Instructional relevance or utility?
IAP Research Agenda

a. Opportunity to learn research
b. Recommendations for future studies
Propensity Score Matching

- For revised comparability analyses and other additional research studies, we have made a matched sample using propensity score matching.
- Covariates: gender, race/ethnicity, migrant status, homeless status, foster care status, military affiliation status, economically disadvantaged status, LEAP scale score from prior year (grade 6 in SY 20-21)
- Nearest neighbor matching without replacement
- Resulting sample has 4,424 students that took IAP in grade 7 in SY 21-22, and 4,424 matched students that took LEAP (selected from 42,182 LEAP test takers)
Opportunity to Learn Research for (G7 from SY 21-22)

● **Background**
  ○ A key goal of IAP is to reduce disparities in reading achievement due to prior background experiences that not all students have
  ○ Thus we are hoping the achievement gap on IAP will be smaller than the achievement gap on LEAP when comparing students of different socioeconomic backgrounds
  ○ At the same time, all steps of the scaling and linking were designed so that IAP scores are comparable to LEAP scores

● **Research Question**
  ○ Is the gap in performance between economically disadvantaged and non-economically disadvantaged students less for students that took the IAP as compared to LEAP?
Opportunity to Learn Research (G7 from SY 21-22)

Question for TAC

1. What analyses does the TAC recommend?
2. How should the research study be framed relative to the comparability analysis?
IAP Research Agenda

a. Opportunity to learn research
b. Recommendations for future studies
Recommendations for Studies

1. What studies does the TAC recommend be conducted to support statewide transition to IAP in 2025?

   In terms of:
   ● Reporting
   ● Assessment Use
   ● Sustainability & Scalability

2. What considerations does the TAC have for maintaining the IAP assessment as part of the LEAP 2025 assessments vs. breaking trend as a new assessment. What analysis might support these considerations?

2. The IAP is intended to better connect state assessment to curricula, and in turn, encourage instructional shifts. What specific studies would help demonstrate the value, or not, of this program relative to LEAP 2025?
Audra Kosh with NWEA and Nathan Dadey with the Center for Assessment led the presentation on the Innovative Assessment Program (IAP). First, they provided the TAC with context about the program, describing the development work to date, the overall structure of the test, and the reporting categories. The IAP program is characterized by three unit assessments during the year which together contribute to a single summative score reported on the current statewide assessment scale. The end-of-unit reports are designed to provide information that helps guide instruction and support during the year.

Audra and Nathan described the operational psychometrics workflow and desired comparability between LEAP and IAP. The comparability claim is that a student would receive the same scale score and performance level regardless of whether they took LEAP or IAP. Nathan described the linking/equating design developed to support that claim. The TAC was asked to advise about whether linking items from the previous year’s IAP should be included as part of the design.

Next, Audra and Nathan reviewed the previous ALD alignment workshop from May 2022 and how that informed plans for the workshop proposed for 2023. The main change is to use existing reporting category cuts and labels from LEAP (strong, moderate, weak). They reviewed the proposed reporting categories describing the team’s plans to generate multiple possible score reports and review with the field. They invited TAC’s feedback regarding questions and methods to guide that review.

Finally, Audra and Nathan briefly reviewed the proposed IAP research agenda which included propensity score matching analyses to examine comparability claims and student group analyses to examine differences between students participating in each assessment program.

TAC Discussion and Feedback

The TAC focused a considerable portion of its discussion on the topic of comparability. They questioned whether the claim of interchangeable scale scores or ‘alternate form comparability’ is an appropriate or achievable target. One advisor asked if the IAP and LEAP are comparable at that level, by what criteria...
would one prefer the innovative assessment? What would have to be better and how much would it have to be better (e.g., with respect to instructional utility) to defend the IAP as an alternative to or replacement for LEAP? The TAC acknowledged that to some extent these are policy questions that go beyond their purview. However, it’s important to be clear about the program’s goals and acceptable conditions/ constraints (e.g., time, cost, comparability). Doing so will help LDOE develop suitable plans and evaluation criteria to guide the next steps.

Evaluation criteria should include evidence that students who lack prior knowledge are demonstrating gains and that educators find it coherent with good instructional practice.

With respect to comparability, some advisors worried that a stringent comparability claim (i.e., student level alternate forms) could limit the extent to which the IAP can achieve its goals. In other words, if perfect comparability is achieved, could the program be regarded as a success? It may be preferable to relax comparability preferring concordance to alternate forms or focus on comparability at an aggregate level instead of a student level.

The LDOE noted that a high degree of comparability is needed to support the IAP’s inclusion in accountability, specifically that scores can be used in the state VAM model in a fair and defensible manner. The TAC countered that comparability could be relaxed without obviating plans to use IAP scores as part of the state growth model.

Another question the TAC discussed at length is whether the current calibration and linking design masks the progress of students taking the IAP. In other words, is there anything we’re doing now that is negating the learning gains for IAP examinees? The TAC responded that there was no clear evidence that current methods are negating learning gains but agreed it should be a focus of ongoing analyses. Such analyses should include model and person fit.

With respect to including prior IAP items in the linking design, the TAC noted this decision should be considered in the context of a number of tradeoffs. For example, will this limit available content for reporting claims or restrict range and precision? Overall, the TAC did not feel there was a compelling case to include prior IAP items, which is at least partially due to the TAC’s concerns that the comparability claims for the IAP should be relaxed as previously described.

With respect to the criteria for dropping items, the TAC noted their advice on this topic in the prior session was applicable in the context of the IAP as well. Obviously, an item that is flawed or incorrect should be dropped. Otherwise, apply statistical criteria and incorporate a qualitative review to inspect items flagged for empirical reasons to make sure the test reflects key content and that adjustments do not mask achievement patterns.

The TAC further noted that the innovative nature of this program suggests that experts should review items to help improve their understanding of high-quality items. Start by having educators/ experts identify features of items that are thought to be high or low quality absent any data. When data are available, does that confirm or call into question their hypotheses about quality? A system of inquiry...
shaped by expert feedback and data analyses can help developers grow in their understanding of the characteristics and features of items that should be prioritized.

With respect to the planned reporting categories, the TAC suggested prioritizing analyses to evaluate reliability (e.g., quadratic weighted kappa). It’s important to make sure the scores are trustworthy and not ‘capitalizing on error.’

For the workshop, the TAC agrees 6-8 educators is an appropriate target and using categories of strong, moderate, and weak classifications should be continued.

The TAC complimented the IAP partners on the work being done to produce report mock-ups and vet them with the field. This is often discussed but rarely implemented in practice. Ensure the reviews focus on how educators understand and would apply the information as opposed to whether they ‘like’ the reports. One advisor suggested that an extension of the evaluation plan might involve trying out different reports in the field and seeing if they lead to different instructional practices.

The TAC also commended the partners on differentiating reports for different audiences (i.e., educators and parents.).

The TAC offered the following suggestions in response to the proposed research and evaluation plan:

- Ensure there is a prominent focus on academic impact overall and by student group and ability, as there may be some students who performed better on IAP versus LEAP which could be masked by examining overall impact only. Mixed effects can help account for school effects.
- Ensure qualitative information is collected via surveys, focus groups, or other approaches. It’s important to ‘go deep’ to understand the specific mechanisms that may promote intended or unintended outcomes.
- Build out a theory of action for the instructional and support actions that are thought to be most important and conduct observations to determine the extent to which these practices are implemented.

Action Items

- Clarify the goals, priorities, and constraints of the IAP and develop evaluation criteria with respect to these conditions to guide future development.
- Based on these goals, priorities, and constraints, consider relaxing the comparability claim of interchangeable scores between IAP and LEAP. TAC members expressed concern this could inhibit innovation.
- Conduct analyses to determine whether and to what extent current calibration and linking decisions are masking the progress of students taking the IAP. Include model and person fit.
- Ensure criteria for evaluating items for inclusion in scoring consider both statistical checks and qualitative reviews.
- To improve understanding of high-quality items, consider having experts articulate characteristics and features of high-quality items before administration and evaluate this feedback with data post-administration. Use feedback to improve item development.
• Evaluate reliability of results produced for each reporting category.
• Continue with plans to vet draft reports with the field. Ensure the reviews focus on how educators understand and would apply the information.
• Make-sure the research and evaluation plan includes a study of academic impact overall and by student group and ability range; consider using mixed-effect models. Develop deeper understanding of the practices that make a difference with qualitative information such as surveys, focus groups, and observations.
Process and Procedure
Overview

Progressive Measurement
Operational Forms (Grades 6-8)
### Instructional Units and Tests/Forms

<table>
<thead>
<tr>
<th>Grade</th>
<th>Unit</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Dust</td>
<td>DUST-A</td>
</tr>
<tr>
<td></td>
<td>Hatchet</td>
<td>HATCHET-A</td>
</tr>
<tr>
<td></td>
<td>Steve Jobs</td>
<td>JOBS-A</td>
</tr>
<tr>
<td></td>
<td>Stones</td>
<td>STONES-A</td>
</tr>
<tr>
<td>7</td>
<td>The Giver</td>
<td>GIVER-C</td>
</tr>
<tr>
<td></td>
<td>Written in Bone</td>
<td>BONE-C</td>
</tr>
<tr>
<td></td>
<td>A Christmas Carol</td>
<td>CC-B</td>
</tr>
<tr>
<td></td>
<td>Behind the Sceness</td>
<td>SCENES-B</td>
</tr>
<tr>
<td>8</td>
<td>Flowers for Algernon</td>
<td>FLOWERS-A</td>
</tr>
<tr>
<td></td>
<td>Call of the Wild</td>
<td>COTW-A</td>
</tr>
<tr>
<td></td>
<td>Tell Tale Heart</td>
<td>TTH-A</td>
</tr>
<tr>
<td></td>
<td>Sugar</td>
<td>SUGAR-A</td>
</tr>
</tbody>
</table>
Number of Items and Points by Reporting Category

On each Form

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Application</th>
<th>Synthesis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Grade</td>
<td>Unit-Form</td>
<td>Administration Window</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>DUST-A</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>HATCHET-A</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>JOBS-A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STONES-A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>GIVER-C</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>BONE-C</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>CC-B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCENES-B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>FLOWERS-A</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COTW-A</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>TTH-A</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>SUGAR-A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Patterns

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Window 1</th>
<th>Window 2</th>
<th>Window 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>DUST-A</td>
<td>HATCHET-A</td>
<td>STONES-A</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>HATCHET-A</td>
<td>DUST-A</td>
<td>STONES-A</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>HATCHET-A</td>
<td>JOBS-A</td>
<td>STONES-A</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>GIVER-C</td>
<td>CC-B</td>
<td>SCENES-B</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>BONE-C</td>
<td>CC-B</td>
<td>SCENES-B</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>BONE-C</td>
<td>GIVER-C</td>
<td>SCENES-B</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>GIVER-C</td>
<td>BONE-C</td>
<td>SCENES-B</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>FLOWERS-A</td>
<td>TTH-A</td>
<td>SUGAR-A</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>FLOWERS-A</td>
<td>COTW-A</td>
<td>SUGAR-A</td>
</tr>
</tbody>
</table>
## LEAP Linking Sets

<table>
<thead>
<tr>
<th>Grade</th>
<th>Link Set</th>
<th>Number Of Items</th>
<th>Administration Window</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Cranes</td>
<td>8</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Woman</td>
<td>7</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Underground</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Happiness</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Summit</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Crossing</td>
<td>8</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Magic</td>
<td>8</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Bird</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Why</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Home</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Chronicles</td>
<td>7</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>6</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Genius</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Archeology</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alyssa</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Courage</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>
Overlay File

• Received from MZD soon after each test administration window
• One line/record per student
• Cumulative
  • W1
  • W1W2
  • W1W2W3
• Flat fixed and/or comma separated
• Student demographics, School, District, Teacher data
• Item data (IAP and LEAP Linking Items)
  • Item ID
  • Response
  • Score
  • Time
Data Analysis
SY 2022-23 Operational Psychometrics Workflow

Data Cleaning and Item Evaluation

- Data Preparation
- Classical Stats and Preliminary Calibration (2PL and GPCM)
- RCWE fixed run
- RCWE free run
- Evaluate Items (operational items and anchor item stability)

Final Item Parameters

- Final Calibration
- RCWE fixed run
- LC fixed run
- Stocking-Lord transformation of LC items to RCWE scale
- Retain parameters from RCWE run, and SL-transformed parameters from LC run

Scoring

- Raw Score to Theta and CSEM Tables
- Raw Score to Scale Score Tables
- Impact Data, ALD Alignment, Reporting Category ALDs
# IDM (Incomplete Data Matrix) Structure

## Grade 7

<table>
<thead>
<tr>
<th>Student Info</th>
<th>The Giver (Form C)</th>
<th>Written in Bone (Form C)</th>
<th>A Christmas Carol (Form B)</th>
<th>Behind the Scenes (Form B)</th>
<th>Linking Passage #1 Magic</th>
<th>Linking Passage #2 Crossing</th>
<th>Linking Passage #3 TBD</th>
<th>Linking Passage #4 TBD</th>
<th>Linking Passage #5 TBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>LASID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Items ordered by AP Sequence (variable on test maps) within each unit assessment*  
*Items ordered by Selection Sequence within each linking passage*
# Inclusion Criteria

<table>
<thead>
<tr>
<th>Data Field</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Taken Flag</td>
<td>Equal to “Y”</td>
</tr>
<tr>
<td>Test Void Flag</td>
<td>Is Blank</td>
</tr>
<tr>
<td>Test Raw Total</td>
<td>$\geq 0$</td>
</tr>
<tr>
<td>ELA Supplement Taken Flag</td>
<td>Equal to “Y”</td>
</tr>
<tr>
<td>Complete</td>
<td>Equal to “Y”</td>
</tr>
<tr>
<td>Accountability Code</td>
<td>Not equal to 86</td>
</tr>
<tr>
<td>Braille Provided</td>
<td>Not Equal to “Y”</td>
</tr>
<tr>
<td>Response to Item 12</td>
<td>Not blank</td>
</tr>
<tr>
<td>LASID</td>
<td>No duplicates</td>
</tr>
<tr>
<td>Test Grade</td>
<td>Same for all Windows</td>
</tr>
</tbody>
</table>

15,683 out of 18,849
Number of Students by Pattern and Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Student Counts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>By Pattern</td>
<td>By Grade</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1206</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2992</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>705</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>3317</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1404</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1714</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2342</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2001</td>
<td></td>
</tr>
</tbody>
</table>
## Number of Students by Unit and Linking Sets

<table>
<thead>
<tr>
<th>Grade</th>
<th>Unit Test</th>
<th>Student Count</th>
<th>Link Set</th>
<th>Student Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>DUST-A</td>
<td>4198</td>
<td>Cranes</td>
<td>4903</td>
</tr>
<tr>
<td></td>
<td>HATCHET-A</td>
<td>4903</td>
<td>Woman</td>
<td>4903</td>
</tr>
<tr>
<td></td>
<td>JOBS-A</td>
<td>705</td>
<td>Underground</td>
<td>1629</td>
</tr>
<tr>
<td></td>
<td>STONES-A</td>
<td>4903</td>
<td>Happiness</td>
<td>1633</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Summit</td>
<td>1641</td>
</tr>
<tr>
<td>7</td>
<td>GIVER-C</td>
<td>6435</td>
<td>Crossing</td>
<td>3240</td>
</tr>
<tr>
<td></td>
<td>BONE-C</td>
<td>3120</td>
<td>Magic</td>
<td>3197</td>
</tr>
<tr>
<td></td>
<td>CC-B</td>
<td>3319</td>
<td>Bird</td>
<td>2142</td>
</tr>
<tr>
<td></td>
<td>SCENES-B</td>
<td>6437</td>
<td>Why</td>
<td>2160</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Home</td>
<td>2135</td>
</tr>
<tr>
<td>8</td>
<td>FLOWERS-A</td>
<td>4343</td>
<td>Chronicles</td>
<td>2165</td>
</tr>
<tr>
<td></td>
<td>COTW-A</td>
<td>2001</td>
<td>Science Pair</td>
<td>2178</td>
</tr>
<tr>
<td></td>
<td>TTH-A</td>
<td>2342</td>
<td>Genius</td>
<td>1083</td>
</tr>
<tr>
<td></td>
<td>SUGAR-A</td>
<td>4343</td>
<td>Archeology</td>
<td>1067</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alyssa</td>
<td>1077</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Courage</td>
<td>1116</td>
</tr>
</tbody>
</table>
Evaluation of Items for Use

1. Classical Test Theory Analyses of items in Unit Tests. Appendix A.
2. Classical Test Theory Analysis of items in Linking Sets. Appendix B.
3. IRT statistics for operational IAP items. Appendix C.
4. IRT fit plots for operational IAP items. Appendix D.
5. Stability analysis of items in Linking Sets. Appendix E.
# Item Actions/Decisions

<table>
<thead>
<tr>
<th>Grade</th>
<th>Unit</th>
<th>Item Reference</th>
<th>Action</th>
<th>Psychometrician Notes</th>
<th>Content Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>JOBS-A</td>
<td>LADOE_malw1oPQU7</td>
<td>Drop</td>
<td>Adj Corr = .103; poor fit; a-parm = .132; large standard errors on difficulty parameters</td>
<td>Part A is tricky: the connection between the quote and the answer is not obvious, but it's the only correct answer.</td>
</tr>
<tr>
<td>8</td>
<td>COTW-A</td>
<td>LADOE_oYXAgVET5e</td>
<td>Drop</td>
<td>Adj Corr = .087; poor fit; a-parm = .119; outlying difficulty parameters with large standard errors</td>
<td>C is a little attractive, but the keys--A and E--are the best answers.</td>
</tr>
<tr>
<td>8</td>
<td>SUGAR-A</td>
<td>MZ_L_432</td>
<td>Drop</td>
<td>Adj_corr = .195; a_parm=.142; poor fit plot;</td>
<td>This item is problematic. The keys are not clear, so we are fine with dropping it.</td>
</tr>
<tr>
<td>6</td>
<td>STONES-A</td>
<td>LADOE_3EPXbym1PG</td>
<td>Drop</td>
<td>This item was a definite &quot;Drop&quot; decision in Dry Run because a 0-point distractor (A:B) had an adj_corr of 0.184, while all 1-point answers had adj_corr &lt; 0. In OP Run the A:B adj_corr = .056, still greater than the adj_corr of any 1-point answer.</td>
<td>We are concerned that A in part A is too attractive since it's a big idea in the whole book and partly addressed in the quote. So, we are fine with dropping this item. If the stem had included &quot;Based on the excerpt,&quot; we would have been more inclined to keep it. (good candidate for revising and including in any future EFT slots)</td>
</tr>
</tbody>
</table>
## Dropped Linking Items

<table>
<thead>
<tr>
<th>Grade</th>
<th>Linking Set</th>
<th>Item Reference</th>
<th>Psychometric Notes</th>
<th>Content Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Happiness</td>
<td>LADOE_09Nrlv9LS</td>
<td>$Z_b &gt; 1.96$ and $WRMSD &gt; 0.075$</td>
<td>Fine to drop, not crucial to the set, but will leave only 5 items in this set moving forward</td>
</tr>
<tr>
<td>6</td>
<td>Summit</td>
<td>LADOE_7hk3QkIIdk</td>
<td>$Z_b &gt; 1.96$ and $WRMSD &gt; 0.075$</td>
<td>I think the set might be fine without this item. However, I am a little hesitant to remove it from the set next year since I think it provides an overview of the passage and may help students in answering other items in the set. It also addresses an earlier part of the passage, and the other items focus more on the second half of the passage, so it fills in a smaller gap.</td>
</tr>
<tr>
<td>8</td>
<td>Science</td>
<td>LADOE_5Zfd5vIDwV</td>
<td>$Z_b &gt; 1.96$ and $WRMSD &gt; 0.075$</td>
<td>For content reasons, we can’t remove this item from the set next year since it is the only crossover item in this pair. We don’t give students a pair without a crossover item because it doesn’t make sense to have a pair if there is not an item that asks students to connect the two related texts.</td>
</tr>
</tbody>
</table>
### Number of Items and Points in Link Sets Before and After Item Action Decisions

<table>
<thead>
<tr>
<th>Grade</th>
<th>Link Set</th>
<th>Before</th>
<th></th>
<th>After</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
</tr>
<tr>
<td>6</td>
<td>Cranes</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Woman</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Underground</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Happiness</td>
<td>6</td>
<td>12</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Summit</td>
<td>8</td>
<td>16</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>Crossing</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Magic</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Bird</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Why</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Home</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>Chronicles</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Science</td>
<td>6</td>
<td>12</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Genius</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Archeology</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Alyssa</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Courage</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Grade</td>
<td>Unit</td>
<td>Before</td>
<td></td>
<td>After</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>--------</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
</tr>
<tr>
<td>6</td>
<td>DUST-A</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>HATCHET-A</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>JOBS-A</td>
<td>13</td>
<td>33</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>STONES-A</td>
<td>13</td>
<td>33</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>GIVER-C</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>BONE-C</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>CC-B</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>SCENES-B</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>FLOWERS-A</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>COTW-A</td>
<td>13</td>
<td>33</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>TTH-A</td>
<td>13</td>
<td>33</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>SUGAR-A</td>
<td>13</td>
<td>33</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

Number of Items and Points per Form Before and After Item Evaluation
### Number of Points per Summary/Overall Assessment by Pattern

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Achievement Level</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Approaching Basic</td>
<td>Basic</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>19</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>24</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>23</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>24</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>24</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>24</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>25</td>
<td>38</td>
</tr>
</tbody>
</table>
### RSSS Table

*(Raw Score to Scale Score)*

- Pattern
- Category
  - Overall
  - Knowledge
  - Application
  - Synthesis
- Raw Score
- Theta
- Theta SEM
- Scale Score
- Scale Score
  - Scale Score SEM
  - Scale Score Upper/Lower Bounds
- Achievement Level
Grade 5 Field Test
IAP Key Findings and Considerations

Progressive Measurement
Operational Forms (Grades 6-8)
## Consistency of Cut Scores over Patterns

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Achievement Level</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Approaching Basic</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Basic</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>20</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>20</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>19</td>
<td>94</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>24</td>
<td>99</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>23</td>
<td>99</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>24</td>
<td>99</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>24</td>
<td>99</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>24</td>
<td>97</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>25</td>
<td>95</td>
</tr>
</tbody>
</table>

| | | Basic | Mastery | Advanced |
|-------|---------|--------|----------|
|       |         |        |          |
| 6     | 1       | 33     | 50       | 76       |
| 6     | 2       | 33     | 50       | 76       |
| 6     | 3       | 33     | 50       | 74       |
| 7     | 1       | 38     | 53       | 71       |
| 7     | 2       | 36     | 51       | 70       |
| 7     | 3       | 37     | 53       | 72       |
| 7     | 4       | 37     | 53       | 72       |
| 8     | 1       | 37     | 52       | 74       |
| 8     | 2       | 38     | 53       | 73       |
### Consistency of Cut Scores over Patterns

#### Knowledge

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Moderate</th>
<th>Strong</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>13</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>13</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>13</td>
<td>17</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>16</td>
<td>20</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>16</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>16</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>14</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>14</td>
<td>17</td>
<td>25</td>
</tr>
</tbody>
</table>

#### Application

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Moderate</th>
<th>Strong</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>16</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>15</td>
<td>22</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>16</td>
<td>22</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>14</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>16</td>
<td>22</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>16</td>
<td>22</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>15</td>
<td>21</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15</td>
<td>20</td>
<td>37</td>
</tr>
</tbody>
</table>

#### Synthesis

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Moderate</th>
<th>Strong</th>
<th>Maximum Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>16</td>
<td>33</td>
</tr>
</tbody>
</table>
Overall Scale Score Means

We don’t expect means to be the same across patterns within grade because students were not randomly assigned to pattern.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>N</th>
<th>Mean</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>1206</td>
<td>740.90</td>
<td>27.85</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2992</td>
<td>735.69</td>
<td>28.18</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>705</td>
<td>731.86</td>
<td>25.97</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>4903</td>
<td>736.42</td>
<td>27.93</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>3317</td>
<td>744.55</td>
<td>32.48</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1404</td>
<td>736.29</td>
<td>32.21</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1714</td>
<td>744.68</td>
<td>31.71</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>6437</td>
<td>742.77</td>
<td>32.39</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2342</td>
<td>735.78</td>
<td>32.90</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2001</td>
<td>740.58</td>
<td>32.41</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td>4343</td>
<td>737.99</td>
<td>32.76</td>
</tr>
</tbody>
</table>
Overall Achievement Level Percentages

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pattern</th>
<th>Percentages by Achievement Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>≥ 3</th>
<th>≥ 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td></td>
<td>9.2</td>
<td>17.2</td>
<td>34.7</td>
<td>36.2</td>
<td>2.7</td>
<td>73.6</td>
<td>38.9</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>11.7</td>
<td>23.8</td>
<td>32.7</td>
<td>29.0</td>
<td>2.8</td>
<td>64.5</td>
<td>31.8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>10.9</td>
<td>28.8</td>
<td>34.8</td>
<td>23.7</td>
<td>1.8</td>
<td>60.3</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td></td>
<td>11.0</td>
<td>22.9</td>
<td>33.5</td>
<td>30.0</td>
<td>2.7</td>
<td>66.2</td>
<td>32.7</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td></td>
<td>8.7</td>
<td>19.1</td>
<td>26.5</td>
<td>34.3</td>
<td>11.3</td>
<td>72.1</td>
<td>45.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>12.9</td>
<td>23.8</td>
<td>28.2</td>
<td>28.2</td>
<td>6.9</td>
<td>63.3</td>
<td>35.1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td>7.9</td>
<td>19.0</td>
<td>27.6</td>
<td>34.7</td>
<td>10.9</td>
<td>73.2</td>
<td>45.6</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td></td>
<td>9.4</td>
<td>20.1</td>
<td>27.2</td>
<td>33.1</td>
<td>10.2</td>
<td>70.5</td>
<td>43.3</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td></td>
<td>14.6</td>
<td>21.0</td>
<td>31.0</td>
<td>29.5</td>
<td>3.9</td>
<td>64.4</td>
<td>33.4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>12.2</td>
<td>17.3</td>
<td>29.9</td>
<td>36.1</td>
<td>4.4</td>
<td>70.4</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>total</td>
<td></td>
<td>13.5</td>
<td>19.3</td>
<td>30.5</td>
<td>32.5</td>
<td>4.1</td>
<td>67.1</td>
<td>36.6</td>
</tr>
</tbody>
</table>

We don’t expect achievement level percentages to be the same across patterns within grade because students were not randomly assigned to pattern.
### Reporting Category Means and Level Percents

#### Knowledge

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>Mean</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>4903</td>
<td>737.81</td>
<td>37.0</td>
<td>27.4</td>
<td>35.6</td>
</tr>
<tr>
<td>7</td>
<td>6437</td>
<td>744.87</td>
<td>33.4</td>
<td>22.8</td>
<td>43.8</td>
</tr>
<tr>
<td>8</td>
<td>4343</td>
<td>740.04</td>
<td>37.0</td>
<td>23.5</td>
<td>39.5</td>
</tr>
</tbody>
</table>

#### Application

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>Mean</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>4903</td>
<td>736.55</td>
<td>36.7</td>
<td>33.1</td>
<td>30.2</td>
</tr>
<tr>
<td>7</td>
<td>6437</td>
<td>742.95</td>
<td>32.1</td>
<td>26.1</td>
<td>41.8</td>
</tr>
<tr>
<td>8</td>
<td>4343</td>
<td>737.88</td>
<td>36.0</td>
<td>26.8</td>
<td>37.2</td>
</tr>
</tbody>
</table>

#### Synthesis

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>Mean</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>4903</td>
<td>730.91</td>
<td>37.4</td>
<td>27.6</td>
<td>35.0</td>
</tr>
<tr>
<td>7</td>
<td>6437</td>
<td>739.83</td>
<td>29.0</td>
<td>27.3</td>
<td>43.7</td>
</tr>
<tr>
<td>8</td>
<td>4343</td>
<td>736.57</td>
<td>36.4</td>
<td>24.0</td>
<td>39.6</td>
</tr>
</tbody>
</table>
## Skill Profiles

### Grade 6

<table>
<thead>
<tr>
<th>Reporting Category</th>
<th>Mean</th>
<th>Percentages by Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>737.81</td>
<td>37.0 27.4 35.6</td>
</tr>
<tr>
<td>Application</td>
<td>736.55</td>
<td>36.7 33.1 30.2</td>
</tr>
<tr>
<td>Synthesis</td>
<td>730.91</td>
<td>37.4 27.6 35.0</td>
</tr>
</tbody>
</table>

### Grade 7

<table>
<thead>
<tr>
<th>Reporting Category</th>
<th>Mean</th>
<th>Percentages by Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>744.87</td>
<td>33.4 22.8 43.8</td>
</tr>
<tr>
<td>Application</td>
<td>742.95</td>
<td>32.1 26.1 41.8</td>
</tr>
<tr>
<td>Synthesis</td>
<td>739.83</td>
<td>29.0 27.3 43.7</td>
</tr>
</tbody>
</table>

### Grade 8

<table>
<thead>
<tr>
<th>Reporting Category</th>
<th>Mean</th>
<th>Percentages by Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>740.04</td>
<td>37.0 23.5 39.5</td>
</tr>
<tr>
<td>Application</td>
<td>737.88</td>
<td>36.0 26.8 37.2</td>
</tr>
<tr>
<td>Synthesis</td>
<td>736.57</td>
<td>36.4 24.0 39.6</td>
</tr>
</tbody>
</table>
Grade 5 Field Test
Comparability
Comparability

Propensity Score Matched-Groups

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>LEAP</th>
<th>IAP</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>4703</td>
<td>739.1</td>
<td>736.5</td>
<td>-3</td>
</tr>
<tr>
<td>7</td>
<td>6159</td>
<td>743.0</td>
<td>743.0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>3893</td>
<td>743.1</td>
<td>738.7</td>
<td>-4</td>
</tr>
</tbody>
</table>
Students who took both IAP and LEAP in 2023

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>LEAP</th>
<th>IAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
<td>703.4</td>
<td>723.0</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>749.0</td>
<td>764.0</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>732.5</td>
<td>723.0</td>
</tr>
<tr>
<td>All</td>
<td>27</td>
<td>725.6</td>
<td>726.0</td>
</tr>
</tbody>
</table>
Use and further development of Pre-equated Forms
## Types of IAP Forms in 2023-24 School Year

<table>
<thead>
<tr>
<th>Form Contents</th>
<th>Number of Items</th>
<th>Form Names in '23-24</th>
<th>Reporting Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OP</td>
<td></td>
<td>Time</td>
</tr>
<tr>
<td></td>
<td>Calibrated</td>
<td>New</td>
<td>EFT</td>
</tr>
<tr>
<td>New (Post-equating)</td>
<td>0</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td><strong>Calibrated</strong></td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calibrated and Embedded</td>
<td>13</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Calibrated</strong> plus New OP</td>
<td>12</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Calibrated</strong> plus New OP and embedded</td>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
## Options

<table>
<thead>
<tr>
<th>Phase</th>
<th>Requirements</th>
<th>NEW</th>
<th>Calibrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intact</td>
</tr>
<tr>
<td>Preparation</td>
<td>OP Test Map</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>LEAP Link Item parameters</td>
<td>2</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>LEAP Link Test Map</td>
<td>3</td>
<td>?</td>
</tr>
<tr>
<td>Evaluation</td>
<td>IDM</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>IAP Item CTT Analyses</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>LEAP Link CTT</td>
<td>6</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>IAP Item IRT &amp; Fit</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Link Stability</td>
<td>8</td>
<td>?</td>
</tr>
<tr>
<td>Scaling</td>
<td>Item Actions</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Equating/Scaling</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>RSSS Tables</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Impact</td>
<td>Student Scores</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Impact Tables</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>
Potential TAC Topics

• Calibrating LC and RCWE items separately vs. jointly
• Use of Item/total vs. Item/Theta correlations
Do we need to calibrate LC and RCWE Items Separately?

• What is the effect of separate vs. joint calibrations on key outcomes?
  ➢ Item calibrations
  ➢ Standard errors
  ➢ What else?

• Joint calibration will save time and reduce chances of error

• Using equating constants that differ from expected values (0, and 1) only by random error bakes this additional source of random error into the calibrations of the LC items.
Equating Constants for LC items in 2023

<table>
<thead>
<tr>
<th>Grade</th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1.001651</td>
<td>-0.001298</td>
</tr>
<tr>
<td>7</td>
<td>0.997382</td>
<td>-0.000253</td>
</tr>
<tr>
<td>8</td>
<td>0.995562</td>
<td>-0.000555</td>
</tr>
</tbody>
</table>
Item/total vs Item/theta correlations

- Item/total correlations are currently based on end of year total raw scores
  - Raw scores are not comparable across “patterns”, so “total score” must be based on the test form in which the item was used
    - Forms have relatively few items and points (13 and 33 respectively)
      - Less reliable.
      - With fewer items/points, difference between adjusted and unadjusted item/total correlations get larger.
- “Item/theta” correlations based on end-of-year theta
  - End-of-year student measures (thetas) are estimated from item calibrations
    - The use of item calibrations to estimate thetas makes thetas comparable across patterns
    - Theta is therefore based on *ALL* items taken by a student (39 items and 99 points)
      - More reliable
      - Difference between unadjusted and adjusted correlation may be trivial enough to ignore the adjusted correlation
  - Theta is a “linear” (unbounded) measure, so it is compatible with the use of a correlation coefficient. (Correlation is a measure of linear association.)
## Item/total vs Item/theta correlations

<table>
<thead>
<tr>
<th>FormName</th>
<th>ItemReference</th>
<th>N</th>
<th>Correlation</th>
<th>Adj_Correlation</th>
<th>Theta</th>
</tr>
</thead>
<tbody>
<tr>
<td>HATCHET-A</td>
<td>MZ_L_108_RCWE</td>
<td>4903</td>
<td>0.74</td>
<td>0.65</td>
<td>0.84</td>
</tr>
<tr>
<td>HATCHET-A</td>
<td>MZ_L_108_LC</td>
<td>4903</td>
<td>0.73</td>
<td>0.65</td>
<td>0.83</td>
</tr>
<tr>
<td>JOBS-A</td>
<td>MZ_L_130</td>
<td>705</td>
<td>0.41</td>
<td>0.35</td>
<td>0.62</td>
</tr>
<tr>
<td>JOBS-A</td>
<td>MZ_L_119</td>
<td>705</td>
<td>0.61</td>
<td>0.48</td>
<td>0.67</td>
</tr>
<tr>
<td>JOBS-A</td>
<td>MZ_L_118</td>
<td>705</td>
<td>0.57</td>
<td>0.47</td>
<td>0.55</td>
</tr>
<tr>
<td>JOBS-A</td>
<td>LADOE_xiNkJspowz</td>
<td>705</td>
<td>0.32</td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td>JOBS-A</td>
<td>MZ_L_126</td>
<td>705</td>
<td>0.42</td>
<td>0.32</td>
<td>0.37</td>
</tr>
<tr>
<td>STONES-A</td>
<td>LADOE_GAGHr6HCDn</td>
<td>4903</td>
<td>0.51</td>
<td>0.43</td>
<td>0.50</td>
</tr>
<tr>
<td>STONES-A</td>
<td>MZ_L_356</td>
<td>4903</td>
<td>0.46</td>
<td>0.36</td>
<td>0.44</td>
</tr>
<tr>
<td>STONES-A</td>
<td>LADOE_xiNkJspowz</td>
<td>4903</td>
<td>0.76</td>
<td>0.66</td>
<td>0.79</td>
</tr>
<tr>
<td>STONES-A</td>
<td>LADOE_h84VgJfa9i_RCWE</td>
<td>4903</td>
<td>0.76</td>
<td>0.69</td>
<td>0.83</td>
</tr>
<tr>
<td>STONES-A</td>
<td>LADOE_h84VgJfa9i_LC</td>
<td>4903</td>
<td>0.78</td>
<td>0.69</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Dropped: JOBS-A LADOE_xiNkJspowz
Topic 5: Enhanced Reporting and Reporting Engagement Plan

August 8, 2023
Agenda

1. Development process in Spring 2023, lessons learned and current work products
2. Next steps and considerations
Part 1.
Development process in Spring 2023, lessons learned and current work products
Window 1
Students take one of the two unit assessments.
Dust
Hatchet
End-of-Unit Score Report

Window 2
Students take one of the three unit assessments.
Dust
Hatchet
Steve Jobs
End-of-Unit Score Report

Window 3
Students take one unit assessment.
Stones
End-of-Year Score Report

LEAP Passage #1
LEAP Passage #2
LEAP Passage #3
LEAP Passage #4
LEAP Passage #5
LEAP Passage #6

Fall  Spring
Presentation Purpose: Detail our approach to developing **score reports and supports** that are meant to be instructionally useful.
Score Reports and Supports

Our work during the Spring of 2023 was focused on both building score reports and also supports that help connect to instructional next steps. The supports, in part, help acknowledge that any score report can, at best, play a supporting role in instructional decision making.

These supports should be robust & extensive
Score Reports and Supports

Our work during the **Spring of 2023** was focused on both building *score reports* and also *supports* that help connect to instructional next steps. The supports, in part, help acknowledge that any score report can, at best, play a supporting role in instructional decision making.

---

**Classroom Report for Teachers**

---

**Classroom Supports**

---

### BAP Score Report Guidance

For students scoring in the moderate or weak score ranges within each domain, use the reflection questions to help design follow-up supports. Note that the domains and each set of reflection questions are ordered intentionally. Weaknesses in multiple domains should be addressed in order of 1) Knowledge, 2) Application, and 3) Synthesis. Additionally, reflection questions should be asked and addressed in the order provided within each domain. For example, if a student scores weak in Knowledge, you must first assess whether or not they have gaps in foundational reading skills before working on building capacity for answering higher-order questions about grade-level texts.

1. **Knowledge**
   1. Do the students have gaps in foundational reading skills or struggle to read the grade level material?
   2. Do the students have gaps in vocabulary or word meaning?
   3. Do the students have gaps in their knowledge about the unit topic?
   4. Do the students have gaps in their knowledge of grade-level texts?

2. **Application**
   1. Do the students have trouble understanding the central ideas, meaning of words and phrases, and main ideas presented in a text?
   2. Do the students have trouble applying their learning to independent tasks?

3. **Synthesis**
   1. Do the students have trouble at the sentence level of writing?
   2. Do the students have trouble understanding the expectations of writing tasks?
Approach

- Develop rough mockups of the reports and supports using a combination of Rmarkdown and Google docs
  - An important intermediate step
- Reports showed student data from each teacher’s classrooms from that year.
  - Hopefully to provide important context
Intended Use for Classroom Reports for Teachers

To support teachers in adjusting and differentiating instruction for small groups of students within the upcoming curricular unit, based on triangulation across the IAP assessment results, curriculum based assessments and classroom interactions.
Elaborating the Use Case

To support teachers in adjusting and differentiating instruction for small groups of students within the upcoming curricular unit, based on triangulation across the IAP assessment results, curriculum based assessments and classroom interactions.

- Students groupings are based on the IAP assessment reporting categories of Knowledge, Application and Synthesis. Each of these have three levels.
- The supports provide reflection questions that are meant to help teachers think holistically about students and strategies for instructional next steps tied to the curriculum.

We are trying to further systemize good instructional practice in ways that enhance teacher capacity.
So let’s dive into the rough prototypes shared with the field in early Spring 2023.

Remember, these rough mockups play an important role in gathering feedback prior to cleaned up versions.
Application

In this reporting category, students read a new text related to the unit content and respond to multiple questions as well as four point constructed response question that measures their ability to apply the key knowledge and skills taught in the unit. The maximum number of points was 13 in this reporting category.

<table>
<thead>
<tr>
<th>Application Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1</strong></td>
</tr>
<tr>
<td>STU T (0)</td>
</tr>
<tr>
<td>STU E (2)</td>
</tr>
<tr>
<td>STU F (2)</td>
</tr>
<tr>
<td>STU G (2)</td>
</tr>
<tr>
<td>STU X (2)</td>
</tr>
<tr>
<td>STU C (3)</td>
</tr>
<tr>
<td>STU W (3)</td>
</tr>
<tr>
<td>STU L (4)</td>
</tr>
<tr>
<td>STU R (4)</td>
</tr>
</tbody>
</table>

For the 14 students in the Level 1 or 2 ratings, consider the following reflection questions:

1. Do the students struggle with certain aspects of text complexity (central ideas, meaning of words and phrases, text structure, background knowledge needed for comprehension)?
2. Do the students have trouble applying their learning to independent tasks?

If you answered "yes" to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all reporting categories.
<table>
<thead>
<tr>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU T (0)</td>
<td>STU P (5)</td>
<td>STU B (7)</td>
</tr>
<tr>
<td>STU E (2)</td>
<td>STU D (6)</td>
<td>STU Q (7)</td>
</tr>
<tr>
<td>STU F (2)</td>
<td>STU I (6)</td>
<td>STU Z (7)</td>
</tr>
</tbody>
</table>

1. Do the students struggle with certain aspects of text complexity (central ideas, meaning of words and phrases, text structure, background knowledge needed for comprehension)?
2. Do the students have trouble applying their learning to independent tasks?

If you answered "yes" to any of these questions, consider the example next steps in [this supporting document](#). This guidance is also available in a more [complete document for all reporting categories](#).
Stakeholder Engagement

We were able to draw on number of ongoing relationships to gather feedback on a draft classroom report and supports.

- Creation and review of classroom score report mockups using real data for four focus groups in early Spring 2023, after window 2
  - Fairly representative groups of 6-9 teachers and leaders
  - Participants part of “IAP Collaboration Groups”
- Ad hoc meetings with:
  - IAP Collaboration Group Leads
  - Parent and Family Engagement Coordinators
Reporting is as much a school and district process as it is a state process.

- Much of the interpretive and training materials are district created and district disseminated.
- There are a number of key touch points throughout the year that materials can be developed to support.
- Many districts would like to consume and report results in their own Learning Management Systems.
Educators emphasized that they used prior reports:
- As part of coaching conversations with students, especially around goal setting.
- To provide intervention and plan instruction.

Teachers wanted a variety of improvements to make the materials more accessible and usable.
- Essay rubrics translated into student friendly language
- More exemplars of the essays.
- Additional detail on essay scores of zero.

Reports are used as a coaching tool as well as planning tool.
Teachers thought the classroom reports and supports had promise.

- The classroom report mockup spanned multiple pages (3-4) and contained text that was meant to help explain the report in friendly terms.
  - Participants responded favorably to these extended reports.
  - Suggests potential for more text based reporting
- They thought the supports document provided useful information in considering next steps
- Additional requests:
  - Comparisons
  - Finer grained information
Subsequent Work and Next Steps

● Partnerships with curriculum experts with expertise in the Guidebooks and Wit & Wisdom curriculum

● Current working drafts
  ○ Classroom score report for teachers (minor revisions pending, LDOE review, development of final template)
  ○ Teacher guide document (minor revisions pending, LDOE review)
  ○ Professional development materials (minor revisions pending, LDOE review)
  ○ Student score report for students and parents (minor revisions pending, LDOE review, development of final template)
  ○ Plus a request for revised .csv files that can be consumed within district learning management systems (including student demographics)
Part 2.
Next steps and considerations
Open Questions

1. All of the materials are moving towards finalization. What are the most important pieces to finalize ASAP?
   a. We are assuming it is the first PD module and the classroom report, so that those materials can be disseminated.

2. Does LDOE want essay scores be provided on the classroom score reports?
   a. Comparisons of rubric score point curves could be compared to see how different the difficulties are.

3. How does LDOE want to share and roll out the new classroom level report & professional development materials, once finalized?
   a. How will the materials be shared? Will there be a orientation webinar? Will LDOE want to meet with district leads to provide training?
   b. Where will the materials be housed?

4. Is LDOE open to providing a link to a feedback form within the score reports or teacher guide?
Considering SY 23-24

- The current scope supports the development of one new report. A key report is an update to the classroom report for Window 2.
  - Particularly important is considering how to communicate about change across windows. For example, the updated score reports could provide transition tables between levels for each reporting category.
  - Two focus groups could be allotted to exploring this design with the IAP collaboration groups.
- A survey or other work could focus more systematically on whether and how instruction changes as a result of score report use
  - A special study could be conducted to expand on this, e.g., to see if some schools are doing much better than compared to their peers.
- Work on this next score report should begin in late August immediately after we finalize the current student score report
- Other areas of focus?
  - An examination of the potential for AI scoring of essays?
Topic 6: Planning Needs for 23-24 and Beyond

August 8, 2023
Agenda

1. Features of the IAP and Future Directions
2. Considering Timelines and “Scale Up”
Part 1.
Features of the IAP and Future Directions
Some Features of the IAP

- Administered in three End-of-Unit/Window assessments
- Choice around which units are administered within the Guidebooks
- Heavy emphasis on writing
- Summative scores produced by weighting each window
- Directly connected to curriculum with a focus on unit texts and the knowledge built from engaging with them
Framing Future Directions

The key questions here are:

1. What combination of features is needed to support LDOE’s future assessment work?
2. What combination of features is sustainable?
3. How can these features be tested out in the crawfish model or otherwise?
Pressing Challenges

- **Curricula Identification.** Designing assessment content to curricula means that we need to identify the curricula to be developed to
  - To date, voluntary collection of curricula information has not resulted a clear picture of statewide implementation
  - Is there some target pool of curricula that needs to be developed to?

- **Size of Item and Form Bank.** What are the total number of forms, and associated items, that are needed to support the design?

- **Choice.** Designing enough choice in assessment options to have all stakeholders see this model as an improvement
<table>
<thead>
<tr>
<th>Option</th>
<th>Main IAP</th>
<th>Crawfish Spring 23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three End-of-Unit/Window assessments</td>
<td>✔️</td>
<td>✔️ (Only final window tested)</td>
</tr>
<tr>
<td>Unit Choice</td>
<td>✔️</td>
<td>TBD</td>
</tr>
<tr>
<td>Heavy emphasis on writing</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Summative scores produced by weighting each window</td>
<td>✔️</td>
<td>TBD</td>
</tr>
<tr>
<td>Directly connected to curriculum w/ focus on unit texts</td>
<td>✔️</td>
<td>Knowledge Directly Connected; Application &amp; Synthesis Common across two curricula</td>
</tr>
</tbody>
</table>
Visualizing Curricula: Main IAP, Grades 6-8

Window 1
Guidebooks
...

Window 2
Guidebooks
...

Window 3
Guidebooks
...

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spring</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Visualizing Curricula: Main IAP, Grade 5

Window 1

Guidebooks

Wit & Wisdom

Window 2

Guidebooks

Wit & Wisdom

Window 3

Guidebooks

Wit & Wisdom

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spring
Visualizing Curricula: Crawfish, Spring 23

Window 1
- Guidebooks
- Wit & Wisdom
- EL Education

Window 2
- Guidebooks
- Wit & Wisdom
- EL Education

Window 3
- Guidebooks
- Wit & Wisdom
- EL Education

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fall
Spring
## Considering Options

<table>
<thead>
<tr>
<th></th>
<th>Main IAP</th>
<th>Crawfish Spring 23</th>
<th>Crawfish Spring 24</th>
<th>Other variations?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three End-of-Unit/Window assessments</td>
<td>✔️</td>
<td>✔️ (Only final window tested)</td>
<td>✔️ (Only final window tested)</td>
<td></td>
</tr>
<tr>
<td>Unit Choice</td>
<td>✔️</td>
<td>TBD</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Heavy emphasis on writing</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Summative scores produced by weighting each window</td>
<td>✔️</td>
<td>TBD</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Directly connected to curriculum w/ focus on unit texts</td>
<td>✔️</td>
<td>Knowledge Directly Connected; Application &amp; Synthesis Common across two curricula</td>
<td>TBD</td>
<td></td>
</tr>
</tbody>
</table>
Now let’s try explore some possible designs.

We share these designs to illustrate what *might* be possible. These designs intentionally stretch our current designs, and do so for the purposes of illustration.
Considering Possibilities: Summary

A. Through-Year Crawfish forms, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

B. Guidebooks with Through-Year and End-of-Year Versions, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

C. Through-Year versions for most popular curricula within the state (say more than 15% of students taking the curricula), with a generic End-of-Year curricula agnostic version
Considering Possibilities: Summary

A. Through-Year Crawfish forms, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

B. Guidebooks with Through-Year and End-of-Year Versions, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

C. Through-Year versions for most popular curricula within the state (say more than 15% of students taking the curricula), with a generic End-of-Year curricula agnostic version
Considering Possibilities: Summary

A. Through-Year Crawfish forms, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

B. Guidebooks with Through-Year and End-of-Year Versions, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

C. Through-Year versions for most popular curricula within the state (say more than 15% of students taking the curricula), with a generic End-of-Year curricula agnostic version
Guidelines for Visualizing Curricula: Guidebooks with Through-Year and EOY

Window 1
- Guidebooks...

Window 2
- Guidebooks...

Window 3
- Guidebooks
- Guidebooks End of Year Form
- Curricula Agnostic End of Year Form

Timeline:
- Fall Spring
Considering Possibilities: Summary

A. Through-Year Crawfish forms, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

B. Guidebooks with Through-Year and End-of-Year Versions, plus a generic End-of-Year curricula agnostic version (potentially pulling on LEAP content)

C. Through-Year versions for most popular curricula within the state (say more than 15% of students taking the curricula), with a generic End-of-Year curricula agnostic version
Visualizing Curricula: Unit Forms for Most Popular Curricula

Window 1
- Guidebooks
- Wit & Wisdom
- EL Education

... 

Window 2
- Guidebooks
- Wit & Wisdom
- EL Education

... 

Window 3
- Guidebooks
- Wit & Wisdom
- EL Education

Curricula Agnostic End of Year Form

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spring</td>
</tr>
</tbody>
</table>
Next Steps

● What dimensions of design are most important to LDOE?
  ○ Three End-of-Unit/Window assessments
  ○ Unit Choice
  ○ Heavy emphasis on writing
  ○ Summative scores produced by weighting each window
  ○ Directly connected to curriculum w/ focus on unit texts

● Which, if any, of the design options should be considered and explored? Or what new designs need to be invented?

● Are there specific special studies of interest to LDOE here? E.g.,
  ○ Alternative approaches to the creation of summative scores?
  ○ Exploration of change across windows?
Part 2.
Considering Timelines and “Scale Up”
## Multi-year Overview

<table>
<thead>
<tr>
<th></th>
<th>SY 21-22</th>
<th>SY 22-23</th>
<th>SY 23-24</th>
<th>SY 24-25</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade 7</strong></td>
<td>5,000</td>
<td>10,000</td>
<td>12,000</td>
<td>12,000 +</td>
</tr>
<tr>
<td><strong>Grades 6 and 8</strong></td>
<td>Field Test (and Grade 7 replenishment)</td>
<td>10,000 (6 and 8)</td>
<td>12,000</td>
<td>12,000 +</td>
</tr>
<tr>
<td><strong>Grade 5 – ELA Guidebooks</strong></td>
<td>Pilot</td>
<td>Field Test (5,000)</td>
<td>6,000</td>
<td>6,000 +</td>
</tr>
<tr>
<td><strong>Grade 5 – Wit and Wisdom</strong></td>
<td>Pilot</td>
<td>Field Test (5,000)</td>
<td>2,000</td>
<td>TBD</td>
</tr>
<tr>
<td><strong>Grades 3 and 4</strong></td>
<td>Development</td>
<td>Development</td>
<td>Field Test</td>
<td>TBD</td>
</tr>
<tr>
<td><strong>Third Option – Grade 5?</strong></td>
<td>Research/ Stakeholder Engagement</td>
<td>Development/Pilot (Spring 2023)</td>
<td>Full Pilot?</td>
<td>TBD</td>
</tr>
</tbody>
</table>

### Operational
- Field Test or Pilot
- Research and Development

*Last year of current waiver*
## Multi-year Overview

<table>
<thead>
<tr>
<th></th>
<th>SY 21-22</th>
<th>SY 22-23</th>
<th>SY 23-24</th>
<th>SY 24-25</th>
<th>SY 25-26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 7</td>
<td>5,000</td>
<td>10,000</td>
<td>12,000</td>
<td>12,000 +</td>
<td></td>
</tr>
<tr>
<td>Grades 6 and 8</td>
<td>Field Test (and Grade 7 replenishment)</td>
<td>10,000 (6 and 8)</td>
<td>12,000</td>
<td>12,000 +</td>
<td></td>
</tr>
<tr>
<td>Grade 5 – ELA Guidebooks</td>
<td>Pilot</td>
<td>Field Test (5,000)</td>
<td>6,000</td>
<td>6,000 +</td>
<td></td>
</tr>
<tr>
<td>Grade 5 – Wit and Wisdom</td>
<td>Pilot</td>
<td>Field Test (5,000)</td>
<td>2,000</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>Grades 3 and 4</td>
<td>Development</td>
<td>Development</td>
<td>Field Test</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>Third Option – Grade 5?</td>
<td>Research/ Stakeholder Engagement</td>
<td>Development/Pilot (Spring 2023)</td>
<td>Full Pilot?</td>
<td>TBD</td>
<td></td>
</tr>
</tbody>
</table>

**Last year of current waiver**

<table>
<thead>
<tr>
<th>Operational</th>
<th>Field Test or Pilot</th>
<th>Research and Development</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TBD</strong> (Statewide Implementation, Waiver Extension, Program Adaptation Program Sunset)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Considering Scale Up

- What grades and curricula should be covered by the IAP by 24-25?
- How many students should be assessed each year?
  - A steady increase to familiarize the field and pressure test the program
  - Intentionally limited to a representative set of schools and districts
- Need for systematic data collection on curricula
  - DTC managed survey as part of this test administration process? If so, done just within IAP or statewide as part of LEAP?
- For the future:
  - Impact of curriculum changes and updates?
  - Linking across forms
  - Reporting scale
Considering Scale Up: Numbers of Students

Possible Grade 7 Trajectories

<table>
<thead>
<tr>
<th></th>
<th>SY 21-22</th>
<th>SY 22-23</th>
<th>SY 23-24</th>
<th>SY 24-25</th>
<th>SY 25-26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series1</td>
<td>4,844</td>
<td>4,903</td>
<td>9,000</td>
<td>14,000</td>
<td>40,000</td>
</tr>
<tr>
<td>Series2</td>
<td>4,844</td>
<td>4,903</td>
<td>9,000</td>
<td>20,000</td>
<td>40,000</td>
</tr>
<tr>
<td>Series3</td>
<td>4,844</td>
<td>4,903</td>
<td>9,000</td>
<td>9,000</td>
<td>40,000</td>
</tr>
<tr>
<td>Series4</td>
<td>4,844</td>
<td>4,903</td>
<td>9,000</td>
<td>9,000</td>
<td>9,000</td>
</tr>
</tbody>
</table>
Research/Psychometric Work Plan

Progressive Measurement

and

The Center
Essential Work Segments & Keystones

• Psychometric Scaling Specifications

• Grades 3 and 4 field tests
  • OP Form Construction
    • Timing analyses (each window?)
    • Data Review in July ‘24
      • EOY CTT, IRT, and DIF analyses

• Grade 5
  • ALD alignment workshop
    • W1W2 Postequating using LEAP items
    • W1W2 RSSS Tables
  • EOY Reporting for Accountability
    • W1W2W3 Postequating using LEAP items
    • Equating Summary
    • RSSS Tables

• Grades 6 to 8
  • RSSS tables and scale score reports each window using OP items calibrated in 22-23
  • Calibration of embedded items and OP items that replace OP items dropped in 22-23
  • EOY RSSS tables
  • Post-equating using LEAP linking sets?
Scaling Specifications

• NWEA sends draft specifications 8/23
• Topic at team meeting 8/24
• LDOE provides feedback 8/31
• NWEA submits revised draft 9/6
• LDOE provides feedback 9/13
• NWEA provides final draft 9/18
Grades 3 and 4 Field Tests

• Each Window
  • Test Map Exchange
    • LDOE Test Maps
    • Research Test Map
  • Timing Data Analysis
• W1W2W3 (EOY)
  • CTT, IRT, and DIF analyses and reports
  • Data Review
LEAP Post-Equating (IAP OPS and EFTs)

Required for Grade 5; Optional for Grades 6-8

- Test Map Exchange
  - From LDOE to Research
  - From NWEA to LDOE

- Item Parameters
  - LEAP-format From LDOE
  - Nominal-format from NWEA

- Test Information Functions

- Post-equating to LEAP
  - CTT analysis
  - Stability Analysis
  - Equating Analysis, Summary, RSSS Tables

Can be done at EOY

Should be done at EOY, using W1W2W3 data
Equating (EFTs and New Ops) through OP IAP
Items Calibrated in ‘22-23

- Stability Analysis
- Equating, Equating 44
  rrr[r Summary, RSSS Tables
- Equating Summary
- RSSS Tables

Should be done at EOY, using W1W2W3 data
Grade 5 Window 1
Grade 5-8, Window 1

• Test Map Exchange (LDOE and Research) 7/5/23
• Test Information Functions 7/18/23
• CTT Analysis (OP and Linking)
Grade 5

End of Window 2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Sequence</th>
<th>Activities/Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation</td>
<td>1</td>
<td>W1 and W2 OP Test Maps</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>LEAP Link Item parameters</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>LEAP Link Test Maps</td>
</tr>
<tr>
<td>Evaluation</td>
<td>4</td>
<td>IDM</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>IAP Item CTT Analyses</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>LEAP Link CTT</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>IAP Item IRT &amp; Fit</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Link Stability</td>
</tr>
<tr>
<td>Scaling</td>
<td>9</td>
<td>Item Actions</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Equating/Scaling</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>RSSS Tables</td>
</tr>
<tr>
<td>Impact</td>
<td>12</td>
<td>Student Scores</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Impact Tables</td>
</tr>
</tbody>
</table>
Grade 5 ALD Alignment

• Item calibrations
• Participants (recruitment)
• Participant Materials
• Slides & other meeting prep
• Meeting
• Report
• ALDs
Grade 5 End of Year

• Repeat end of Window 2 analyses with Window 3 added on
• Equating summary
Grades 6, 7, 8 Precalibrated and Embedded

- Pre-Administration (Each Window)
  - Create RSSS table for each form to be administered using only the pre-calibrated operational items on the forms
- Post-Administration
# Types of IAP Forms in 2023-24 School Year

<table>
<thead>
<tr>
<th>Form Contents</th>
<th>OP Calibrated</th>
<th>OP New</th>
<th>EFT</th>
<th>Form Names in '23-24</th>
<th>Reporting Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>New (Post-equating)</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>All Grade 5 forms</td>
<td>4 weeks</td>
</tr>
<tr>
<td><strong>Calibrated</strong></td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>DUST-A</td>
<td>No delay</td>
</tr>
<tr>
<td><strong>Calibrated</strong> and Embedded</td>
<td>13</td>
<td>0</td>
<td>2</td>
<td>DUST-A</td>
<td>No delay</td>
</tr>
<tr>
<td><strong>Calibrated</strong> plus New OP</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>JOBS-B</td>
<td>No delay</td>
</tr>
<tr>
<td><strong>Calibrated</strong> plus New OP and embedded</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>JOBS-BE</td>
<td>No delay</td>
</tr>
</tbody>
</table>
## Options

<table>
<thead>
<tr>
<th>Phase</th>
<th>Requirements</th>
<th>NEW</th>
<th>Calibrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intact</td>
</tr>
<tr>
<td>Preparation</td>
<td>OP Test Map</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>LEAP Link Item parameters</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>LEAP Link Test Map</td>
<td>3</td>
<td>?</td>
</tr>
<tr>
<td>Evaluation</td>
<td>IDM</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>IAP Item CTT Analyses</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>LEAP Link CTT</td>
<td>6</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>IAP Item IRT &amp; Fit</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Link Stability</td>
<td>8</td>
<td>?</td>
</tr>
<tr>
<td>Scaling</td>
<td>Item Actions</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Equating/Scaling</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>RSSS Tables</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Impact</td>
<td>Student Scores</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Impact Tables</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>
# Grades 6 to 8 Operational Test Administration Plan

<table>
<thead>
<tr>
<th>Grade</th>
<th>Unit-Form</th>
<th>Administration Window</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>DUST-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HATCHET-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JOBS-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STONES-A</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>GIVER-C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BONE-C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC-B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCENES-B</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>FLOWERS-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COTW-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTH-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUGAR-A</td>
<td></td>
</tr>
</tbody>
</table>
Introduction
Since 2018, the state of Louisiana, along with its partners (Center for Assessment, Johns Hopkins University, Odell Education, MZ Development, Strategic Measurement and Evaluation), has invested in the creation of a test design that is unlike any other state assessment in the nation. Louisiana employs a high-quality curriculum-integrated through-year approach consisting of several assessments administered throughout the school year, so educators and students receive ongoing feedback to inform teaching and learning.

This Innovative Assessment is explicitly aligned to the English language arts (ELA) curriculum content that the students have studied. It measures both content knowledge and skills students derive from studying specific texts and topics. Of the few states that participated in the Innovative Assessment Demonstration Authority (IADA), Louisiana has been cited as a state that was the most explicit in their attempt to use IADA as an instrument to improve educational equity.

Specifically, the purpose of the Innovative Assessment is to improve the following:
- Integration: Assessment design strengthens the connection between instruction and assessment by using several brief curriculum-connected assessments throughout the year, which measure students’ comprehension of materials and texts that they have studied in class.
- Focus: Teachers can focus instruction on background knowledge and making meaning of full texts.
- Equal Access: All students have the opportunity to develop background knowledge together so that no student is at a disadvantage due to a lack of life experiences.
- Preservation of Local Control: School systems continue to decide which ELA curriculum is used during instruction and which Innovative Assessments students take.

Current Test Delivery Status
During the 2021-22 school year, the department “re-started” the Innovative Assessment after both the 2019-20 and the 2020-21 “pause” years due to the pandemic. The grades assessed and the administration type are noted in Table 1.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Curriculum</th>
<th>Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Wit and Wisdom*</td>
<td>Pilot</td>
</tr>
<tr>
<td>5</td>
<td>ELA Guidebooks</td>
<td>Pilot</td>
</tr>
<tr>
<td>6</td>
<td>ELA Guidebooks</td>
<td>Field Test</td>
</tr>
<tr>
<td>7</td>
<td>ELA Guidebooks</td>
<td>Field Test</td>
</tr>
<tr>
<td>8</td>
<td>ELA Guidebooks</td>
<td>Field Test</td>
</tr>
<tr>
<td>7</td>
<td>ELA Guidebooks</td>
<td>Operational</td>
</tr>
</tbody>
</table>

*Great Minds, creator of the Wit and Wisdom curriculum became a partner in 2021.

Grade 7 was administered as an operational assessment to 45 schools across nine districts, as well as a field test to additional districts. Three administrations were given throughout the year (Fall, Winter, and Spring). Approximately 24,000 students across all the grades participated, with nearly 5,000 of those students in the operational assessment.
During the 2022-23 school year, it is planned that the current assessments in Grade 5 will move to a field test administration, and assessments in Grades 6-8 will all be operational. In addition, new content will be developed across Grades 3 and 4, with the expectation that it will be field tested in 2023-24.

**Defining Comparability**

The 2021-22 school year marked the first time the Innovative Assessment was administered operationally for three administrations. As part of the state’s IADA application, comparability has been defined as a student who took the Innovative Assessment Program (IAP) would receive the same achievement level on the statewide LEAP 2025 assessment. To support that premise, the Louisiana Department of Education implemented two activities - an Achievement Level Descriptor (ALD) alignment meeting, and a scaling methodology to produce scale scores for IAP on the same scale as LEAP 2025 assessment.

**ALD Alignment Meeting**

In May 2022, a group of six educators (herein referred to as Panelists) met to determine the degree of alignment between the Grade 7 Innovative Assessments and the LEAP 2025 ELA ALDs. Over the course of three days, the Panelists went through a series of process steps that culminated in judgments regarding the degree of alignment.

The Panelists generally found strong alignment between the items in the Innovative Assessment and the LEAP 2025 assessment. As further noted in Table 2, while there was some judgment of weak alignment, in all cases at least at least 80% of the participants rated the alignment as Moderate or Strong.

<table>
<thead>
<tr>
<th>Level</th>
<th>Achievement Area</th>
<th>Counts (N)</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weak (1)</td>
<td>Moderate (2)</td>
<td>Strong (3)</td>
</tr>
<tr>
<td>2</td>
<td>Reading</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Writing</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Reading</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Writing*</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Reading</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Writing</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Reading</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Writing</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*One panelist did not rate this category.*
Scaling Methodology and Outcomes

All steps of the scaling process were designed to support comparability between IAP scale scores and LEAP 2025 scale scores. Data with the percent of students in each achievement level for LEAP 2025 and IAP show a similar pattern across both assessments, thereby demonstrating strong comparability of LEAP 2025 and IAP (see Table 3).

<table>
<thead>
<tr>
<th>Grade 7 ELA Performance Levels for 2022</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEAP 2025</td>
<td>15</td>
<td>17</td>
<td>24</td>
<td>31</td>
<td>14</td>
</tr>
<tr>
<td>IAP</td>
<td>10</td>
<td>17</td>
<td>27</td>
<td>33</td>
<td>13</td>
</tr>
</tbody>
</table>

Based on these analyses, Louisiana believes it is reasonable to use the Innovative Assessment as the annual determination of student performance within Louisiana’s ESEA-ESSA compliant system of school identification and support.

Creating a Curriculum-Relevant Assessment

As part of Louisiana’s current work in creating an Innovative Assessment – one that is aligned to a specific curriculum - a gap has been identified. Many school systems adopt more than one high-quality curriculum since curriculum is a local choice. In other cases, some school systems may choose not to use high-quality instructional materials. In addition, there are logistical challenges of tying assessment to curriculum: movement of students from school system to another school system, and new students from outside the state coming in at different times.

In creating a curriculum-aligned assessment there are the logistical challenges of tying assessment to curriculum; for the Louisiana Department of Education to fully scale their Innovative Assessment, and continue to allow for local control, a new test design must be created; one that is less dependent on specific curricula, and at the same time is designed to provide the same level of meaningful information to educators that incentivizes deep student engagement in the material and texts throughout the year.

At the heart of this work is the creation of an ELA curriculum-relevant assessment that will allow Louisiana to scale statewide, and that will complement or replace the current curriculum-specific assessments. These curriculum-relevant assessments are expected to:

• Provide the same level of, or better, information to educators; instructionally-oriented and practical, actionable information, and

• Incentivize deep engagement in the material and texts throughout the year.

Initial data from Louisiana noted five school districts to target with a curriculum-relevant assessment. Early outreach indicates all five districts are interested in participating in this work. Those five districts, including their current curriculum use, include:

- KIPP—EL (Expeditionary Learning) Education
- Lafourche—CKLA (Core Knowledge Language Arts)
- LaSalle—CKLA (Core Knowledge Language Arts)
- Morris Jeff—mixture of Guidebooks, Wit and Wisdom
- Red River—EL (Expeditionary Learning) Education
Reflections
As the partners look to build this curriculum-relevant assessment, a critical component of the team’s work forward is a reflection on the past. As part of that reflection, we have noted a set of Lessons Learned, as well as begun to dig into past survey and focus group data.

Lessons Learned
This section of the memo not only calls out some of the Lessons Learned from the Louisiana work since 2018, but attempts to note how these lessons have, or potentially will have, an impact on the work toward building a curriculum-relevant assessment.

- **Develop a strong theory of action from the start and reflect on it often.**
  - **Work so Far:** As we have developed the Innovative Assessment, many decisions have been made, often quickly, and sometimes requiring the team to pivot. The theory of action provides the logical chain of reasoning and a roadmap to ensure decisions are being made that align to the action that is ultimately desired. The current theory of action is articulated in two parts, which correspond to the way in which scores are reported within the assessment program. The first part is the summative logic model that shows how the final, summative scores from the Innovative Assessment are meant to be used by educators, in conjunction with training and support, to improve student learning. The second part is the end-of-unit logic model that shows how the scores from the preceding unit are meant to inform instruction in the subsequent unit.

  **Next Steps:** These models will impact a roadmap of investigations and desired outcomes. For example, we expect to explore the impact of the reports created for this program and how they are being utilized by teachers in the classroom; are they having the desired effect, and if not, what can be changed to ensure they are.

**Figure 1: Summative Logic Model**
• **Involve stakeholders in all aspects of the development.**
  o **Work so Far:** As part of this work, we have involved the people most proximate to the problem in many aspects of development, which include item reviews, passage reviews, data webinars, focus groups, surveys, and score report guidance development. In addition, an Advisory Group has been created for the program that consists of administrators, teachers, and district leaders from the participating districts, whose main job is to offer advice and provide feedback on all aspects of the program – from the test platform, to the test administration windows, to the unit progression.

  **Next Steps:** Our plan is to analyze and understand the themes that have come up from past engagement with stakeholders, expand where appropriate, and ensure that we have incorporated the voices of those most proximate to the problem.

**General Considerations Regarding Lessons Learned**

• **Collaboration**
  o Approaching more curriculum-relevant state assessment designs necessarily opens the door to more shared ownership of assessment design and curriculum decisions. This can lead to more coherence between teaching, learning, and assessment can be a win for everyone involved, especially teachers and students.

• **Be Open to Iteration**
  o The mindset of the team needs to be open to changes and understand that iteration is a natural occurrence within innovation.
    ▪ Our research plan for this work incorporates iteration from the beginning – from iterating on the current test design to new designs, to gathering feedback from those most proximate to the problem, and to making necessary adjustments to better meet education demands and needs.
• **Planning to Plan**
  o Recognizing the program schedule as an important tool to ensure that the teams know what needs to be done, when it needs to be done, and the timeframe the task should take. While typical state summative programs are administered once a year, with plans and schedules being developed in early July and running into the following July, through-year programs, are administered multiple times a year (i.e., three times in the case of Louisiana). This means that at least three times a year teams are working on such components as development, validation, administration, scoring, and reporting, in addition to all of the supplemental and accommodation materials.

• **Capacity is Critical**
  o While smaller chunks of work may seem easier to work with at first glance, it is important to realize that the effort is greater for the State and all teams.
  o Convening stakeholders as early as possible to discuss resources, reviews, back-up plans, and schedule risks. Assumptions that may have worked for the traditional program may need to be adjusted; communication is key.

**Survey and Focus Group Data**
Throughout the different design and implementation phases of the Louisiana IAP a series of student and teacher outreach efforts via surveys and focus groups were conducted. The goal was to engage an initial subset of the stakeholders most proximate to the problem to identify key decision points for assessment and report design that could be implemented. Analyses and themes from these efforts will be used to create a plan for additional outreach efforts during the 2022-23 school year in a principled manner (see details in Table 4).

**Table 4: Survey and Focus Group Data Collections to Date**

<table>
<thead>
<tr>
<th>Audience</th>
<th>Tool</th>
<th>Timing</th>
<th>Grade - Count (N)</th>
<th>Key Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>Survey</td>
<td>Spring 2019</td>
<td>Grade 7 - 35</td>
<td>Reactions to delivery system, use of text, overall test design</td>
</tr>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Spring 2019</td>
<td>Grade 7 - 40</td>
<td>Instructional approaches, materials to support the transition to a new assessment, use of text</td>
</tr>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Fall 2019</td>
<td>Grade 7 - 87</td>
<td>Instructional approaches, use of tools on the assessment materials to support transition, experience compared to LEAP</td>
</tr>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Winter 2020</td>
<td>Grade 7 - 177</td>
<td>Instructional approaches, use of tools on the assessment materials to support transition, experience compared to LEAP</td>
</tr>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Fall 2021</td>
<td>Grade 7 - 38</td>
<td>ELA classroom practice, Guidebooks use, sources of guidance for practice</td>
</tr>
<tr>
<td>Teachers</td>
<td></td>
<td></td>
<td>Grades 6, 7, 8, - 168</td>
<td></td>
</tr>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Winter 2022</td>
<td>Grade 7 - 184</td>
<td>ELA classroom practice, Guidebooks use, sources of guidance for practice</td>
</tr>
</tbody>
</table>
### Teachers Survey
#### Winter 2022
- Grade 6, 7, 8 - 63
- Grade 5 - 30
- Implementation survey: End of year feedback on how things went – related to system, manuals, and what could be done to improve

<table>
<thead>
<tr>
<th>Teachers</th>
<th>Survey</th>
<th>Grade</th>
<th>Implementation survey: End of year feedback on how things went – related to system, manuals, and what could be done to improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td>Survey</td>
<td>Spring 2022</td>
<td>All grades - 34</td>
</tr>
<tr>
<td>Students</td>
<td>Survey</td>
<td>Spring 2022</td>
<td>All grades - 950</td>
</tr>
<tr>
<td>Teachers</td>
<td>Focus Group</td>
<td>Summer 2019</td>
<td>30</td>
</tr>
<tr>
<td>Teachers</td>
<td>Focus Group</td>
<td>Early Winter 2020</td>
<td>10</td>
</tr>
<tr>
<td>Administrators</td>
<td>Focus Group</td>
<td>Early Winter 2020</td>
<td>10</td>
</tr>
</tbody>
</table>

### Teacher Surveys
A recent analysis of the teacher surveys throughout the 2021-22 school year uncovered some key themes that will be critical to our work as we move toward a curriculum-relevant assessment. These include:

- Teachers widely supported tying state assessment to ELA unit texts
- Teachers also supported the creation of a test design that allows for a deep focus on building knowledge
- Teachers engage in a variety of practices that are well-aligned with the design of the Innovative Assessment although using texts to make connections across content areas happens least

In addition, key needs that have bubbled up include:

- Teachers wanted more specific information about student performance and what to do next, especially in the area of Writing
- Teachers wanted more support for the effective use of data to better address student needs

Feedback from earlier implementation surveys was addressed through revisions to the IAP assessment program and more recent surveys show desired improvements or the assessment experience for both teachers and students.

### Student Survey
The recent Student Survey from Spring 2022 uncovered that a large percentage of the students preferred the Innovative Assessment to the regular LEAP 2025 assessment (66% to 34%). For those students who preferred the IAP, many called out that the preference was related to the fact that the test was administered after the unit of study (37%). Students also reported being confident (44% and 22% very confident) in answering the questions on the Innovative Assessment; only 6% reported feeling not confident. Students' responses overlapped with the written survey responses and illustrated that
testing shortly after lessons, with a shorter time, increased focus, raised confidence, and reduced test anxiety for most students.

**Next Steps**

Gaps, too, have emerged. We currently have no systematic data from parents/families beyond anecdotal evidence. And while we have test and school administrators on the Advisory Group and actively solicit feedback during the monthly “office hour” and the Advisory Group, it seems that prioritizing this group, or possibly school board members, might be appropriate. We are currently developing a data collection and analysis protocol for the next 10 months and will be leveraging the best practices in sampling, data collection, and statistical analysis methodology from our partner institutions.
Overview of the Research Design
At the core, our research design is simple – it attempts to overlay the curriculum analysis with feedback from the people most proximate to the problem (see Figure 3).

Figure 3: Research Design Phases

As noted in the previous section, we are analyzing our previous data collections to better understand current themes and possible gaps, so we can develop meaningful research questions that recognize and honor the time stakeholder groups have already provided to this program. In Phase II, we expect to be able to use those research questions to drive both empathy interviews and small focus groups. During Phase II, we also expect to show test design models and gather feedback regarding what works and what doesn’t work in those models. Finally, in Phase III, we expect to finalize our test design model, and with it more broadly engage with the people most proximate to the problem, and dig into areas such as use of data, reimagining score reports, and making instructional connections.

The Curriculum Analysis
The five curricula noted by Louisiana as the most used in Grades 3-5 (with ELA Guidebooks as the most used) throughout the State included:

- Amplify Core Knowledge Language Arts
- Great Minds Wit and Wisdom
- HMH Into Reading
- Louisiana ELA Guidebooks
- Open Up Resources OR Learnzillion EL (Expeditionary Learning) Education
As part of their work, Johns Hopkins implemented an analysis of the five curricula and analyzed domain, topical, and sub-topical coverage. In addition, they looked at recurring texts across the five curricula. The analysis included both anchor and core texts (assuming they were clearly noted or specified), but the analysis did not take into account any extension activities.

At a high level, their conclusions included the following:
- No texts were shared across all five curricula
  - While some texts were shared between two or three curricula at Grades 3 and 4, there was no overlap in Grade 5 for all five curricula.
- Some topics were shared across the curricula, such as marine animals and emotions, but overwhelmingly there were minimal shared topics
- Very little was shared across the curricula at the sub-topic level

**Design Challenges**

Getting to the administration will require some deep thinking based on the findings of the curriculum analysis, as well as reflection on themes and feedback from the stakeholder engagement plan. Key questions regarding the specific content and test design that the team will need to address include:
- Given that there are no common texts across the five curricula, how do we create a test design that still honors and focuses on deep engagement in texts?
  - And at the same time, provides instructionally-relevant information to educators?
  - How should we address different teaching sequences?
- What communications will be needed to ensure that school districts and students can have a successful testing experience?
- What other considerations should there be related to how, what, and when?

  ▪ Possible considerations regarding **what** is assessed:
    - Items and tasks that are directly related to the content experienced within one or more curricula (e.g., directly leveraging common texts or phenomena students have experienced or interacted with, or designing for intentional transfer);
    - Ensuring consistent interpretations of standards within instruction and assessment (e.g., similar use of standards bundling, unpacking across assessment and instructional materials); and
    - Designing assessments to follow similar within- and across-year learning progressions for how students are expected to develop the targeted knowledge and skills.

  ▪ Possible considerations regarding **how** it is assessed:
    - Explore the inclusion of tasks that engage students in similar sense-making and reasoning routines as those used to develop disciplinary proficiency within instruction;
    - Provide structures on an assessment for students to make their thinking visible that are familiar to students; and
    - Attend to the ways students have experienced transferring understanding to new contexts.

  ▪ Possible considerations regarding **when** assessment takes place
    - Explore structuring the assessed domain across multiple assessments and allow for flexibility in administration.
The team will be using weekly team meetings over the course of the next two months to brainstorm around the results of the analysis; a planning session is scheduled for late October to create test design sketches to present to key stakeholder groups.

**Engaging with People Most Proximate to the Problem**
While the team engaged with both students and teachers in Spring 2022, more engagement is needed to ensure a solid understanding of the needs identified by the people most proximate to the problem. Current thinking of the team includes engaging with the Louisiana IAP Advisory Group or the Future of Assessment Advisory Group for continuous feedback and oversight. In addition, the team expects to conduct additional engagement activities with teachers, parents, and administrators, starting in the fall and continuing through early winter. In addition, a separate Teacher Survey is planned in early fall which will complement the curriculum analysis that has been completed. Engagement questions and protocols will be dictated by the analyses conducted on past surveys (see Table 5).

**Table 5: Engagement Plan**

<table>
<thead>
<tr>
<th>Wave</th>
<th>Action</th>
<th>Who</th>
<th>When</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Wave</td>
<td>Advisory Group</td>
<td>~5–8 members representing perspectives of educators, business leaders, and parents</td>
<td>August</td>
<td>Continuous feedback and oversight</td>
</tr>
<tr>
<td>Wave 1</td>
<td>Teacher Engagement</td>
<td>Focus groups/interviews of ~16 teachers</td>
<td>September &amp; October</td>
<td>Triangulate key themes and address missing data from the surveys</td>
</tr>
<tr>
<td>Wave 2a</td>
<td>Administrator Engagement</td>
<td>Focus groups/interviews of ~6–8 administrators responsible for curriculum design / selection</td>
<td>November</td>
<td>High level feedback on test design sketches and use cases</td>
</tr>
<tr>
<td>Wave 2b</td>
<td>Parent Engagement</td>
<td>Focus groups/interviews with ~25 parents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 3</td>
<td>All stakeholders - revisit engagement</td>
<td>Focus groups interviews</td>
<td>January February</td>
<td>High level feedback on final pilot design and reported information</td>
</tr>
</tbody>
</table>

**Prototype Plan**
At a high level, the team expects to administer a curriculum-relevant pilot assessment to somewhere between 300–900 Grade 5 students during the Louisiana Innovative Assessment Spring Administration (April 25 – May 26, 2023). Students will come from any of the five participating districts; one to three forms will be administered, depending on total student count.

High-level milestones for this work include:
- Planning Session – 10/26/22 and 10/27/22
- Content Development Begins (target) – 1/9/23
- Content Finalized – 4/1/23
- Administration Window – 4/25/23 – 5/26/23
As the work on the test design becomes more defined, the team will add more details to these high-level milestones.
Section 1: Project Plan Overview

The ongoing work under the Louisiana Innovative Assessment (IAP) Program continues to wrestle with the varied curricula in the state, as well as accessing those whose teaching and learning is based on less popular curricula. The current IAP includes curriculum-specific assessments for the English language arts (ELA) Guidebooks and Wit & Wisdom curricula - and most schools use these two curricula. However, not every school does. Maintaining choice and local control are priorities in Louisiana, and the Crawfish Project aims to address these state needs.

The Crawfish Project attempts to create solutions to the varied curricula problem and address how the department can fully scale their Innovative Assessment, by continuing to allow for local control that is less dependent on specific curricula, and at the same time, provide the same level of meaningful information to educators. This model draws on the Innovative Assessment principles to incentivize deep student engagement in the material and texts throughout the year.

In creating a curriculum-aligned assessment there are the logistical challenges of tying assessment to curriculum. To fully scale the Innovative Assessment and continue to allow for local control that is less dependent on specific curricula, and at the same time, provide the same level of meaningful information to educators. This model draws on the Innovative Assessment principles to incentivize deep student engagement in the material and texts throughout the year.

Background
Since 2018, the state of Louisiana, along with its partners (NWEA, Center for Assessment, Johns Hopkins University, Odell Education, MZ Development, Strategic Measurement and Evaluation), has invested in the creation of a test design that is unlike any other state assessment in the nation. Louisiana employs a high-quality curriculum-integrated through-year approach consisting of several assessments administered throughout the school year, so educators and students receive ongoing feedback to inform teaching and learning.

This Innovative Assessment is explicitly aligned to the ELA curriculum content that the students have studied. It measures both content knowledge and skills students derived from studying specific texts and topics. Of the few states that participated in the Innovative Assessment Demonstration Authority (IADA), Louisiana has been cited as a state that was the most explicit in their attempt to use IADA as an instrument to improve educational equity.
Crawfish Purpose:

Specifically, the purpose of the Innovative Assessment is to improve the following:

- **Integration**: Assessment design strengthens the connection between instruction and assessment by using several brief curriculum-connected assessments throughout the year, which measure students' comprehension of materials and texts that they have studied in class.
- **Focus**: Teachers can focus instruction on background knowledge and making meaning of full texts.
- **Equal Access**: All students have the opportunity to develop background knowledge together so that no student is at a disadvantage due to a lack of life experiences.
- **Preservation of Local Control**: School systems continue to decide which ELA curriculum is used during instruction and which Innovative Assessments students take.

At the heart of this work is the creation of an ELA curriculum-relevant assessment that will allow Louisiana to scale statewide, and that will complement, or replace the current curriculum-specific assessments. These curriculum-relevant assessments are expected to:

- Provide the same level of, or better, information to educators; instructionally-oriented and practical, actionable information, and
- Incentivize deep engagement in the material and texts throughout the year.

Project Plan Overview

Since the release of Memo One, two key outcomes have been realized as part of the work on this program. The first outcome was the creation of multiple test design sketches and the second outcome was a set of focus groups with educators to discuss those sketches.

**Outcome I: Test Design Sketches**

As noted in Memo One, as part of their work, Johns Hopkins University implemented an analysis of the five curricula and analyzed domain, topic, and sub-topic coverage. In addition, they looked at recurring texts across the five curricula. The analysis included both anchor and core texts (assuming they were clearly noted or specified), but the analysis did not take into account any extension activities.

At a high level, their conclusions included the following:

- No texts were shared across all five curricula.
  - While some texts were shared between two or three curricula at Grades 3 and 4, there was no overlap in Grade 5 for all five curricula.
- Some topics were shared across the curricula, such as marine animals and emotions, but overwhelmingly there were minimal shared topics.
• Very little was shared across the curricula at the sub-topic level.

Based on these outcomes, further investigation continued. The content partners dug deeper into what could be gleaned from the analysis, as well as, what was available and known about the curricula. Through many brainstorming discussions throughout the fall, two design proofs were created, and a “bucket” of questions and considerations for the partners to contemplate was created. The two proofs were the CrawFish Funnel, and the Build Together Knowledge Model. From these design conversations, many questions and considerations have been highlighted. The evolution of those thoughts is referred to as the “Option 3 Bucket.”

The CrawFish Funnel Design

The CrawFish funnel looks across the various units of study across the curricula and incorporates sections that are both unique and common.

Specifically, the funnel model includes:
• unique, unit-based “Knowledge” sections
• a common thematically related warm read text and item set in the “Application” section, and
• a common writing prompt in the “Synthesis” section.

This model continues to honor the background knowledge taught in the classroom, but then “funnels” students into a common “Application” and “Synthesis” sections. See an illustration of the design in Figure 1 below.
Build Together Knowledge Design

The core set of beliefs around this design was that students would have the ability to build knowledge together on the test by having the passages across the test, both informational and literary, working as a set that models most curricular structures.

Specifically, it was suggested that this model:

- helps mitigate knowledge gaps by allowing students to show off their reading and writing knowledge;
- should be easy to create content since it is not directly tied to curriculum but to curricular approach (intentional building of knowledge using literature and informational texts); and
- brings in other content knowledge (science, social studies), which will encourage elementary teachers to devote more time to other subjects, which has been proven to improve literacy.

See an illustration of the design in Figure 2 below.
The team and the Department have recognized that it's possible to develop and administer ELA assessments aligned to multiple ELA curricula, and that the current approach (multiple curriculum-aligned assessments) is not sustainable. As part of the design discussions, the team has discussed other options based on the cross-curriculum analysis such as building testlets aligned to content commonality through genres, thematic focus, topics, standard clusters, and writing modes.

Some current thinking includes:

Designing one through-course assessment relevant to all 5 curricula used at grade 5, might be possible, but the assessment would lose many of the important pieces of what we have embraced thus far in the Innovative Assessment.

These include:
- The assessment would be so broad that any real connection to ELA curriculum and the knowledge built through that curriculum likely would be lost.
- Reporting would not be as relevant or useful toward impacting instruction.

Connecting the ELA assessment to knowledge built in other disciplines, might also be possible, but includes both advantages and disadvantages:
- The advantage of this approach is that the assessment can draw on the same background knowledge for all students through the disciplines of Social Studies and Science. (Note: Social studies standards in Louisiana are constructed around specific knowledge taught at specific times throughout the year. Science standards contain specific knowledge requirements but do not specify time of year.)
• The disadvantages are that this model relies on knowledge built through another class and possibly another teacher, nor does it honor the knowledge gained through the high-quality ELA curriculum.

Through the IADA waiver, the team has redefined ELA content as the books and ancillary materials read in a class (curriculum), not the skills defined in the ELA standards. This disconnect between ELA "content" and ELA standards separates ELA from other subject areas like math, science, and social studies, and is the challenge for a sustainable ELA assessment system. However, the test design developed values this disconnect, and has shaped the vision.

The two givens on this program thus far have included:

1. Core Design Principle: ELA curriculum content-based assessment system
2. Biggest challenge to successfully developing a sustainable ELA assessment with the core design principle: Multiple ELA curriculum implemented across the state

Creating a sustainable ELA assessment that continues to honor the core design principle will require changes. Questions discussed with the department include:

• Limiting local choice?
• Mandating one common ELA curriculum? Choice between 2 - 3 ELA curriculums?
• Establishing and enforcing a year-long scope and sequence for ELA content and standards?

Recognizing that limiting local choice may not be an acceptable change, the team has considered other changes to successfully develop a sustainable ELA assessment that continues to include the core design principle.

In design, current considerations include:

• Is there a need for the three current sections: Knowledge, Application, and Synthesis
• Should the number of points and item types be reconsidered?
• Should we alter the dependency on the use of fixed forms
  ○ Developing and banking sections independently
  ○ Tagging one Application section to multiple Knowledge sections according to content alignment

In development, current considerations include:

• Creating multiple curriculum-aligned Item Banks.
  ○ Develop a high volume of content aligned items coded to unit focus standards - across multiple curriculums.
• Creating multiple thematically aligned Item Banks.
  ○ Develop a high volume of passage sets tagged to thematically aligned units/modules - across multiple curricula.
Other discussions in administration and reporting are underway. It is expected that the team will continue to work through each of the questions and considerations, determining what, if any, the department is willing to reconsider.

Outcome 2: Involving the People Most Proximate to the Problem

To date, the second key outcome of Phase 2 has been the involvement of those most proximate to the problem in determining a curriculum-relevant design. While the team was unable to fully implement the proposed “Engagement Plan” as described in Table 5 of Memo One, many activities have been implemented. Those activities that were implemented are noted below in Table 1.

Table 1: Phase 2 Engagement Activities to Date

<table>
<thead>
<tr>
<th>Wave</th>
<th>Action</th>
<th>Who</th>
<th>When</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Teacher Engagement</td>
<td>Focus Group with 6 teachers</td>
<td>October</td>
<td>Discuss End of Unit reporting</td>
</tr>
<tr>
<td>2</td>
<td>Teacher Engagement</td>
<td>Focus groups/interviews of ~20 teachers</td>
<td>December - February</td>
<td>High level feedback on test design sketches and reporting considerations</td>
</tr>
<tr>
<td>Wave 2b</td>
<td>Teacher Engagement</td>
<td>Survey</td>
<td>December - February</td>
<td>Analyze teachers' classroom practices and utilization of instructional materials. The information gathered in this survey will demonstrate how ELA teachers in grades 3-8 understand and utilize their curricula</td>
</tr>
<tr>
<td>3</td>
<td>Student Engagement</td>
<td>Survey questions</td>
<td>January - February</td>
<td>Feedback on reports and results; how they engage in learning more about reading and writing</td>
</tr>
</tbody>
</table>
In addition, the team is working closely with the department to institute a “Design Review Committee” that can interact on a fairly informal basis with the team to answer a whole host of questions, as well as ask questions in return. Along with that, the team continues to interact with the “BeEngaged Initiative” at the state, a group focused on parent engagement. While to date no specific activity has been planned, the team is attempting to align its work with that of the “BeEngaged Initiative” and utilize its planned activities as appropriate. Finally, while the proposed plan called for engagement with the Advisory Group that assisted with the development of Louisiana’s “Future of Assessment” that work has been paused given that the state has not yet released the “Future of Assessment Report”. A summary of the engagement activities this fall have been provided in Section 2.

**Major Elements Planned for Phase 2**

The major elements planned for the remainder of Phase 2 include:

- Content development and future recommendations
- Prototype testing of the funnel design and score reports, involving (1) administration of funnel test forms, (2) real data prototyping of score reports, and (3) focus groups, expert panels and cognitive laboratories

**Content Development and Future Recommendations**

This work will include the content creation needed to employ the Funnel model. While the team will use existing “Knowledge” sections of the ELA Guidebooks and Wit and Wisdom spring units, additional “Knowledge” sections will be built for those curricula determined to be used in Grade 5, as well as a new “Application” and “Synthesis” sections. Given the need for structure around the use of the Funnel model in the future, the content team will allocate time to do a deeper dive into the curriculum analysis, identifying like units, common themes, and make recommendations to the department regarding how best to implement a full year of the Funnel model.

**Prototype Testing and Reporting**

Recruitment for the implementation of the prototype will be critical leading up to the spring test administration (April 25 – May 26). While the department has identified the top curricula being utilized in the state, where those curricula are being used has been more challenging. A survey sent out to all districts in late 2022 will help to identify
curriculum usage patterns and from there the state is hoping to recruit at least 200-300 students per funnel.

In addition to recruitment, the team also will focus on user support. While much of what is created for the “main” administration (Grades 6-8 operational tests and Grade 5 field tests) can be utilized – such as “How to” guides and online training tools, there likely will be additional user needs for this prototype. Current thinking is at minimum a webinar for these participants and unique testing manuals.

In addition, the team is eager to plan a set of cognitive labs where we will have the ability to engage with students and allow them to think aloud as they are responding to questions. Possible areas for investigation include:

- How learners like to learn generally
- Which kind of test design they prefer in terms of texts they know or don’t know and the topics they are connected to
- What kinds of experiences they have in school and outside of school that build up content knowledge / familiarity with certain topics / genres / discourse communities
- What kinds of general feedback they have about test design
- What kind of information they want from the test
- What they plan to do with that
- What they believe about themselves as writers and readers.
Section 2: Research Update

LDOE and its partners conducted a series of focus groups between October 2022 and January 2023 to collect educator feedback on score reporting options and potential test designs.

In October, a focus group was held with five Louisiana educators participating in the IAP pilot. The purpose of the focus group was to receive educator feedback on two different score reporting options: an end-of-unit report used in school year 2021-22 and a proposed, revised end-of-unit report for school year 2022-23. The main difference between the reports was the metrics.

The original report (i.e., from school year 2021-22) reported student raw scores on reporting categories (knowledge, application, synthesis) which included the student’s raw rubric score on the writing component of the essay, whereas the revised report (i.e., from school year 2022-23) removed the raw scores and instead reported performance levels (i.e., “Meets Expectations” or “Working Toward Expectations”) in each reporting category.

Educators were asked which takeaways they got from the original report and how they acted on those takeaways, if at all, and what takeaways they might get and what actions they might take from the revised report. Educators were also asked which benefits they saw from shifting to the revised report, if any. Overall, participants responded favorably to the performance levels available on the revised report but expressed dissatisfaction around the removal of the raw scores. Educators liked the raw scores from the original report because they said the raw scores provided more granular information than simply two performance levels. The feedback highlighted the need to meet educators requests for benchmarking student performance against a set expectation while still providing granular information to inform instruction.

In December 2022 and January 2023, four additional focus groups were held with approximately 22 Louisiana educators to collect educator feedback on two possible test designs. In test design option 1 (noted as the Build Together Knowledge model in Section 1), content on the ELA test would be topically related to science and social studies curriculum. For example, if students were learning about outer space in science, their ELA test could have a passage about space exploration. Critically, this test would still assess ELA standards, while drawing on ideas from science and social studies. In test design option 2 (noted as the Funnel model in Section 1), the test would retain the current design of knowledge, application, and synthesis sections where the knowledge section includes texts the students received in their curriculum, regardless of which curriculum the district uses.

1 Demographic information was not collected for participants in the October 2023 focus group.
2 All educators identified as female; 18 identified as White, 3 as Black or African American, 1 as American Indian or Alaskan Native
The application and synthesis sections, however, would include texts thematically aligned across the different curricula. Three focus groups in December 2022, representing educators from three Louisiana parishes, were presented with test design option 1 and asked what benefits or concerns they had with the design and if they would want their students assessed with the design, along with open conversation for any other reactions.

Educators expressed concerns with the scope and sequence of science and social studies curricula and how they would know which topics would be potentially included on the ELA test, as well as concern with not having control over the teaching of that content. Similarly, educators commented that this design seemed similar to the existing LEAP test, which already includes passages with science and social studies ideas. An overarching theme of these groups was that they were extremely satisfied with the test design of the current IAP whereby content is directly related to curriculum.

Given the unfavorable reactions to test design option 1, another focus group was held in January 2022 with educators from a fourth Louisiana parish to present a second option. These educators were presented with both designs - option 1 and option 2 as described above - and asked their feedback on each design as well as to indicate which design they preferred. This group echoed the concerns raised by the first three groups with test design option 1 regarding lack of control over how the science and social studies teacher teaches. Educators unanimously favored test design option 2 over test design option 1.

The findings across all groups highlight that educators value granular information about student performance and assessments that tightly align to curriculum. Other states considering through-year assessments may find that these learnings apply in their states as well. As for next steps with the prototyping team, more investigation into the multiple curricula in Louisiana need to be conducted to determine if test design option two is a viable option from a test development perspective.
Section 3: Prototyping Plan/Summary

The prototyping plan involves: (1) piloting a test form built around the funnel design and (2) developing score reports and supports that directly link assessment, instruction, and curriculum together.

**Funnel Design Prototype.** The team is currently developing the funnel design based assessment forms for the three curriculum, EL Education, which is a new curriculum, as well as for ELA Guidebooks and Wit & Wisdom. The purpose of this prototype is to (1) examine student and teacher perceptions of this new design, in terms of questions like: is it useful?, does it support instruction?, and (2) to examine the comparability of these three forms, and specifically to examine the viability of psychometrically linking these three forms together to support the creation of a single summative score to be used in the state’s accountability system. The former, (1), will be addressed via surveys, focus groups and interviews, whereas the latter will be addressed by applying item response theory to see if all the forms can be placed on a common scale. Ultimately, this prototype will help inform LDOE’s next steps towards scaling up the program statewide. One motivating question behind this prototype is around what to do with other curricula, and this prototype will provide evidence about a possible solution to addressing curricular variation in the state.

These new forms will be administered to samples of students with a target size of 300 students. For ELA Guidebooks and Wit & Wisdom, the samples of students will be drawn from larger pools of students participating in the IAP program. All students participating in the IAP are currently rostered within the testing platform, allowing the team to carefully construct a representative sample through classroom level sampling based on the October 2022 Enrollment Data, with emphasis on inclusion of students of color, economically disadvantaged students, and students with limited English proficiency, migrant students, and students with a 504 plan. That is, the team will reach out to individual teachers and invite them to have one or more of their classrooms participate, with appropriate resampling. The sample for EL guidebooks is much less certain, since LDOE is currently recruiting from what appears to be a very small set of implementing schools and districts. Most schools use either Guidebooks or Wit & Wisdom, so finding those using other curriculum is a challenge. At minimum, the team will work individually with teachers and administrators to select classes that best represent diversity in the state.

**Score Reports and Supports Prototypes.** In addition to the development of a test design that can work well with the diversity of curricula in the state, the team has also continued to explore approaches to score reporting. Based on feedback from an October 2022 focus group with five educators and leaders, feedback from Louisiana’s Technical Advisory Committee and ongoing internal conversation, the team has decided to target development on two key score reports:

---

3 These descriptions are drawn directly from the state data.
• A student-level report for parents and caregivers that helps parents and caregivers understand if their student is doing well or not, so that they can support their student as they move into the next unit of instruction, and
• A classroom-level report for teachers that captures classroom level performance in ways that (1) help educators better understand the variation in performance in their classrooms and (2) connects student performance to instructional next steps.

Critically, for the student-level parent report, the team is moving away from an individual score report (ISR) meant to serve multiple audiences to a report that communicates clearly to parents, and in doing so may omit some of the previously provided information. The goal is also to connect these reports to supports. For example, for the student-level parent report, these supports could include (a) a process or protocol for teachers on having conversations with parents on their students’ reports for those who want follow up and (b) supplemental activities that can be flexibly assigned to students to engage with during the upcoming unit of instruction. Similarly, for the classroom-level teacher report, the team is exploring supports that connects current classroom performance to the next unit of instruction, potentially including guidance on what should be prioritized in instruction moving forward. Delivery of these materials is also critically important, as a focus group in February 2023 revealed that educators may not be accessing the currently available supports - that if resources are not integrated directly into assessment or instructional materials, they will not be used.

The team is now currently organizing a “real data” prototyping effort in which teachers and leaders will interact with score report prototypes that are based on their own students’ data. These reports will be either simple PDFs produced by hand or via an interactive R markdown file. Either way, these reports are meant to be quick “napkin” sketches that help teachers, and if possible parents, understand possibilities. Often, without something to react to, focus groups and interviews around reporting can fail to produce feedback that can guide future design.

The real data report prototyping will be conducted with teachers who are part of the “IAP Collaboration Groups” towards the end of February and the beginning of March via focus groups in which teachers will be provided their score reports via a secure file transfer site. During the focus groups teachers will have a chance to review the results, share their interpretations of the reports, consider what next instructional steps they would take, and recommend needed instructional supports and report revisions. The team is implementing these reports and supports in late February and early March so that they can, potentially be used to inform instruction within upcoming units of instruction. We plan to follow up with these teachers during their instruction within the next unit to explore whether and how they were able to use the results provided in the reports as well as the accompanying supports.

These focus groups will be conducted with a set of approximately twenty-two teachers in four parishes. These teachers generally represent the population of teachers in the state, but slightly underrepresented teachers of color in the state (18.1% in our sample
vs. 26.1% in the state) as well as in geographic diversity (mostly from northern districts). The team is now working towards recruiting an additional district or districts to increase representation.

In addition to real data prototyping with teachers and leaders, the team has also implemented a student survey for those in the operational program. This survey asks about student experience with the score reports, as well as about their experiences with reading and writing more generally. The team is also working with the “IAP Advisory Group” to better reach students and parents. One approach that is currently being developed is to recruit students to act as an “expert panel” in which a period of class time is devoted to providing feedback on the innovative assessment program, and in particular its score reports. This approach positions students as the experts and, ideally, also acts in a manner akin to a research-practice partnership with students.
Section 4: Lessons learned and implications for implementation

There are a number of high-level lessons learned from this work:

- **Connecting to Multiple Curricula.** Developing assessments and supports that connect closely with multiple curricula require deep inspection of curriculum and leverage overlap in skills and content. Often, these overlaps are idiosyncratic - meaning that thematic or skills-based overlap on some units of curricula are often happenstance. This means that, ultimately, an assessment design that is closely embedded in curriculum requires enough of these idiosyncratic overlaps to support assessment or that curricula itself be modified. Alternatively, an assessment design may be less closely embedded in curricula, for example be mostly skills based or only partially reference the unit as in the funnel design. Here the team has tried to maintain an element of the embedded approach through the funnel design. The question is still whether such a model is viable at scale, which in turn is related to the number of curricula the program will support. Currently, the funnel design involves at least one unique set of items per unit, for each curriculum. For example, for a single grade we might have 5 unit assessments for each of 5 curricula, resulting in 25 assessments. Assuming the same set up across grades 3 through 8, this would mean 150 assessments, without accounting for multiple versions of the same unit assessment. Importantly, item response theory does not work well with small samples, meaning that curricula that are instructed to very few students will be very difficult to develop viable assessments for.

- **State Curricula Implementation.** Supporting the program at scale requires knowing what curriculum is being taught in each and every classroom. Without that information, it is difficult to judge what assessments will need to be ended. Critically, this kind of information almost invariably cannot be collected at scale without a policy mandate. Voluntary collection in the state, to date, has produced information that is incomplete. Such a collection could be required alongside annual statewide testing or via a separate required collection, but would likely face pushback from various stakeholders.

- **No Shortcuts to Curriculum Relevance.** The approach within this work is to connect the content, timing, reporting and supports of assessment to curricula within the state. The Build Together Design was an attempt to loosen the connection between curriculum and assessment. That design was widely regarded as less desirable than the current model, leading the team to return to approaches that tied assessment content more directly to curriculum. As noted by one teacher during a focus group, “Can assessment not tied to the curriculum truly be innovative?”.

- **Robust Reports and Supports.** The currently provided reports and supports are just the tip of the iceberg in terms of what is needed. The field seems hungry for materials that (1) are easily accessible and (2) connects curriculum to assessment. In one focus group we discovered that educators were using an interim assessment product and related instructional supports alongside the Wit & Wisdom curriculum. This interim product took a skills-based approach to reading comprehension and written expression, whereas the curriculum weaves
content and skills together. This means the program has the opportunity to fill this kind of gap via supports that are more coherent with curriculum and instruction. In addition, one educator noted that it wasn’t the standards-based instruction that was of most value, but the granularity of information that was provided that was of value. Again, this suggests next steps for report and support development.

**Success Conditions.** Ultimately, all of this work is ramping up toward a potential full statewide implementation of the program once the waiver is complete, in 2025-26. In terms of the Funnel Design Prototype, success would involve (1) that educators find similar value, or even better, improved, value in the Funnel design relative to the main program design, and (2) that the items from the three funnel design modules can be placed onto the same scale and that resulting scores are comparable enough to support annual determinations under the state’s ESSA system. Based on this success, the team would then need to find a way to scale up this design to at least field test status by 2024-25 to support operational roll out in 2025-26. Scaling up would involve not only developing enough content to cover the curricula used in the state, but also understanding the implemented curriculum throughout the state.

In terms of the Score Reports and Supports Prototypes, success would entail that the score reports are interpreted well, but also actually support instructional next steps. Restated, that teachers not only find the reports and supports useful, but actually change their instruction for the better as a result. We are providing these reports and supports in late February and early March so that we can examine what happens in the following unit of instruction.

**Student Experiences.** The funnel design prototype is being developed following the principles of Understanding by Design and in compliance with the state’s accommodations and accessibility guidelines. Given this, students will receive appropriate accommodations and supports as with the statewide program. All students will also receive a survey after the administration. The score reports and supports prototypes will be explored with students via expert panels in which students will be asked to comment on the reports and supports. The team plans to develop a protocol to encourage student voice and ensure that all voices are heard. Ideally, this involves an initial panel to help cement the final design.

**Enabling Conditions and Barriers.** The Louisiana Department of Education has taken great strides to develop rapport with the field. This includes a “IAP Advisory Committee” made up of teachers and leaders which meets monthly to provide feedback and guidance on implementation and four “IAP Collaboration Groups” which are also groups of teachers who meet on a semi-regular basis to react to specific topics of interest. These touch points with the field have enabled the team to have ongoing interactions to better understand school and district context and shape the IAP. In addition, the Louisiana Department of Education is highly interested in this project and continues to engage deeply with the work. Finally, this work is enabled by Louisiana’s work on high-quality curriculum, which is supported through state legislation and involves both state funding and state review of curriculum.
There are two key barriers to the future success of this program. First is engagement, or lack thereof, of schools and districts using curricula other than Guidebooks and Wit & Wisdom. Challengingly, these schools and districts are few and far between and ways to identify them are limited. Second are concerns around the scalability and feasibility of the prototype. While a clear plan is in place for spring 2023, what is less clear is how the Funnel model can be scaled for a full year implementation at grade 5, and then beyond. As noted earlier in the memo, the content experts will begin to do deeper exploration of the curricula used and attempt to make recommendations to the department.

**Needed Support.** There are a few areas of support that would be helpful, including:

- Specific protocols for structuring engagements with students or youth
- Suggestions around policy guidance for balancing local control with a curriculum relevant solution.
Section 5: Budget Update

The budget update is provided as a separate document.

Highlights include:

- Much of the activity to date has been related to both design creation and engagement with the people most proximate to the problem. Specific activities have included: creating survey tools, analyzing past and new surveys, creating focus group protocols, conducting focus groups and analyzing results, design exploration, and curriculum analysis.
- No dollars have been spent thus far for meetings or travel. Engagement activities to date have been all virtual. It is expected that dollars in this category will be spent in the spring for the implementation of cognitive labs and interviews.
- No dollars have been spent toward Permissions or License fees to date. Expenditures will be made in these categories as both passage needs are determined and sample size is finalized.
- Content creation will begin in earnest starting early February. Given the importance of understanding the funnel model for an entire year, some dollars will be allocated to more investigation in the curricula to make recommendations to the department regarding what will be required to make this model work for a full year. Other possible designs may also be determined. Any dollars remaining will be utilized for additional new content.
Milestone Memo 3

Section 1: Project Plan Overview

The ongoing work under the Louisiana Innovative Assessment (IAP) Program continues to wrestle with the varied curricula in the state, as well as accessing those whose teaching and learning is based on less popular curricula. The current IAP includes curriculum-specific assessments for the English language arts (ELA) Guidebooks and Wit & Wisdom curricula - and most schools use these two curricula. However, not every school does. Maintaining choice and local control are priorities in Louisiana, and the Crawfish Project aims to address these state needs.

The Crawfish Project attempts to create solutions to the varied curricula problem and address how the department can fully scale their Innovative Assessment, by continuing to allow for local control that is less dependent on specific curricula, and at the same time, provide the same level of meaningful information to educators. This model draws on the Innovative Assessment principles to incentivize deep student engagement in the material and texts throughout the year.

In creating a curriculum-aligned assessment there are the logistical challenges of tying assessment to curriculum. To fully scale the Innovative Assessment and continue to allow for local control, a new test design must be created that is less dependent on specific curricula. As part of Louisiana’s current work in creating an Innovative Assessment – one that is aligned to a specific curriculum - a gap has been identified. Many school systems adopt more than one high-quality curriculum since curriculum is a local choice. In other cases, some school systems may choose not to use high-quality instructional materials. In addition, there are logistical challenges of tying assessment to curriculum. For example, movement of students from one school system to another school system and new students from outside the state coming in at different times.

At the heart of this work is the creation of an ELA curriculum-relevant assessment that will allow Louisiana to scale statewide, and that will complement, or replace the current curriculum-aligned assessments. These curriculum-relevant assessments are expected to:

- Provide the same level of, or better, information (than the curriculum-aligned assessments) to educators; instructionally-oriented and practical, actionable information, and
- Incentivize deep engagement in the material and texts throughout the year.

Since the release of Memo Two, two key outcomes have been realized as part of the work on this program. The first outcome was the creation of collaboration groups that focused on reporting and supporting the question “What do I do next?”; the second outcome was the creation, administration, and initial analyses of the CrawFish forms.
Major Outcomes

Collaboration with Educators and Students on Reports

In February, a series of focus groups were held with educators where the department committed to developing the innovative assessments in partnership with Louisiana educators and leaders. These focus groups were an opportunity for educators and leaders to provide feedback on the current innovative assessment and were also an opportunity to engage more deeply on score reports. Specifically, focus areas included:

- Reactions to the last set of results and what adjustments were made in instruction
- Feedback on the design and usefulness of the current score reports
- Ideas around possible instructional tools, materials or supports that can help teachers improve instruction within the IAP

Based on the feedback, report prototypes were created. In March, the team conducted follow-up with the focus group participants. This time the participants were asked to reflect on the prototypes to get an initial round of feedback and highlight areas where more work was needed. Questions included:

- Did you find anything confusing, unclear, etc.?
- Do you have any questions on the report?
- In general, are the reports useful?
- Can any of these results inform your instruction in your current instructional unit? If so how?

Also related to this topic, a student survey was administered during the winter window, asking questions related to reports and reporting. Over 13,000 students participated across grades 6-8.

More details regarding the findings from the collaboration focus groups, as well as the identified next steps can be found in Section 3.

Creation and Administration of CrawFish Forms

As noted in Milestone Memo Two, the CrawFish funnel design looks across the various units of study across the curricula and incorporates sections that are both unique and common.

Specifically, the funnel model includes:

- unique, unit-based “Knowledge” sections
- a common thematically related warm read text and item set in the “Application” section, and
- a common writing prompt in the “Synthesis” section.
This model continues to honor the background knowledge taught in the classroom, but then “funnels” students into a common “Application” and “Synthesis” sections. See an illustration of the design in Figure 1 below.

**Figure 1: Illustration of the CrawFish Funnel Design**

While the original plan was to funnel three curricula – Guidebooks, Wit and Wisdom, and EL – the district using EL (Red River) decided not to participate in the pilot, leaving only Guidebooks and Wit and Wisdom.

After multiple discussions with the team, the decision was to create the CrawFish pilot form (Form C) for Guidebooks and Wit and Wisdom and spiral with the curriculum-aligned field test forms (Forms A and B) at the class level. More details regarding the sampling design can be found in Section 3.

**Major Elements Planned for Phase 2**

The major elements planned for the remainder of Phase 2 include:

- Completion of the administration
- Administration of a Grade 5 Student Survey
- Student Writing Task response analysis
- Item analyses
- Student focus groups
- Parent survey regarding reports

All of the work has been completed and/or will be completed and is discussed in greater detail in the following sections.

**Section 2: Summary of feedback and input from people proximate to the problem**

**Feedback and Input from the People Most Proximate to the Problem**

The team has put in great effort to engage with the people most proximate to the problem and has been honored to have had the many voices who have contributed. Our focus to date has been mainly educators and students, but in late March the team was provided with the opportunity to engage with parents/caregivers.

Overall impacts are noted in Table 1, but those that were most defining included:
- The educator focus groups that focused on the test design sketches. This feedback was used to finalize the prototype design.
- The educator focus groups and follow up that focused on class report designs.
- The student survey from Winter which will support the main tenants for a redesigned individual student report, and
- The student survey from spring 2023, which will help the team better understand if student experiences across curricula are similar
- New relationships were formed with the Parent and Family Engagement Coordinators.

**Table 1: Engagement Activities to Date**

<table>
<thead>
<tr>
<th>Action</th>
<th>Who</th>
<th>When</th>
<th>Purpose</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement with Teachers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Focus Group with 6 teachers</td>
<td>October 2022</td>
<td>Discuss End of Unit reporting</td>
<td>Initial information to create new way to organize information in Class Report</td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Focus groups/interviews of ~20 teachers</td>
<td>December 2022 – February 2023</td>
<td>High level feedback on test design sketches and reporting considerations</td>
<td>Supported the decision in using the Funnel model for the spring administration</td>
</tr>
<tr>
<td><strong>Teacher Engagement</strong></td>
<td><strong>Survey</strong></td>
<td><strong>December 2022 – April 2023</strong></td>
<td><strong>Analyze teachers’ classroom practices and utilization of instructional materials. The information gathered in this survey will demonstrate how ELA teachers in grades 3-8 understand and utilize their curricula</strong></td>
<td><strong>Better understand curriculum use in the state</strong></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------</td>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td><strong>Teacher Engagement</strong></td>
<td><strong>Collaboration Groups</strong></td>
<td><strong>February – March 2023</strong></td>
<td><strong>To support the work of a report redesign at the class level</strong></td>
<td><strong>Redefined class level report</strong></td>
</tr>
<tr>
<td><strong>Teacher Engagement</strong></td>
<td><strong>Collaboration Groups</strong></td>
<td><strong>May – July 2023</strong></td>
<td><strong>Receive feedback on redefined class level report</strong></td>
<td><strong>Create final adjustments for use in 2023-24</strong></td>
</tr>
</tbody>
</table>

**Student Engagement**

<p>| <strong>Student Engagement</strong> | <strong>Survey</strong> | <strong>May 2022</strong> | <strong>Focused on engagement, demonstration of learning, assessment preferences, experience, empowerment, preparation, and results</strong> | <strong>Gain an understanding of student’s general thoughts regarding the innovative assessment</strong> |</p>
<table>
<thead>
<tr>
<th>Student Engagement</th>
<th>Survey</th>
<th>January-February 2023</th>
<th>Feedback on reports and results; how they engage in learning more about reading and writing</th>
<th>Will support the creation of an updated Individual Student Report; will also be the basis for a Student Focus Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Engagement</td>
<td>Student Survey (Optional 5th grade survey)</td>
<td>May 2023</td>
<td>The focus will be on their experience with the assessment. Our proposed analysis is aimed at comparing and contrasting student experiences across the W&amp;W and Guidebook forms.</td>
<td>Evidence to determine if the assessments related to the different curricula are providing similar student experiences.</td>
</tr>
<tr>
<td>Student Engagement</td>
<td>Focus Groups</td>
<td>Summer/fall 2023</td>
<td>Expected to support the work of a report redesign at the individual student level</td>
<td>Redesign individual student report for use in 2023-24</td>
</tr>
</tbody>
</table>

### District Engagement

| District Engagement | Survey | January 2023 | Collection of information regarding curriculum usage at the district level; gauging general interest level in participating in pilot. | Supported the creation of “Collaboration Groups” that supported the work on a report redesign |

### Parent Engagement

| Parent Engagement | Presentation to the Parent and Family Engagement | March 2023 | To introduce the PAFE team to state assessment in Louisiana, and discuss possibly opportunities | New relationships created; intend to utilize interested coordinators in |
### Parent Engagement

**Assessment Collaboration Meeting with PAFE Coordinators**

- **May 2023**
  - To discuss with the coordinators:
    - What conversations are you having with parents?
    - How can we enter into these conversations with parents’?
  - Continue to develop relationships with the PAFE coordinators; understand better how to engage with parents in the established framework.

**Parent Engagement Survey**

- **Summer and Fall 2023**
  - Expected to support an understanding of parent perspective on reports.
  - Expected to support the redesign of an individual student report.

---

High-level plans for next school year include the continued engagement with the people most proximate to the problem through focus groups, surveys, and potential cog labs. While the research questions are not fully defined at this time, it is expected that they will center on utility-based questions such as:

- Do teachers, students and parents find value in a curriculum relevant approach to state testing? If so, how and what?
- What kinds of information in the score reports do students, teachers, and parents/caregivers pay most and least attention to?
- Which interpretations are appropriate and which ones reflect misunderstandings about the assessment purpose and design?
- How valuable are the supports for further learning and/or instructional guidance provided alongside the reports to help students, teachers, and parents/caregivers? How effective are these actions for supporting actual improvement?

---

### Section 3: Prototyping Summary

Prior to the start of this grant, as noted multiple times in the various memos, Louisiana already had in place an Innovative Assessment Program (IAP) for its “curriculum-aligned” assessments. This created some key efficiencies with incorporating the CrawFish pilot, such as:
- Integrating CrawFish/Form C as part of the ongoing Grade 5 field test
- Use of the same platform
- Use of the same technical support tools and Help Desk
- Use of the same manuals
- Use of the same communications

At the core, incorporating the pilot form into the Grade 5 field test meant not needing to treat the pilot differently. The team saw Form C as another form of the same unit, therefore lessening the need for unique, or even possibly confusing, communication to the field.

As noted earlier, the decision was to create the CrawFish pilot form (Form C) for Guidebooks and Wit & Wisdom and spiral with the curriculum-aligned field test forms (Forms A and B) at the class level.

The sampling design details, found in Table 2, were the following:

<table>
<thead>
<tr>
<th>Administration</th>
<th>Total Actual Count</th>
<th>Form A #</th>
<th>Form A %</th>
<th>Form B #</th>
<th>Form B %</th>
<th>Form C #</th>
<th>Form C %</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW</td>
<td>1,413</td>
<td>608</td>
<td>43</td>
<td>608</td>
<td>43</td>
<td>198</td>
<td>14</td>
</tr>
<tr>
<td>Guidebooks</td>
<td>2,422</td>
<td>1,114</td>
<td>46</td>
<td>1,114</td>
<td>46</td>
<td>194</td>
<td>8</td>
</tr>
</tbody>
</table>

All forms were randomly assigned based on overall district curriculum choice – ELA Guidebooks or Wit & Wisdom. The sampling design allowed the team to have the necessary counts for the field test and subsequent analyses (Forms A and B), but also provided a reasonable sample size for the CrawFish form (Form C) with which to analyze performance on both the machine-scored items and the writing task.

**Final Prototype Design**

As noted earlier, the funnel design looks across the various units of study across the curricula and incorporates sections that are both unique and common.

Specifically, the funnel model includes:

- unique, unit-based “Knowledge” sections
- a common thematically related warm read text and item set in the “Application” section, and
- a common writing prompt in the “Synthesis” section.
For the spring 2023 pilot administration, the CrawFish form for both curricula followed the design pattern as shown in Table 3.

**Table 3: Comparison Between the Two Curricula**

<table>
<thead>
<tr>
<th>Test Section</th>
<th>ELA Guidebooks: Shutting out the Sky, Form C</th>
<th>Wit and Wisdom: A War Between Us, Form C</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Unit text pairings; unique items</td>
<td>Unit text pairings; unique items</td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>Same paired excerpts; common items</td>
<td>Same paired excerpts; common items</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>Common Opinion Writing Task (WT) based on same warm read pairing (from the Application section) and one unit text (different unit texts)</td>
<td>Common Opinion Writing Task (WT) based on same warm read pairing (from the Application section) and one unit text (different unit texts)</td>
<td><em>The content of the ELA Guidebooks and Wit &amp;Wisdom Crawfish opinion writing prompts are the same, however the directions are different, and the unit texts are different - just as they are in the Innovative Assessment Program for ELA Guidebooks and Wit &amp; Wisdom</em></td>
</tr>
</tbody>
</table>

From an administration perspective, as noted earlier, the team was able to utilize the already occurring Innovative Assessment Program (IAP) Grade 5 field test, as well as key components of the program including same communication channels, same test manuals, same technical support tools and protocols, and the same platform. While results will not be shared with the schools or students from the CrawFish pilot (i.e., score reports), results from the writing task and machine-scored items will be utilized to inform future CrawFish planning and thinking.
Analyses Planned and Conducted

Student Results

The key research questions the team will attempt to answer include the following:

- How does student performance on each section of the “funnel” (knowledge, application, synthesis) differ by students’ curriculum type and/or by student demographic category?
- How do the Crawfish writing responses compare to trends on IAP rubric scores?
- Are we seeing differences in responses - compared to IAP and possibly compared across CrawFish forms? Are they possibly attributed to test design?

To answer these questions, the team plans to analyze both the writing and the machine-scored components of the assessment. Due to the timelines associated with the memo, the team has identified both short- and long-term analyses as noted below.

Short-Term Analyses:
- **Hand-Scored Items:** Understand how students are performing on the Writing Task on the curriculum relevant assessment (CrawFish)
- **Machine-Scored Items:** Understand how students are performing on Form C

Long-Term Analyses:
- Understand how performance on Form C compares to the performance on Form A and Form B (IAP)

Initial Review of Writing Responses

For the Writing Tasks, we examined a sample of student responses and reviewed for patterns and trends, using the rubric as the guide. We looked to see how/if the student used the warm read in the essay and followed the prompt directions. Initially these analyses looked across the CrawFish form across the two curricula.

This initial review seemed to indicate that the distribution of responses to Form C (for both curriculum) are pretty typical to the responses observed on the IAP from its inception. The majority are in the low score bucket, some are in the medium score bucket, and a few in the high score bucket. Some specific observations that could influence the prompt for next year’s pilot are the following:

- Students often bring in the warm read into their response, but many are not including the anchor text(s);
- Many students focused on the quote that was provided as part of the prompt and did not address the prompt itself.
Based on these findings, the team will determine what may need to change with the creation of future prompts. Future analyses will explore how the performance on the CrawFish forms versus the performance on the non-CrawFish forms.

<table>
<thead>
<tr>
<th>SOTS Form C</th>
<th>War M3 Form C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td><strong>Low</strong></td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td><strong>Medium</strong></td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>Grand Total</strong></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>60</td>
<td>52</td>
</tr>
<tr>
<td>27</td>
<td>24</td>
</tr>
<tr>
<td><strong>Total Form C Counts:</strong> 126</td>
<td><strong>Total Form C Counts:</strong> 142</td>
</tr>
<tr>
<td><strong>Percent Captured:</strong> 75.39%</td>
<td><strong>Percent Captured:</strong> 57.74%</td>
</tr>
</tbody>
</table>

**Initial Review of Machine-scored Items**

For the machine-scored items, we will be looking at overall p-values and correlations to determine how student performance on each item, and how well the item differentiates between high and low achievers. From that review, adjustments will be made to items accordingly. While the initial analyses will look specifically at Form C, future analyses will look across the grade 5 forms, and compare performance on the innovative assessment forms to the CrawFish forms, in an attempt to answer if performance on the funnel design/curriculum-relevant design is similar to that of the Innovative Assessment Program.

*Note: The initial review of machine-scored items is still being implemented and will be forwarded to Ed First as soon as it is completed.*

**Stakeholder Engagement**

**Educators**

As noted in the table in Section 2, there has been considerable engagement with the people most proximate to the problem during the spring and early summer of 2023. While there was some engagement with students and parents/caregivers, most of the engagement this spring focused on educators and centered on how best to report results and the kinds of tools needed to support adjustment to instruction.

Some key take-aways as it related to instructional supports included:
- Hearing how other teachers use different activities in their classes was considered very helpful by the educators. They called out the importance of opening up conversations between other parishes and educators, and how it would be an excellent way to learn from each other.
- Noting the importance of supports that provide opportunities for success for students who are grade levels below in reading, as well as those that have special needs.
- More exemplars, ideally from real students (and more practice tests or released items, potentially with explanations on how the examples or practice tests connect to the curriculum and the student activities therein)
- Consideration or explanation of the alignment between the innovative writing task (on the assessment) and the culminating writing task in the Guidebooks curriculum (i.e., students to refer to one piece of text in the culminating writing task whereas the innovative assessment requires multiple texts)
- Better connections to current resources
- Videos of teachers implementing the same lessons and same standards as teachers are currently implementing
- Process for student self-scoring and reflection
- Guidance for students about why they are taking this test and why it matters

As it relates to reports, some key take-aways included:

- A desire to understand the patterns of performance across the classroom, and clear next steps based on those patterns.
- A process/protocol for having conversations with students on their results from a growth mindset perspective.
- An explanation or articulation that connects the reporting categories to the standards, in a much detail as possible
- Next steps for students who are strong on the reporting categories
- More detail on students who are scoring zeros and well as recommendations for instruction
- Having indications of “on track” and progress

Initial discussions asked for feedback on various prototypes for reports. Educators were then provided samples of the prototypes (using their data from the Window 2 administration) to utilize in the classrooms. During the summer and into the coming school year, the plan is to go back to the educators and interview them on how useful they found the reports, how they engaged with the data included, and what suggestions they have to improve them. Outcomes of those interviews will be used to finalize a new class level report which will be in production for the fall 2023 window.
Students – Spring 2023

As part of the Spring 2023 administration, a Grade 5 optional survey was made available to all 5th graders participating – those who took the curriculum-aligned versions of the assessment, and those who took CrawFish (curriculum-relevant). The survey focused on the student experience with the assessment. Of the nearly 4,000 5th grade students who participated, only about 200 students responded to the survey.

The survey and detailed results can be found in Appendix 1 and Appendix 2.

High-level results across both curricula include:

- Majority of the students (>85%) said that they tried hard to do well on the assessment
- Greater than 85% were confident that they did well.
- Greater than 80% said that they are familiar with the type of writing that they needed to do on the assessment.

Students – Winter 2023

During the Winter 2023 administration, a student survey was implemented as part of the assessment. While not required, it was highly recommended. Over 13,000 students responded across grades 6-8.

Questions focused on reporting – and what information students would like to receive, what they typically do with that information, what they want to share with their parents/caregivers, and what they know about how to become a better reader and writer. The survey questions, a link to the results, and a Power Point can be found in Appendix 3 and Appendix 4.

Some key highlights include:

- Students want to know how they do on the assessment
- Students appear to want their performance recognized by their parents/caregivers
- They have different expectations regarding discussions with teachers after the assessment – some want more feedback about what they can be doing better, while others are content with just moving on.
- Many students have some ideas about how to become better readers and writers but would benefit from guidance from their teacher.

These findings are being used to support the creation of a new student report for the 2023-24 school year.

Areas for Further Development

For spring 2023, the model worked well. A deeper dive into the curriculum allowed the team to find common themes across the units for ELA Guidebooks, Wit & Wisdom (including EL too, but
Red River opted not to participate), allowing for unique items in the Knowledge section, and common items across the Application and Synthesis sections.

However, there are areas of concern with the current model.

- The inherent nature of the model (funnel) requires the ability to cluster like units together to ensure common themes with which to build the Application and Synthesis sections. This in turn requires the state putting in place a de facto scope and sequence. While the state does impose a scope and sequence for the ELA Guidebooks (Level I, Level II, Level III), and Wit & Wisdom recommends a clear ordering of their modules, will/can the state impose a scope and sequence for all curricula that are used? How does this impact local choice?
- The model is not the solution for all scenarios. This is because it assumes a level of fidelity with the determined curricula. What happens with students that migrate into the state mid-year? What happens with students who are unable to attend school for crucial learning of the curriculum?

Based on these concerns, the team expects to continue to dig into the curriculum analysis and think through additional prototype designs for review and discussion. A key area for discussion will include the role of knowledge in the assessment, and how/where should that knowledge be built.

In addition, we were hampered by the lack of data being collected regarding curriculum and curriculum use in the state.

- We don’t have a good understanding of the curriculum being used in the state. While the state initially considered 5 curricula for the grant study, the state could not identify anyone using two of the curricula, and only about 80 or so students in Grade 5 using the third curriculum.
  - If the state intends to pursue a curriculum aligned assessment model, a yearly systematized data collection on the use of curriculum in the state is required.
- While we have some self-reported data on the level of fidelity the curricula are implemented, it is not clear if the curricula are being implemented at the level of fidelity required to scale a curriculum-aligned or curriculum-relevant assessment.
  - As with bullet number one, this is another area where a systematized data collection is required.

Section 4: Lessons Learned and Implications for Implementation

Primary lessons
Key to this work has been our engagement with the people most proximate to the problem. From the early focus groups on designs, to the collaboration with teachers on reports, to the
student surveys and focus groups, what we have learned has provided us with opportunity to adjust along the way. It has also afforded the assessment and data teams at the state with the opportunity to engage with groups that they have not in the past. For example, the Be Engaged Initiative at the department has opened doors to parents and caregivers the department has not had previously. For the state of Louisiana, stakeholder engagement, and a true focus on the people most proximate to the problem, has been a new experience. An invaluable experience. The hope is that the relationships that have been built this year, will continue well into the future.

As noted earlier some key impacts from our engagement have included:

- The educator focus groups that focused on the test design sketches. This feedback was used to finalize the prototype design.
- The educator focus groups and follow up that focused on class report designs.
- The student survey from Winter which will support the main tenants for a redesigned individual student report, and
- The student survey from spring 2023, which will help the team better understand if student experiences across curricula are similar.

The role of engaging with the people most proximate to the problem will continue to be a significant effort into the coming school year. A current highlight of engagement activities to come include:

- Parent/caregiver engagement on reports
- Collaboration with educators on support tools
- Continued engagement with students, as expert panels and through cog labs

Criteria

There are different criteria for different aspects of the project, and our work has been guided by these implicit criteria. This project has operated from a continuous improvement model, meaning that even when we meet these criteria, we are still looking for ways to improve.

- End Of Unit Report and Supports Prototyping:
  o Do teachers and leaders report that (1) they understand the reports and supports, (2) that they think the reports and supports will be helpful in guiding instruction, and (3) they have feedback or recommendations on revision.
  o Do teachers report, after a unit of instruction, that they used the information within the reports and supports to inform their instruction? If so, how. And if not, why not? And in either case, do they have feedback or recommendations revisions?
- Comparability of Crawfish Forms:
Finding that students have the same patterns of writing between Crawfish and Main forms

- That the shared items across the forms function similarly with respect to the crawfish model

- **Curricula Overlap**
  - That curricula information can be collected across either the full state, or a representative sample (i.e., curriculum of interest can be identified)
  - That there is sufficient overlap between identified curricula to support common application and synthesis sections.

**Summative Score Creation**

At this time, it is assumed that creation of summative scores from the CrawFish design will match what has been used on the larger Innovative Assessment Program. The approach for the IAP is to pool data across units and estimate scale scores using an IRT model, essentially producing an average theta across assessment administrations. In doing so, the IAP intentionally shifts the claim away from performance at the end of the year. The IAP draws on a value that each unit is critical to student learning and should therefore be factored into annual determinations.

**Enabling Conditions and Barriers**

As discussed in Section 3, data collection focused on both curriculum use, and fidelity of implementation has been a huge barrier to the team this year. It is hoped that the state can put into place a systematized data collection process that will allow real-time understanding of what curricula are being used, and how they are being used, state-wide.

Another key area is the connection and collaboration between curriculum and assessment. In the state of Louisiana where the ELA Guidebooks have been created, maintained, and updated, connections between the two departments are ongoing. While there is room for improvement, generally the assessment team understands the curriculum roadmap, and the curriculum team is involved in reviews. But Louisiana is a special case. In no other state in the country is this happening. How does one create the relationships outside of the state with a curriculum vendor? What influence and control does a state/should a state have? Those are much more difficult questions to answer and will need to be figured out if the department wants to continue with a curriculum aligned model.

**Implementation Plan**

The current plan for school year 2023-24 is to continue with Grade 5 and do a full-year pilot. At this time, given the information related to curricula used in the state, the plan will be as shown in Table 4.

**Table 4: Implementation Plan for School Year 2023-24**
<table>
<thead>
<tr>
<th>Test Design Administration</th>
<th>Funnel Model</th>
<th>New Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall Window</td>
<td>X (ELA Guidebooks and Wit &amp; Wisdom at a minimum)</td>
<td></td>
</tr>
<tr>
<td>Winter Window</td>
<td>X (ELA Guidebooks and Wit &amp; Wisdom at a minimum)</td>
<td></td>
</tr>
<tr>
<td>Spring Window</td>
<td>X (Tentative; will be dependent on what is discovered during the year)</td>
<td></td>
</tr>
</tbody>
</table>

A more detailed implementation plan for CrawFish for school year 23-24 can be found in Appendix 5.

**Section 5: Budget Update**

The budget update reflects the budget amendment recently approved.

As noted in Memo 2, the funds set aside for meetings, travel, and permissions were not spent as expected. The budget amendment reflected the following:

- To support the qualitative research analyses related to engagement activities between now and the end of the grant period.
  - Specifically, the team plans to continue with the teacher collaborative groups, begin a parent collaborative group, host a parent survey, conduct a student survey focused on test experience, and engage in a student focus groups.
- To support the cost of an additional program manager to support the coordination and support of these activities.
- To support the cost of two Inquiry Research activities focused on both educator and administrator reactions to assessments connected to curriculum.

In addition, NWEA contributed additional labor to the program (approximately $23,000) as seen in the Grantee contribution column.

The final budget is attached as a separate document and can be found in Appendix 6.
Education First Final Memo

Through-Year Curriculum-Connected Assessment Grant Program

June 30, 2023
Brief Project Overview

The Louisiana Department of Education (LDOE) through-year assessment project focuses primarily on the Scope and Sequence aligned model (Model 2), but also honors school system choice of curriculum resulting in some elements of the Curriculum-aligned model (Model 1) in the LDOE design. The project also includes two different disciplines—mathematics and English language arts—with different national partnerships for both programs.

Since 2018, Louisiana has been developing through-year assessments with approval under the Innovative Assessment Demonstration Authority (IADA). The LDOE focused on a design for ELA assessments that would be more instructionally relevant. That application and subsequent development, called the Innovative Assessment Program (IAP), focused on building a through-year assessment in ELA that was aligned to high-quality instructional materials from a single curriculum, ELA Guidebooks.

Louisiana’s theory of action is grounded in the belief that by aligning and integrating state assessments with curriculum, the assessment system can better support high-quality classroom instruction. With a through-year model students take a test soon after a unit is completed and educators can use the results to inform the next unit of instruction. The curriculum-aligned assessments are designed to tap into the knowledge built in the units of study. Overall, the program encourages the reduction of testing through the implementation of this comprehensive assessment system. The LDOE hopes to use what we have learned from this project to design improved assessment models that can be adopted statewide.

This project enables the LDOE to design and pilot ELA & Math assessments relevant to other high-quality curricula in order to begin scaling through-year assessments for all Louisiana school systems. For the ELA project, LDOE continued to work with a consortium of experts led by NWEA. While NWEA helped organize and manage the thought partners, representatives from the National Center for the Improvement of Educational Assessments, Johns Hopkins University, Great Minds, Odell Education, MZ Development, Strategic Measurement and Evaluation contributed to the design, administration, and scoring of this curriculum-connected assessment.

For the Math pilot Louisiana partnered with New Meridian during the 2022-2023 school year to design and administer math “testlets” for grades 5 and 7. Seven school systems from across the state volunteered to administer the assessments throughout the school year. The math pilot is designed to focus on the scope and sequence model, allowing schools to test students immediately after instruction. New Meridian brought together educators from Louisiana and Montana, also piloting the assessment, to help create items for the pilot.

LDOE continues to leverage these strong partnerships and research base to help envision how this model could be more widely adopted.. This project has helped conceptualize how to bridge the gap between Model 1 (curriculum-aligned assessments) and Model 2 (scope and sequence aligned assessments). At the center of the work is the belief that providing meaningful data, reports, and supports throughout the year gives educators crucial information to inform instruction. The LDOE continues to use what has been learned about student progress, instructional impact, reporting, and scaling from the through-year assessments in Louisiana to help create a vision for a cohesive grade 3-8 state system.
Summary of Feedback and Input from People Proximate to the Problem

The entire design of the ELA through-year assessment is centered on reaching students from underserved populations. By designing assessments that focus on knowledge gained through ELA classroom instruction, the LDOE hopes to remove barriers for students without the background knowledge to find success. Teachers and school leaders who serve these students were not only involved through passage and item reviews, but also through focus groups and discussions centered around the assessment design, reports, and instructional support materials.

This year, the ELA team has put in great effort to engage with the people most proximate to the problem and has been honored by the many voices that have contributed. Our focus to date has been mainly educators and students, but in late March the team was provided with the opportunity to engage with district coordinators and a parents/caregiver group.

Overall impacts are noted in Table 1, but those that were most defining included:

- Educator focus groups met to review test design sketches which influenced the final prototype design.
- Other educator focus groups and follow up sessions concentrated on class report designs.
- The student survey from Winter will support the main tenets for a redesigned individual student report.
- The student survey from spring 2023 will help the team better understand if student experiences across curricula are similar.
- New relationships were formed with the Parent and Family Engagement Coordinators which will influence future planning.

Plans for next school year also include engagement with the people most proximate to the problem through focus groups, surveys, and potentially cog labs. While the research questions are not fully defined at this time, it is expected that they will center on utility-based questions such as:

- Do teachers, students and parents find value in a curriculum relevant approach to state testing? If so, how and what?
- What kinds of information in the score reports do students, teachers, and parents/caregivers pay most and least attention to?
- Which interpretations are appropriate and which ones reflect misunderstandings about the assessment purpose and design?
- How valuable are the supports for further learning and/or instructional guidance provided alongside the reports to help students, teachers, and parents/caregivers? How effective are these actions for supporting actual improvement?
For the math pilot program, New Meridian engaged educators in empathy interviews to get initial input on the design of the assessments. Included in those educator groups were teachers of students who are impoverished, have learning differences, as well as teachers of students who are English language learners. Overall impacts of engagement activities for the math pilot are noted in Table 2 in the appendix. The impacts from this valuable work included the following:

- report design and categories
- report timing and structure
- reporting features
- student perceptions of items
- information on educator reporting needs

The LDOE also worked to ensure that the student demographics of school systems in the pilots was representative of the state of Louisiana’s student population. Through student, educator, and district surveys and focus groups, the program will continue to engage all stakeholders, especially those most proximate to the problem.
Summary of Pilot Tests

For the ELA pilot, during the 2022-2023 pilot administration year, the LDOE assessment and instructional teams met with national partners, educators, and school system leaders to finalize the blueprint for the assessments. At the heart of the work was an analysis of the various units of study across high quality curricula in order to determine how to create an assessment relevant to multiple curricula. The team also set out to continue honoring the background knowledge students gained from the units of study in their ELA classrooms. Through this design principle, this pilot attempts to remove the barriers often faced by historically disadvantaged student groups.

The resulting pilot forms used what we have labeled the funnel model. The model includes:

- unique, unit-based “Knowledge” sections
- a common thematically related warm read text and item set in the “Application” section, and
- a common writing prompt in the “Synthesis” section.

This model continues to honor the background knowledge taught in the classroom, but then “funnels” students into a common “Application” and “Synthesis” section. See an illustration of the design in Figure 1 in the Appendix.

While the original plan was to funnel three curricula – Guidebooks, Wit and Wisdom, and Expeditionary Learning (EL), the district using EL decided not to participate in the pilot this spring, leaving only Guidebooks and Wit and Wisdom. After many discussions with the team, the decision was to create this curriculum-relevant pilot form (Form C) for Guidebooks and Wit and Wisdom and spiral with the curriculum-aligned field test forms (Forms A and B) at the class level.

That earlier work with through-year assessments created some key efficiencies with incorporating this new pilot, such as:

- Integrating this pilot administrations as Form C of the ongoing Grade 5 field test
- Use of the same platform
- Use of the same technical support tools and Help Desk
- Use of the same manuals
- Use of the same communications

All forms were randomly assigned based on overall district curriculum choice – ELA Guidebooks or Wit & Wisdom. Table 3 in the appendix shows the number and percentage of students who participated in the pilot. The sampling design allowed the team to have the necessary counts for the field test and subsequent analyses (Forms A and B), but also provided a reasonable sample size for the new pilot form (Form C) with which to analyze performance on both the machine-scored items and the writing task.

For the spring 2023 pilot administration, the form for both curricula followed the design pattern as shown in Table 3. While results will not be shared with the schools or students from this curriculum-relevant pilot (i.e., score reports), results from the writing task and machine-scored items will be utilized to inform future curriculum-relevant pilot planning and thinking.
The key research questions the team will attempt to answer include the following:

- How does student performance on each section of the “funnel” (knowledge, application, synthesis) differ by students’ curriculum type and/or by student demographic category?
- How do the writing responses compare to trends on IAP rubric scores?
- Are we seeing differences in responses - compared to IAP and possibly compared across the pilot forms? Are they possibly attributed to test design?

To answer these questions, the team plans to analyze both the writing and the machine-scored components of the assessment. Due to the timelines associated with the memo, the team has determined both short- and long-term goals for the analyses.

Short-Term Goals:

- Hand-Scored: Understand how students are performing on the Writing Task on the curriculum- relevant assessment (Form C)
- Machine-Scored: Understand how students are performing on Form C

Long-Term Goals:

- Understand how performance on Form C compares to the performance on Form A and Form B

These findings will inform the development of future prompts and forms. For the machine-scored items, we will be looking at overall p-values and correlations to determine how students perform on each item, and how well the item differentiates between high and low achievers. From that review, adjustments will be made to items accordingly. While the initial analyses will look specifically at Form C, future analyses will look across the grade 5 forms in an attempt to answer if performance on the funnel design/curriculum-relevant design is similar to that of the curriculum-aligned Innovative Assessment Program.

The math pilot prototype consists of 12 individual testlets that each primarily assess a single content strand. These math testlets are evenly spread across four administration windows, with three testlets administered in each window. New Meridian designed items to assess different levels of cognitive depth within each strand. All students participating in the pilots received the same testlets. The total testing time for each administration window is estimated to be approximately 45 minutes (totaling three hours of testing time throughout the year). Table 4 lists the pilot school systems and number of students who participated.

The math testlets assess content that could reasonably be taught in a variety of different sequences. During the prototype The order of administration was predetermined during Phase Two due to technical and logistical considerations. An expanded pilot will include a configuration tool that will permit the custom ordering of math testlets at the district or school level.

The math pilot is restricted to machine-scored items. The future piloting will include at least one constructed response math performance task as part of a fifth administration, and the final design is still intended to include several performance task-based testlets. Future testlet piloting may also include several items designed to assess content that is below-grade level but is necessary for understanding on-grade content.
Since both pilots had forms administered in the spring window that ended on May 25, 2023, data analysis is still being conducted. For the math pilot, New Meridian is collecting more detailed scope-and-sequence materials from each school system to set the order of the testlets.

Similarly, the ELA pilot team continues to work toward understanding exactly what units and texts school systems are teaching. The LDOE does not collect this specific information from school systems, so the assessment team continues to explore how to get curriculum usage details and how to recruit systems who are using less popular ELA curricula. Both pilot programs will continue to recruit pilot school systems which will ensure the process provides reliable information on all student groups.

**Pilot Implementation and Work Plan**

The LDOE continues to share and collect information for both pilots through presentations, newsletters, office hours, and other communication outlets. Team members have also shared information about the pilots at the Teacher Leader Summit and through parent and family engagement coordinators. Recruitment questionnaires have been distributed through LDOE webinars and office hour presentations. Reporting webinars and monthly office hours sessions help to keep all ELA systems updated and engaged.

A weekly Innovative Assessment Program team meeting at LDOE includes team members from assessment content, accountability, reporting, and psychometrics meet each week along with the mathematics and ELA instructional team members. Members contribute to the agenda and support the program by giving valuable feedback and expertise to this complex project.

The LDOE continues to work toward funding these projects through additional revenues from a CGSA grant. No additional full-time staff members will be added for the 2023-2024 academic year. The math team is hoping to ease some of the strain from additional projects through hiring part-time workers. The IAP Program Coordinator’s role has shifted to include duties associated with the math program.

Both programs continue to engage school and district leaders and educators to gather feedback on the assessment design, administration, reporting, and timing of the pilots. Tables 1 and 2 in the Appendix outline many of the outreach activities. Additional examples can be found in Slide Decks 1 and 2 as an illustration of the kinds of educator focus groups that have been conducted. Student survey samples can also be found in the appendix.

Teams from both pilots continue to study timing data and assessment results as we move forward with plans for the coming year. Parent, educator, student, and district surveys have helped inform next steps in the assessment and report designs as well as instructional support documents for the ELA pilot. Through educator focus groups and interviews, the ELA team works to engage educators in assessment revisions and report design decisions. Both programs hope to conduct cognitive labs with students in the fall 2023 school year if proper state and district permissions can be obtained. A parent survey will aid the ELA team in understanding the caregiver perspective on reports.

The LDOE has offered professional development throughout the 2022-2023 school year and will continue to expand on this work in the coming year. Passage and item review committees are crucial to the design process and help educators understand the structure of the assessment. New Meridian and LDOE continue to work with teams of educators to design items and review standards alignment. Rangefinding and data review
committees also offer educators valuable training on rubrics and item performance. In addition to these activities, the LDOE ELA instructional and assessment teams have offered professional development on curriculum implementation and writing rubrics, scoring, and feedback. The LDOE is planning a series of professional development workshops next year especially for new systems.

For spring 2023 the ELA pilot assessment model worked well. A deeper dive into the curriculum allowed the team to find common themes across the units for ELA Guidebooks, Wit & Wisdom, and EL, allowing for unique items in the Knowledge section, and common items across the Application and Synthesis sections.

However, there are areas of concern with the current model.

- The inherent nature of the model (funnel) requires the ability to cluster like units together to ensure common themes with which to build the Application and Synthesis sections. Requiring a particular scope and sequence is not aligned with our philosophy of local choice. How does this impact local choice?
- Similarly, this model assumes a level of fidelity with the determined curricula. What happens with students that migrate into the state mid-year? What happens with students who are unable to attend school for crucial learning of the curriculum?
- The psychometric model, design, and alignment with the varied competing priorities will require additional study and decision making.

Based on these concerns, the team will focus on curriculum analysis, study additional prototype designs for review and discussion, and clarify its theory of action over the coming months. A key area for discussion will include the role of knowledge in the assessment, and how/where should that knowledge be built. The LDOE has continued to build the infrastructure and systems needed to support and sustain these through-year models. Establishing meeting protocols and communication plans with the many partners in this work will be key to the success of the programs. In turn, those same supports must be offered to school systems who are taking on these pilot programs. By offering help desk contacts, guidance documents, administration and report webinars, monthly office hours, and numerous avenues for support, the LDOE hopes to provide the structure and guidance systems need for program success.

**Lessons Learned**

Key to this work has been our engagement with the people most proximate to the problem. From the early focus groups on designs, to the collaboration with teachers on reports, to the student surveys and focus groups, what we have learned has provided us with opportunity to adjust along the way. It has also afforded the assessment and data teams at the state with the opportunity to engage with groups that they have not in the past. For example, the Be Engaged Initiative at the department has opened doors to parents and caregivers the department has not had previously. For the state of Louisiana, stakeholder engagement, and a true focus on the people most proximate to the problem, has been a new experience. An invaluable experience. The hope is that the relationships that have been built this year, will continue to grow.
As noted earlier some key impacts from our engagement have included:

- The educator focus groups that focused on the test design sketches. This feedback was used to finalize the prototype design.
- The educator focus groups and follow up that focused on class report designs.
- The student survey from Winter which will support the main tenants for a redesigned individual student report, and
- The student survey from spring 2023, which will help the team better understand if student experiences across curricula are similar.

The role of engaging with the people most proximate to the problem will continue to be a significant effort into the coming school year. A current highlight of engagement activities to come include:

- Parent/caregiver engagement on reports
- Collaboration with educators on support tools
- Continued engagement with students, as expert panels and through cog labs

Those most proximate to the problem are not all aligned behind this theory of action, or may be unwilling to risk foregoing current practices which result in more testing time. Schools worry about the impact of a “new test” on their accountability scores.

For the math program, we have experienced great difficulty with matching New Meridian’s testlet sequence with each system’s scope and sequence. We learned that good communication and planning are keys to success for any new assessment.

Without a doubt, through-year assessments can cause a strain on personnel at school, district, and state levels, especially if well-developed plans and good communication are not in place. Team members from both pilot programs are working to improve communication, meeting structures, clearly defined deliverables, and concrete goals and deadlines.

Communication is also key to growing any new assessment. Additional professional development will help ensure this assessment is used for its intended purpose. Additionally guidance and professional development in how to use the reports to inform instructional decisions. Through monthly professional development sessions in the coming year, the LDOE hopes to relieve some of the tension around these pilots.
Appendix

Figure 1: Illustration of the ELA Funnel Design

Table 1: ELA Pilot Engagement Activities to Date

<table>
<thead>
<tr>
<th>Action</th>
<th>Who</th>
<th>When</th>
<th>Purpose</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement with Teachers</td>
<td>Focus Group with 6 teachers</td>
<td>October 2022</td>
<td>Discuss End of Unit reporting</td>
<td>Initial information to create new way to organize information in Class Report</td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Focus groups/interviews of ~20 teachers</td>
<td>December 2022 – February 2023</td>
<td>High level feedback on test design sketches and reporting considerations</td>
<td>Supported the decision in using the Funnel model for the spring administration</td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Survey</td>
<td>December 2022 – April 2023</td>
<td>Analyze teachers’ classroom practices and utilization of instructional materials. The information gathered in this survey will demonstrate how ELA teachers in grades 3-8 understand and utilize their curricula</td>
<td>Better understand curriculum use in the state</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>---------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Collaboration Groups</td>
<td>February – March 2023</td>
<td>To support the work of a report redesign at the class level</td>
<td>Redefined class level report</td>
</tr>
<tr>
<td>Teacher Engagement</td>
<td>Collaboration Groups</td>
<td>May – July 2023</td>
<td>Receive feedback on redefined class level report</td>
<td>Create final adjustments for use in 2023-24</td>
</tr>
</tbody>
</table>

### Student Engagement

<table>
<thead>
<tr>
<th>Student Engagement</th>
<th>Survey</th>
<th>May 2022</th>
<th>Focused on engagement, demonstration of learning, assessment preferences, experience, empowerment, preparation, and results</th>
<th>Gain an understanding of students’ general thoughts regarding the innovative assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Engagement</td>
<td>Survey</td>
<td>January- February 2023</td>
<td>Feedback on reports and results; how they engage in learning more about reading and writing</td>
<td>Will support the creation of an updated Individual Student Report; will also be the basis for a Student Focus Groups</td>
</tr>
<tr>
<td>Student Engagement</td>
<td>Student Survey (Optional 5th grade survey)</td>
<td>May 2023</td>
<td>The focus will be on their experience with the assessment. Our proposed analysis is aimed at comparing and contrasting</td>
<td>Evidence to determine if the assessments related to the different curricula are</td>
</tr>
<tr>
<td>Student Engagement</td>
<td>Focus Groups</td>
<td>June 2023</td>
<td>To support the work of a report redesign at the individual student level</td>
<td>Redesigned individual student report for use in 2023-24</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>---------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
</tbody>
</table>

**District Engagement**

<table>
<thead>
<tr>
<th>District Engagement</th>
<th>Survey</th>
<th>January 2023</th>
<th>Collection of information regarding curriculum usage at the district level; gauging general interest level in participating in pilot.</th>
<th>Supported the creation of “Collaboration Groups” that supported the work on a report redesign</th>
</tr>
</thead>
</table>

**Parent Engagement**

<table>
<thead>
<tr>
<th>Parent Engagement</th>
<th>Presentation to the Parent and Family Engagement Coordinators (PAFE)</th>
<th>March 2023</th>
<th>To introduce the PAFE team to state assessment in Louisiana, and discuss possible opportunities to partner and next steps</th>
<th>New relationships created; intend to utilize interested coordinators in supporting reviews of reports.</th>
</tr>
</thead>
</table>
Table 2: Math Pilot Engagement Activities to Date

<table>
<thead>
<tr>
<th>Engagement Activity</th>
<th>Key Question Addressed</th>
<th>Impact of Activity on Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus Groups (Summer 2022)</td>
<td>Does the data collected via the IAAS have a positive impact on educators’ instructional decisions?</td>
<td>This was the first feedback that New Meridian received on potential report designs and categories. It informed the baseline for the cross-functional team to begin the iterative report development and refinement process.</td>
</tr>
<tr>
<td>Empathy Interviews (Summer 2022)</td>
<td>Does the data collected via the IAAS have a positive impact on educators’ instructional decisions?</td>
<td>Used to refine user personas and further support many of the design decisions that New Meridian made in Phase One. These were critical in informing decisions on report timing and the underlying IT infrastructure needed to support it.</td>
</tr>
<tr>
<td>Educator Post-Admin Surveys (Throughout the administration year)</td>
<td>Does the data collected via the IAAS have a positive impact on educators’ instructional decisions?</td>
<td>Encouraged a more rapid development of the configuration tool so that New Meridian can include a beta version in an expanded pilot in the next academic year. Informed the expanded reporting feature that New Meridian intends to release in the expanded pilot during the next academic year.</td>
</tr>
<tr>
<td>Student Post-Admin Surveys (After second and fourth administrations)</td>
<td>Does the data collected via the IAAS have a positive impact on educators’ instructional decisions?</td>
<td>These provided student-level perspectives on the items and forms themselves. They provided context for timing and performance data that New Meridian used during the recalibration of the testlets following the final administration in the pilot year.</td>
</tr>
<tr>
<td>Focus Groups (Spring 2023)</td>
<td>Does the data collected via the IAAS have a positive impact on educators’ instructional decisions?</td>
<td>This was the most recent set of feedback on the prototype configurator tool and reports. The feedback was critical in determining what information educators needed/wanted to see when scheduling testlets. It also provided</td>
</tr>
</tbody>
</table>
invaluable insight into the reports that educators need to see (e.g., at the classroom level), and they provided great feedback on how New Meridian should change the prototype summative score predictive reporting.

Table 3: ELA Pilot Administration Details

<table>
<thead>
<tr>
<th>Administration</th>
<th>Total Actual Count</th>
<th>Form A #</th>
<th>Form A %</th>
<th>Form B #</th>
<th>Form B %</th>
<th>Form C #</th>
<th>Form C %</th>
</tr>
</thead>
<tbody>
<tr>
<td>W&amp;W</td>
<td>1,413</td>
<td>608</td>
<td>43</td>
<td>608</td>
<td>43</td>
<td>198</td>
<td>14</td>
</tr>
<tr>
<td>Guidebooks</td>
<td>2,422</td>
<td>1,114</td>
<td>46</td>
<td>1,114</td>
<td>46</td>
<td>194</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 4: Math Pilot School Systems

<table>
<thead>
<tr>
<th>District</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claiborne</td>
<td>34</td>
</tr>
<tr>
<td>Iberia Parish</td>
<td>492</td>
</tr>
<tr>
<td>Iberville Parish</td>
<td>69</td>
</tr>
<tr>
<td>St. Bernard Parish</td>
<td>349</td>
</tr>
<tr>
<td>Vermilion Parish</td>
<td>142</td>
</tr>
<tr>
<td>Vernon Parish</td>
<td>508</td>
</tr>
<tr>
<td>Winn Parish</td>
<td>245</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>1839</strong></td>
</tr>
</tbody>
</table>
Table 5: ELA Pilot Design

<table>
<thead>
<tr>
<th>Reporting Category</th>
<th>ELA Guidebooks: Shutting out the Sky, Form C</th>
<th>Wit and Wisdom: A War Between Us, Form C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Unit text pairings; unique items</td>
<td>Module text pairings; unique items</td>
</tr>
<tr>
<td>Application</td>
<td>Common Unit/Module related excerpt(s); <strong>common passage(s) and items</strong> for all curricula</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>Common Opinion Writing Task (WT) based on common warm read pairing (from the Application section) and one unit/module text</td>
<td></td>
</tr>
</tbody>
</table>

**Survey Exhibits**

- Student Survey Sample 1
- Student Survey Sample 2
Student Survey Sample 1

Window 2 | Winter 2023 Student Survey  
Target Audience: Grades 6-8 Students

Since you were in third grade, you and your classmates have likely taken the state English language arts (ELA) assessments called the Louisiana Educational Assessment Program (or LEAP 2025), which is given at the end of the school year.

This school year you are participating in the Innovative Assessment Program (IAP) ELA Guidebooks Assessment that is given to students like you three times during the school year in the Fall, Winter, and Spring after your class has completed an ELA Guidebooks unit. The following questions ask about your experience regarding the end-of-unit IAP ELA Guidebooks Assessment.

The following questions ask about your experience regarding assessments. We are interested in better understanding how we can use these tests to add value to your education by learning how you plan to use the results.

<table>
<thead>
<tr>
<th>Q1. Part A What do you want to know about how you did on this test? [Select all that apply.]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Options:</strong></td>
</tr>
<tr>
<td>• My overall performance</td>
</tr>
<tr>
<td>• My performance on each part of the test</td>
</tr>
<tr>
<td>• My performance compared to students in my school</td>
</tr>
<tr>
<td>• My performance compared to students in other schools</td>
</tr>
<tr>
<td>• I want feedback on my writing</td>
</tr>
<tr>
<td>• I want feedback on my reading</td>
</tr>
<tr>
<td>• Other</td>
</tr>
</tbody>
</table>

Q1. Part B If you answered “Other” above, please explain in the space below.

[Open Text]

<table>
<thead>
<tr>
<th>Q2. Part A What do you want your family or caregivers to know about how you did on this test? [Select all that apply.]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Options:</strong></td>
</tr>
<tr>
<td>• My overall performance</td>
</tr>
<tr>
<td>• My performance on each part of the test</td>
</tr>
<tr>
<td>• My performance compared to students in my school</td>
</tr>
<tr>
<td>• My performance compared to students in other schools</td>
</tr>
<tr>
<td>• I want feedback on my reading shared with them.</td>
</tr>
<tr>
<td>• I want feedback on my writing shared with them.</td>
</tr>
<tr>
<td>• Other</td>
</tr>
</tbody>
</table>

Q2. Part B If you answered “Other” above, please explain in the space below.

[Open Text]

<table>
<thead>
<tr>
<th>Q3. Part A After taking this test, what would you like your teacher to do? [Select all that apply.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Explain what I can do to be a better reader and writer?</td>
</tr>
</tbody>
</table>
Q3. B If you answered “Other” above, please explain in the space below.
[Open Text]

Q4. Earlier this school year you may have taken a similar test on your first ELA Guidebooks Unit.

Q4. Part A. Were your results on that test shared with you?
[Select all that apply.]
Options:
- Yes. I received a report with my results.
- Yes. My teacher went over my results with me.
- Yes. My teacher went over the results with the class together.
- No. I didn’t take a test earlier this year on my first Guidebooks unit.
- No. My results were not shared with me.
- Other

Q4 Part B. If you answered “Other” in Question 8, please explain in the space below.
[Open Text]

Q5. How important is it that you know how you performed on this test?
[Choose only one.]
Options:
- A. It is very important to me.
- B. It is somewhat important to me.
- C. I’m neutral on this either way.
- D. It is somewhat unimportant to me.
- E. It is not at all important to me.

Your Experiences with Reading and Writing

Q6. Part A What do you know about how you can become a better reader?
Options:
- I have a few general ideas on how to become a better reader.
- I have a clear plan of what I need to do to become a better reader.
- I would like more information from my ELA Teacher about what I can work on to become a better reader.
- I would like my ELA Teacher to ask me about my ideas on how I can become a better reader.
- I want someone to ask me what topics I am interested in to become a better reader.
Q6. Part B If you answered “Other” above, please explain in the space below.
[Open Text]

Q7. Part A What do you know about how you can become a better writer?
[Select all that apply.]

Options:
- I have a few general ideas on how to become a better writer.
- I have a clear plan of what I need to do to become a better writer.
- I would like more information from my ELA Teacher about what I can work on to become a better writer.
- I would like my ELA Teacher to ask me about my ideas on how I can become a better writer.
- I want someone to ask me about topics I am interested in to become a better writer.
- I don’t really think about ways that I can become a better writer.
- I am not quite sure what to do in order to become a better writer.

Q7. Part B If you answered “Other” above, please explain in the space below.
[Open Text]
Survey Questions for Grade 5 Administration—Spring 2023

Introduction

This survey asks you a few questions about a test you recently took. This test was about what you just learned in class during the module A War Between Us. This survey should take no more than 5 minutes to complete.

Your answers will not be shared with your teacher and they do not affect your grade. They only help us make the assessment better. Thanks so much for taking this survey!

Questions

The test you took was on what you learned in the module A War Between Us.

During the module A War Between Us, you read a lot of texts in class. The test asked you about some of these texts.

1. How well did you understand the module texts? [Very well, Somewhat, Not very well at all]

2. How difficult was it to answer the test questions about the module texts? [Not difficult at all, Somewhat difficult, Very difficult]

3. How hard did you try to do well on the test questions? [Very hard, Somewhat, Not very hard]

4. How confident are you that you did well on these questions? [Very confident, Somewhat confident, Not very confident]

After answering test questions about what you read in class, you then read two new texts.

5. How well did you understand these new texts? [Very well, Somewhat, Not very well at all]

6. How difficult was it to answer the questions about these new texts? [Not difficult at all, Somewhat difficult, Very difficult]

7. How hard did you try to do well on these questions? [Very hard, Somewhat, Not very hard]

8. How confident are you that you did well on the test questions? [Very Confident, Somewhat confident, Not very confident]

Finally, you were asked to write an essay as part of the test.

9. How difficult was it to write the essay? [Not difficult, Somewhat difficult, Very difficult]
10. How familiar are you with this kind of writing? [Very familiar, Somewhat familiar, Not very familiar]

11. How hard did you try to write a good essay? [Very hard, Somewhat, Not very hard]

12. How confident are you that you wrote a good essay? [Very confident, Somewhat confident, Not very confident]

**General Questions about You**

13. How often do you read outside of school? [All the time, Often, Sometimes, Not very often, Never]

14. How often do you write outside of school? [All the time, Often, Sometimes, Not very often, Never]
Slide Deck Exhibits

- Slide Deck 1–Focus Group Slides–ELA Design (pages 21-25)
- Slide Deck 2–Focus Group Slides–ELA Reports (pages 26-29)
- Slide Deck 3–ELA Professional Development (pages 30-37)

Slide Deck 1- Focus Group Slides - ELA Design

Collecting Feedback on a Possible State Test Design in Louisiana
Focus Group
January 4, 2023

Agenda
1. Introduction & Overview (15 minutes)
2. Introducing new test design possibilities, and your feedback (45 minutes)
3. Wrap-up (15 minutes)

Goal: To hear from you about a possible design for the IAP for other curriculums.

1. Introduction & Overview

Introductions
- Audra Koch, NIMDA Facilitator
- Carl Davis, OME Educator Facilitator
- Ruth Collof, LDDE Observer
- David Hopkins, LDDE Observer

Your Turn!
- Name
- School
- Experience with the Innovative Assessment Program

Norms
- We want your honest opinion.
- If an idea is great, please tell us.
- If an idea is awful, please tell us.
- And please tell us all of your other thoughts and reactions.
Recording & Notetaking
- We will be recording this session as well as using Microsoft Teams' automatic transcription feature.
- If you are uncomfortable being recorded, please let me know and I will suspend the recording and rely on transcription alone.

On Materials
- Today we will be talking about ideas that are currently under development.
  - We ask that you not share these ideas and the supporting materials with those outside of this meeting.
  - We are working hard to make the program the best it can be, so things can and will change.

The IAP, Next Steps, and Today's Meeting

2. Introducing New Test Design Possibilities

Test Design Option 1
Introducing a New Test Design Possibility

- Current S.A.P. (Standard Alignment to Specific Curricula) (ELA Guidebooks and VM and include)
- This approach makes connections to topics students are familiar with across the curriculum (ELA, Social Studies, Sciences)
- Covers ELA reading and writing standards
- Intentional building of knowledge through literary and informational texts on the same topic
- Supports cross-curricular connections

Possible Design Elements

- Thematic year Instead of 3 times a year
- Focus content-based topics related to social studies (geography, literature) for analysis, narratives, topics related to science (exposition)
- Each text block
  - Includes ELA reading and writing standards
  - Includes both informational and literary texts on the same topic
- Knowledge Section: Informational text
- Application Section: Literary text
- Synthesis Section: Writing task that asks students to write about the texts provided on the test

Examples of Topically Related Cross-Curriculum Tests

Social Studies - Content-based topics related to social studies
- Content-related reading and writing standards
- Thematic year Instead of 3 times a year
- Focus on content-based topics related to social studies
- Application Section: Literary text
- Synthesis Section: Writing task that asks students to write about the texts provided on the test

Clarification Questions?

3. Group Conversation

What benefits or concerns do you see with this design?
Question #2

Would you want your students assessed with this test design? Why or why not?

Test Design Option 2

Introducing Another Test Design Possibility

Each through course test would have meta-data:
Each test would include:
- a short essay test
- a concept knowledge section with an excerpt from a specific curriculum unit test and a set of questions
- a concept application section with a common test (and set of questions) that is also conceptually aligned with the knowledge section across multiple units
- a concept synthesis section with a writing test that requires students to use information from both the concept knowledge section and the concept application section

<table>
<thead>
<tr>
<th>Section</th>
<th>Curriculum A Test</th>
<th>Curriculum B Test</th>
<th>Curriculum C Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Except from unit test &amp; question set</td>
<td>Except from unit test &amp; question set</td>
<td>Except from unit test &amp; question set</td>
</tr>
<tr>
<td>Application</td>
<td>Except from theoretically aligned test and question set (the same test and question set would appear on all three tests)</td>
<td>Except from unit test &amp; question set</td>
<td>Except from unit test &amp; question set</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Writing task based on both the unit test (knowledge section) and the new test (application section)</td>
<td>Except from unit test &amp; question set</td>
<td>Except from unit test &amp; question set</td>
</tr>
</tbody>
</table>

Clarification Questions?

3. Group Conversation
Question #1
What benefits or concerns do you see with this design?

Question #2
Would you want your students assessed with this test design? Why or why not?

4. Wrap Up

Next Steps
We are looking to continue to engage around the next steps for the KRP, in terms of:
- Providing feedback on the design of a new assessment for other curricula.
- Providing feedback on the design of current and possible end-of-unit reports and instructional supports.
- Engaging with students and parents in your schools and districts on the Innovative Assessment Program.

Please complete this exit survey before leaving.
Slide Deck 2- Focus Group Slides - ELA Reports

Open Feedback Session: IAP Collaboration Groups
Score Reporting Prototypes
March 29, 2023

Introductions

- Assessment Development Partners
  - Nathan Davis, Center for Assessment, Evaluation, and Data
  - Audra Kost, NAMPA, Role Talker

- State Assessment & Accountability
  - Content Experts
    - Shatrini Lee, LDDE
    - Erin Huskey, LDDE
  - Observers
    - David Hawkins, LDDE
    - Pam Colford, LDDE
    - Kathy Ayala, LDDE
    - Jennifer Baird, LDDE

Recording & Note-taking

- We will be recording this session as well as using Zoom's automatic transcription service.
- If you are uncomfortable being recorded, please message me privately during the call, and I will suspend the recording and rely on transcription alone.

Context

The department is developing the ELA in a committee.
- Developing the IAP in collaboration with Districts, Administrators, and teachers.
- Today's meeting is part of that commitment.

Today's Meeting

- LDDE and its partners are working to improve score reporting at the end of the school year
- We've developed a number of "proof of concept" score reports that are wrapped up in a single HTML file that you can open with your browser.
- Goal:
  - Introduce you to these reports and get your initial feedback
  - Find the areas we need to do more work

Caveats

These reports reflect early and experimental thinking:
- These reports have not been subjected to the extensive quality assurance procedures that all other products have.
- It may take several areas to implement.
- These reports are based on the state's color codes provided in the IAP document.
- While they should reflect the student's prior classroom, they may not perfectly line up with current classroom assessment.
- These reports are not intended to be a report card but may be used to inform your planning instruction.

Caveats, Part 2

These reports reflect early and experimental thinking:
- What is shown in these reports may not be in the final report.
- Please use the reports and support their integration in your planning instruction.
- Today, we are focusing on classroom reports for teachers.
- Individual student reports are ready for development.
Technology Check

- You should have received an email invite from a website called [name] which will
  guide you to set up a password to login and download the files.
- There should be one file per classroom.
- Each file can be opened in your browser of choice and explored just like a webpage.

If you have issues, please chat or direct chat, Nathan.

Meeting Structure

- Individual Review: Take about 15 minutes to explore the reports.
- Write down questions, ideas, comments - whatever comes to mind.
- Take your hand on more often.

Whole Group Conversation & Feedback

Whole Group Conversation & Feedback

1. Did you find anything confusing, unclear, etc.?
   - Do you have any questions on these?
2. In general, are the reports useful?
   - Can any of these results inform your instruction in your current instructional unit? If so, how?

Whole Group Conversation & Feedback: Guidance Page

- What is useful in this section? Not useful?
- How easy, or hard, is it to look through the knowledge, application, and synthesis sections?
- Do you see any sections that need more information?
- Do these: you would change, add or remove?

Whole Group Conversation & Feedback: Guidance Page

- What is useful in this section? Not useful?
- How easy, or hard, is it to look through the knowledge, application, and synthesis sections?
- Do you see any sections that need more information?
- Do these: you would change, add or remove?
Meeting Structure

Whole Group Conversation & Feedback
- What report or name do you find useful? Why?
- Are any of the results surprising? Are any of the results in line with your expectations?
- Can any of these results inform your instruction in the behind the scenes unit?
- If so, how?
- What do you wish you had in these reports? How would it be useful?

Next Steps

We will continue to improve the score reports and hope to have a follow up with you at the end of your institution on the current unit.

Please take a few minutes to fill out the web survey to let us follow up with you.
https://www.surveymonkey.com/s/FinalNotes2023
Slide Deck 3- ELA Professional Development

Innovative Assessment Program
Assumption Parish
Professional Development
May 24, 2023

Presenters
- Kathy Judy, Supervisor of ELA Assessment Content
  - kathy.judy@la.doe
- Barbara Holland, ELA Assessment Coordinator
  - Barbara.Holland@la.doe
- Ruth Callahan, Innovative Program Assessment Coordinator
  - Ruth.Callahan@la.doe

Agenda
- Assessment: Overview
  - Background of Program
  - Assessment Design
  - Reports and Supports
- Preparing for 2023-2024
- Resources and Professional Development
- Contact Information/Questions

Assessment Overview

The Need for Innovative ELA Assessments

Why Did Louisiana Begin This Journey?

Research supports the importance of knowledge when reading.
- “Children’s success in reading comprehension depends heavily on knowledge. By failing to provide a solid foundation in basic subjects we inadvertently hobble children’s ability in reading comprehension.”
  - Why Do Some Students Read Better at Reading in 30 Years

- “Whether or not readers understand a text depends more on how much background knowledge and vocabulary they have relating to the topic than on how much they’ve practiced comprehension skills.”
  - Why American Students Read Better at Reading in 30 Years
Background of Program

- Louisiana submitted a proposal for the program in April 2012 as a response to the George W. Bush Foundation’s request for innovative assessment systems. Louisiana’s submission was selected as one of the four states to participate in the Bush Foundation’s Innovation in Assessment (BIA) Program.

Louisiana’s Goals and Priorities

**SIX CRITICAL GOALS**

1. **Student-centered curriculum:** Students will use student-centered curricula that are aligned to the Common Core State Standards.
2. **Data-driven instruction:** Teachers will use data to improve instruction and student outcomes.
3. **Assessment for learning:** Students will receive feedback on their learning progress.
4. **Professional development:** Teachers will receive ongoing professional development.
5. **Technology integration:** Schools will integrate technology into instruction and assessment.
6. **Parent and community engagement:** Families and communities will be involved in the education process.

**EDUCATIONAL PRIORITIES**

1. **Equitable access:** Ensure every student has access to high-quality education, no matter their background.
2. **Strong teacher-leaders:** Ensure every student has a qualified teacher in every classroom.
3. **High-quality student assessments:** Ensure every student has access to high-quality assessments.
4. **Continuous improvement:** Ensure every student has access to continuous improvement opportunities.
5. **Rigorous standards:** Ensure every student has access to rigorous standards.
6. **Engaged and informed communities:** Ensure every student has access to engaged and informed communities.

Louisiana’s Academic Strategy

- **ENSURE COHESION AND QUALITY:** Align standards and assessments with the highest quality.

- **BUILD TRUST IN THE FIELD:** Build trust through ongoing support, collaboration, and communication among school systems, principals, and teachers.

- **FACILITATE STRATEGIC PARTNERSHIPS:** Facilitate partnerships between school systems and education partners to ensure teachers and students have access to instructional materials and professional development.

Louisiana’s Theory of Action and Strategy

- **Louisiana believes in ensuring coherence by aligning student achievement with curriculum, instruction, and assessment.**

- **Louisiana is working to support high-quality classroom instruction.**

- The Louisiana Assessment Program is designed as a foundation to support student learning and teachers.

- The curriculum and instruction is designed to build on the knowledge built in the units/modules of study.

- The program includes feedback through the implementation of this comprehensive assessment system.

National Partnerships

Louisiana partnered with national experts and school systems to build and field an innovative assessment system that ensures all students have an opportunity to succeed.

- **NWEA**
- **Edvision**
- **Millsaps College**
- **Mississippi School of the Arts**
- **Mississippi State University**
- **University of Southern Mississippi**
- **The National Center for the Improvement of Educational Assessments**
- **KEND**
- **Georgetown University**
- **University of Virginia**
- **Armed Services Vocational Aptitude Battery**

And Louisiana continues to build and field innovative assessment tools and new opportunities for partnerships have emerged through Germs and Bistro.

Who Are Our Stakeholders?

- Parents
- Teachers
- Students
2022-2023 IAP School System Partners
- Burnet Heights Elementary
- Catalina Heights Elementary
- Carol Stream Elementary
- Carol Stream Intermediate
- Carol Stream Junior High
- Carol Stream High School
- Carol Stream North
- Carol Stream South
- Carol Stream West
- Carol Stream East
- Carol Stream Central
- Carol Stream North Intermediate
- Carol Stream South Intermediate
- Carol Stream West Intermediate
- Carol Stream East Intermediate
- Carol Stream Central Intermediate
- Carol Stream North Junior High
- Carol Stream South Junior High
- Carol Stream West Junior High
- Carol Stream East Junior High
- Carol Stream Central Junior High
- Carol Stream North Senior High
- Carol Stream South Senior High
- Carol Stream West Senior High
- Carol Stream East Senior High
- Carol Stream Central Senior High

2022-2023 School System Partners and Engagement
- 25 School Systems—over 300,000 kids, 33,000 students
- How we are engaging stakeholders:
  - Student surveys
  - Faculty surveys
  - Student surveys
  - Parent surveys
  - Staff surveys
  - Stakeholder engagement
  - Data analysis
  - Professional development
  - Professional development

Educator Feedback
- 65% of the educators who responded to the spring 2022 survey said that they felt very confident about their students doing well on the IAP.
- “I believe the students were more comfortable and confident taking the assessment than they were with the test.”
- “I enjoyed the smaller test windows. Giving the students four weeks to complete the assessment was beneficial.”
- “I think that students who were tested in the fall, rather than at the end of the year.”
- “The students felt more confident with the test content.”
- “By administering the test earlier in the district, I believe that students will feel more prepared in completing the assessments. They will have the opportunity to work at their own pace.”

Student Feedback
- 65% of the students who responded to the spring 2022 survey said that they felt very confident about their students doing well on the IAP.
- “I believe the students were more comfortable and confident taking the assessment than they were with the test.”
- “I enjoyed the smaller test windows. Giving the students four weeks to complete the assessment was beneficial.”
- “I think that students who were tested in the fall, rather than at the end of the year.”
- “The students felt more confident with the test content.”
- “By administering the test earlier in the district, I believe that students will feel more prepared in completing the assessments. They will have the opportunity to work at their own pace.”

ELA Innovative Assessment Program Highlights
- How the assessment is structured:
  - Structure: Several assessment structures are designed to measure student knowledge of skills and concepts they have studied in class.
  - Flexibility: Schools systems continue to decide which skills are used for instruction and which assessments students take.

ELA Innovative Assessment Program Highlights
- Integration: Several brief assessments are conducted throughout the year to measure student understanding of skills and concepts they have studied in class with a real-time integration of instruction and assessment.
- Flexibility: Teachers can focus instruction on background knowledge and student understanding.
- Accessibility: All students have the opportunity to develop shared background knowledge based on individual needs.
- Flexibility: Schools systems continue to decide which skills are used for instruction and which assessments students take.
Operational Administrations (Grades 6-8)

<table>
<thead>
<tr>
<th>TEST SECTION</th>
<th>TOTAL OF ITEMS</th>
<th>TIMING</th>
<th>TEST ADMINISTRATION</th>
<th>TEST INSTRUCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1</td>
<td>60 questions</td>
<td>60 min</td>
<td>2 people</td>
<td>Read the passage.</td>
</tr>
<tr>
<td>Section 2</td>
<td>60 questions</td>
<td>60 min</td>
<td>2 people</td>
<td>Write a response.</td>
</tr>
<tr>
<td>Section 3</td>
<td>60 questions</td>
<td>60 min</td>
<td>2 people</td>
<td>Write a response.</td>
</tr>
</tbody>
</table>

What Kind of Writing Is Required on the Assessment?

- Students respond to two different writing prompts for each unit assessment.
- Constructed Response—section 1
  - Students respond to a constructed response question that allows them to show their understanding of a new text(s) and how the new text(s) connect to the knowledge of the unit.
- Essay Prompt—section 2
  - Students respond to a two-prompt that requires students to use the knowledge they have built in the text and the new text(s) in section 3 of the assessment. In order for students to demonstrate synthesis and expression of knowledge across texts, students are asked to include relevant information from three texts including the new text.

General Rubric for Constructed Response Items

<table>
<thead>
<tr>
<th>SCORE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>The response is a clear and coherent explanation of ideas supported by relevant details and sufficient evidence from the text.</td>
</tr>
<tr>
<td>3</td>
<td>The response is a clear and coherent explanation of ideas supported by relevant details and some evidence from the text.</td>
</tr>
<tr>
<td>2</td>
<td>The response is a clear but less coherent explanation of ideas supported by relevant details and limited evidence from the text.</td>
</tr>
<tr>
<td>1</td>
<td>The response is a less clear explanation of ideas supported by relevant details and minimal evidence from the text.</td>
</tr>
<tr>
<td>0</td>
<td>The response lacks clarity, interest, or contains insufficient information to demonstrate comprehension.</td>
</tr>
</tbody>
</table>

Rubric for End-of-Unit Essay:
Reading Comprehension and Written Expression

This rubric provides a detailed framework for evaluating the end-of-unit essay, ensuring that students not only demonstrate comprehension but also effectively synthesize and express their knowledge across texts.
Rubric for End-of-Unit Essay: Knowledge and Use of Language Conventions

A Different Type of Reporting

The LEAP ELA Guidelines assessment provides the following unique opportunities for reporting information about student progress:

- **Frequency**: Although reports are not provided at the end of each test, they are provided at the end of the year to provide information to schools.
- **Detailed Reports**: Reports provide information on student performance on exams and performance in each scoring category.
- **Comparative Reports**: Reports provide averages for comparison by student, class, school, school system, and LEAP average. The LEAP average is the average score of other students across the state who participated in the same unit assessment in that particular year.

End-of-Unit Reports

- **End-of-Unit Reports are sent approximately 4 weeks after the test window ends.**
- **Section One (selected response and constructed responses)**
  - Unit Tests ("Hot tests"), measure knowledge of Unit Module.
  - Unit-Based Tests ("Warm tests"), measure Application of Unit Knowledge.
- **Section Two (writing prompt)**
  - Show student's performance of hot and warm results.
  - Assumes implementation of guidelines within a unit and across tests.

Reports and Guidance Documents

- **Guidance for School Reports**
  - Provides instructional suggestions.
- **Student Score Report**
  - Provides scores for all reporting categories.
  - School, system, LEAP averages.
- **Guidance for Score Reports**
  - Explains the assessment and report categories.

End-of-Year Reports

The LEAP ELA Guidelines include the following:

- **Student scale scores and achievement levels.**
- **School-wide average and the LEAP average**
- **Description of each student's performance by reporting category:**
  - Knowledge of Unit Tests.
  - Application of Unit Knowledge.
  - Synthesis and Expression of Knowledge Across Tests.

*The LEAP average is the average scale score and achievement level for students across the state who participated in the same assessment.*
ELA Innovative Assessment Program 2023-2024

EDC has been working to pilot the Innovative Assessment Program to include additional grades and curricula. For 2023-2024, we are planning to administer the following assessments:

- LEAP ELA Grade 10 Assessment: Grades 5, 6, 7, and 9 Operational
- LEAP ELA Grade 12 Assessment: Grades 5 Operational
- LEAP ELA Grade 12 Assessment: Grades 3 and 4 Field Tests
- LEAP ELAP-OL Grade 5 Field Tests

General Guidelines for Operational Assessments

- Schools may choose to participate in the LEAP-stimulated assessment at one or more of the grade levels: 5, 6, 7, 8.
- Schools may choose to participate in the LEAP-stimulated assessments at one or more of the grade levels: 5, 6, 7, 8.
- Schools may choose to participate in the LEAP-stimulated assessments at one or more of the grade levels: 5, 6, 7, 8.
- Schools may choose to participate in the LEAP-stimulated assessments at one or more of the grade levels: 5, 6, 7, 8.

2023-2024 IAP Administration Windows

- Window 1 (Fall)
  - October 23-November 3, 2023
- Window 2 (Winter)
  - January 29-February 9, 2024
- Window 3 (Spring)
  - April 15-May 17, 2024

General Guidelines for Field Testing

- Individual schools may participate in IAP-stimulated testing if the entire school system is not.
- Systems may participate in IAP-stimulated testing if the entire school system is not.
- All schools that choose to participate in IAP-stimulated testing should be included in professional development throughout the year.
Grades 6-8 ELA Guidebooks
IAP Unit Progression Plan 2023-2024
• Grade 6
  0. Window 2: Island
  0. Window 3: The Bluff or Out of the Box
  0. Window 2: The Bluff or Out of the Box
  0. Window 3: The Bluff or Out of the Box
• Grade 7
  0. Window 1: The King
  0. Window 2A: The King or Out of the Box
  0. Window 2A: The King or Out of the Box
  0. Window 3: The King or Out of the Box
• Grade 8
  0. Window 2: The King
  0. Window 1: The King
  0. Window 3: The King

Grades 5 ELA Operational Assessments 2023-2024
• ELA Guidebooks (2022) Unit Progression
  0. Window 1: The Lion, the Witch, and the Wardrobe
  0. Window 2: The Lion, the Witch, and the Wardrobe
  0. Window 3: The Lion, the Witch, and the Wardrobe
  0. Window 4: The Lion, the Witch, and the Wardrobe
  0. Window 5: The Lion, the Witch, and the Wardrobe
  0. Window 6: The Lion, the Witch, and the Wardrobe

Note: Because we did not have enough students field test the Scaredie and Wanderloot assessment, we are unable to administer those assessments operationally next year. If you are interested in helping field test those assessments for us, please let us know.

IAP Field Testing 2023-2024
• ELA Guidebooks Grade 3 (2022)
  0. Window 1: The Tortoise and the Hare
  0. Window 2: The Tortoise and the Hare
  0. Window 3: The Tortoise and the Hare
• ELA Guidebooks Grade 4 (2022)
  0. Window 1: The Tortoise and the Hare
  0. Window 2: The Tortoise and the Hare
  0. Window 3: The Tortoise and the Hare
  0. Window 4: The Tortoise and the Hare

ELA Innovative Assessment Program
School System Interest Survey 2023-2024
The ELA Innovative Assessment Program will be operational for grades 5-8 next year. Please complete this survey if your system is interested in participating in any of the following through-year assessments:
• Grades 5-8 ELA Guidebooks Operational Assessments
• Grades 5-8 ELA Guidebooks Field Tests
• Grades 5-8 ELA Guidebooks Field Tests

This interest survey is due by May 31st. Systems will be asked to confirm participation and submit a proposal with superintendent approval in the Fall. For more information, contact Paula.Collins@FkSD.org.

Grade 5 IAP-CR Pilot Assessment
The LESE is working to develop ELA through-year assessments relevant to any high quality ELA curriculum.
• The pilot assessments will be administered during three windows in 2023-2024.
• Systems that are interested in helping to design the assessment and administration of the pilot should complete this survey.

Upcoming Dates
Resources and Support
Professional Development Opportunities

- AP Grade 9 ELA Guidebooks Unit 10: Reading/WRiting/Essay Review (July 17-21, 1 p.
- AP Grade 9 ELA Guide and Study Guide: Reading/Data Review (July 17-21, 1 p.)
- AP Grade 9 ELA Guidebooks Unit Review (July 20-27, 1 p., 1 p.
- AP Grade 9 ELA Guidebooks Unit Review (July 20-27, 1 p., 1 p.
- Teacher Leader Summit - 4/14 to 5/10 June 1

IAP Resources

- IAP Grade 9 ELA Guidebooks Unit Review (July 17-21, 1 p.)
- IAP Guide and Study Guide: Reading/Data Review (July 17-21, 1 p.)
- IAP Grade 9 ELA Guidebooks Unit Review (July 20-27, 1 p., 1 p.
- IAP Grade 9 ELA Guidebooks Unit Review (July 20-27, 1 p., 1 p.
- Teacher Leader Summit - 4/14 to 5/10 June 1

LDOE Resources

- LDOE Guidebook: offers comprehensive information about implementing and implementing the standards.
- LDOE Guidebook: guides for planning and implementing quality ELA instruction, which include a variety of instructional strategies.
- LDOE Guidebook: a whole-class curriculum made by teachers for teachers that ensures all students can read, understand, and express their understanding of complex, grade-level texts.
- LDOE Guidebook: contains a collection of resources for supporting English learners.
- Assessment Library: a guide to all assessment materials.

More Resources

- Great Minds Resources: with resources and videos for Wit and Wisdom IAP support.
- ADAM (Online Testing System)
  - Online Tools: provides a variety of activities for students to practice using the online tools to prepare for the test.
  - Online Testing Platform: provides a variety of activities for students to practice using the online testing platform.
  - Online Testing Platform: resources can be downloaded from this site; see Resources and Documents.

Questions

- Please post any questions to the chat or unMinute to ask a question.
- Email: ldoe.aip@la.gov or assessment@la.gov with any Rhee Assessment Program questions.
- Email: czechowski@la.gov with any ELA Guidebooks curricula questions.
Overall Guidance

Introduction

The Word Play unit assessment, like all unit assessments in the IAP, measures how well students understood what they read in during the unit and how well they can apply and build on that knowledge. Each of the three sections of the test - (1) knowledge, (2) application, and (3) synthesis - has a specific purpose, and each is meant to help you better understand where students are at in their learning. In each section below, student scores are provided in terms of strong, moderate and weak ratings:

- Strong rating requires similar knowledge and ability of at least the Mastery achievement level on the LEAP 2025 ELA Assessment;
- Moderate rating requires similar knowledge and ability as the Basic achievement level on the LEAP 2025 ELA Assessment; and
- Weak rating is comparable to the knowledge and ability required below the Basic achievement levels on the LEAP 2025 ELA Assessment.

Note that the score needed to obtain each performance rating within a category or subcategory can vary from unit to unit.

Guidance

In the classroom with the classroom code Example Classroom within the ADAM platform, there were 30 students, of which 28 had a score on the Word Play unit assessment in window 2. The table below summarizes performance by each section. Each row is a section, and the counts within each row are the number of students who are weak, moderate or strong within that section. So in the Knowledge section, 9 student(s) had scores at the Weak level, 3 student(s) had scores at the Moderate level, and 12 student(s) had scores at the Strong level.

<table>
<thead>
<tr>
<th>Section</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Application</td>
<td>9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Synthesis</td>
<td>6</td>
<td>13</td>
<td>5</td>
</tr>
</tbody>
</table>

For students in the moderate or weak score ranges, use the reflection questions in each section to help design or select follow-up supports. For students who are weak or moderate in multiple sections, start with Knowledge, then Application and finally Synthesis. Within each section the reflection questions should be considered in order. In this classroom, there are 12 students were in the Weak or Moderate Levels for Knowledge, 14 students for Application, and 19 for Synthesis.

Knowledge
In this section, students answered questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit. The table below provides a list of students by level, along with the points each student scored within parentheses. The maximum number of points was 9 in this section.

<table>
<thead>
<tr>
<th>Knowledge Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
</tr>
<tr>
<td>STU E (1)</td>
</tr>
<tr>
<td>STU L (2)</td>
</tr>
<tr>
<td>STU X (2)</td>
</tr>
<tr>
<td>STU G (3)</td>
</tr>
<tr>
<td>STU U (3)</td>
</tr>
<tr>
<td>STU P (4)</td>
</tr>
<tr>
<td>STU T (4)</td>
</tr>
<tr>
<td>STU Y (4)</td>
</tr>
<tr>
<td>STU F (5)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

For the 12 students in the weak or moderate levels, consider the following reflection questions:

1. Do students understand what they are being asked to do?
2. Do students understand the academic and domain specific language in the text and item/task?
3. Do students understand how to select evidence that best supports an idea?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.

**Application**

In this section, students read a new text(s) related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit. The maximum number of points was 13 in this section.

<table>
<thead>
<tr>
<th>Application Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
</tr>
<tr>
<td>STU T (0)</td>
</tr>
<tr>
<td>STU E (2)</td>
</tr>
<tr>
<td>STU F (2)</td>
</tr>
<tr>
<td>STU G (2)</td>
</tr>
<tr>
<td>STU X (2)</td>
</tr>
<tr>
<td>STU C (3)</td>
</tr>
<tr>
<td>STU W (3)</td>
</tr>
<tr>
<td>STU L (4)</td>
</tr>
</tbody>
</table>
For the 14 students in the weak or moderate levels, consider the following reflection questions:

4. **Do students understand how to independently use reading strategies to tackle a new text?**

If you answered yes to any of these questions, consider the example next steps in [this supporting document](#). This guidance is also available in a more complete document for all sections.

**Synthesis**

In this section, students write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit. The maximum number of points was 11 in this section, based on 4 points from the RCWE rubric times 2 plus 3 points from the LC rubric.

**Synthesis Levels**

<table>
<thead>
<tr>
<th></th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU L (0)</td>
<td>STU E (3)</td>
<td>STU D (6)</td>
<td></td>
</tr>
<tr>
<td>STU N (0)</td>
<td>STU J (3)</td>
<td>STU M (6)</td>
<td></td>
</tr>
<tr>
<td>STU T (0)</td>
<td>STU P (3)</td>
<td>STU R (6)</td>
<td></td>
</tr>
<tr>
<td>STU W (1)</td>
<td>STU A (4)</td>
<td>STU V (6)</td>
<td></td>
</tr>
<tr>
<td>STU C (2)</td>
<td>STU B (4)</td>
<td>STU O (7)</td>
<td></td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU F (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU G (4)</td>
<td>STU I (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU K (4)</td>
<td>STU Q (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU U (4)</td>
<td>STU Y (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU Z (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the 19 students in the weak or moderate levels, consider the following reflection questions:

5. **Do students understand how to independently unpack the demands of a writing task?**
6. **Do students understand what writing genre and organizational structure to use in response to the prompt (e.g. opinion, informational, or narrative)?**
7. **Do students understand how to connect ideas using transition words and complex sentences?**
8. **Do students have knowledge of grade-level appropriate spelling, grammar, and conventions?**
If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.

Profiles

Classroom Profiles

The table below summarizes the patterns of Strong, Moderate and Weak across the the Knowledge, Application and Synthesis sections for the 28 students who had a score on the Word Play unit assessment in window 2.

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Application</th>
<th>Synthesis</th>
<th>Number of Students</th>
<th>Student Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>3</td>
<td>STU L, STU T, STU X</td>
</tr>
<tr>
<td>Weak</td>
<td>Weak</td>
<td>Moderate</td>
<td>3</td>
<td>STU E, STU F, STU G</td>
</tr>
<tr>
<td>Weak</td>
<td>Moderate</td>
<td>Moderate</td>
<td>2</td>
<td>STU Y, STU U</td>
</tr>
<tr>
<td>Weak</td>
<td>Strong</td>
<td>Moderate</td>
<td>1</td>
<td>STU I</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>2</td>
<td>STU Q, STU Z</td>
</tr>
<tr>
<td>Strong</td>
<td>Weak</td>
<td>Weak</td>
<td>2</td>
<td>STU C, STU W</td>
</tr>
<tr>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
<td>1</td>
<td>STU R</td>
</tr>
<tr>
<td>Strong</td>
<td>Moderate</td>
<td>Weak</td>
<td>1</td>
<td>STU N</td>
</tr>
<tr>
<td>Strong</td>
<td>Moderate</td>
<td>Strong</td>
<td>1</td>
<td>STU D</td>
</tr>
<tr>
<td>Strong</td>
<td>Strong</td>
<td>Moderate</td>
<td>4</td>
<td>STU A, STU B, STU J, STU K, STU NA</td>
</tr>
<tr>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
<td>3</td>
<td>STU M, STU O, STU V, STU NA</td>
</tr>
</tbody>
</table>

Resources

- Parent Guide
- Parent Overviews
  - Guidebooks: 6, 7 and 8
- Score Report Guidance
- Grades 6-8 Operational Assessment Guide
- Supports Flow Chart
IAP Window 2: *Word Play* | Results and Guidance

**Overview**

The Innovative Assessment Program (IAP) measures knowledge of the module topic, reading comprehension, and application of skills and knowledge. The assessment has three sections:

1. Knowledge
2. Application
3. Synthesis

In each section, students receive a strong, moderate, or weak rating:

- Strong aligns to the Mastery achievement level on the LEAP 2025 ELA Assessment;
- Moderate aligns to the Basic achievement level on the LEAP 2025 ELA Assessment; and
- Weak aligns below the Basic achievement levels on the LEAP 2025 ELA Assessment.

The score to obtain each performance rating can vary from unit to unit.

**Results**

Assessment: Window 2—Grade 5 Module 2 *Word Play*
Classroom: Example Classroom
Students: 30 total; 28 with results on one or more section

<table>
<thead>
<tr>
<th>Section</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Application</td>
<td>9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Synthesis</td>
<td>6</td>
<td>13</td>
<td>5</td>
</tr>
</tbody>
</table>

**Summary**

Overall classroom performance can show areas of learning that were more or less successful. This can inform future instructional planning. In this classroom, overall performance shows that students performed higher on knowledge of the module topic and application of reading comprehension skills to a new passage than they did for synthesizing their knowledge and skills.
The number of students in the Weak performance level for Knowledge and Application suggests that some students may not have built foundational knowledge of the module topic.

Guidance

Students who score at the Strong Level—Continue with instruction, offering extensions as appropriate.

Students who score at the Moderate or Weak Levels—Consider performance in the Knowledge Section first, then Application, and finally Synthesis. For each category, use the following reflection questions in sequence in order to design or select follow-up supports.

Knowledge

In this section, students demonstrate their understanding of key knowledge taught in the unit by answering questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit. The table below provides a list of students in the Example Classroom by level, along with the points each student scored within parentheses. The maximum number of points was 9 in this section.

<table>
<thead>
<tr>
<th>Knowledge Levels</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU E (1)</td>
<td>STU I (6)</td>
<td>STU B (7)</td>
<td></td>
</tr>
<tr>
<td>STU L (2)</td>
<td>STU Q (6)</td>
<td>STU M (7)</td>
<td></td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU Z (6)</td>
<td>STU N (7)</td>
<td></td>
</tr>
<tr>
<td>STU G (3)</td>
<td>STU V (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU U (3)</td>
<td>STU W (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU P (4)</td>
<td>STU C (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU T (4)</td>
<td>STU K (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU Y (4)</td>
<td>STU R (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU F (5)</td>
<td>STU A (9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STU D (9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STU J (9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STU O (9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the 12 students in the weak or moderate levels, consider the following reflection questions:

1. Do students understand what they are being asked to do?
2. Do students understand the academic and domain specific language in the text and item/task?
3. Do students understand how to select evidence that best supports an idea?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.
Application

In this section, students demonstrate their ability to apply key knowledge and skills taught in the unit by reading a new topic-related text, responding to questions about it, and producing a short written response to a prompt. Students read a new text(s) related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit. The maximum number of points was 13 in this section.

### Application Levels

<table>
<thead>
<tr>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU T (0)</td>
<td>STU P (5)</td>
<td>STU B (7)</td>
</tr>
<tr>
<td>STU E (2)</td>
<td>STU D (6)</td>
<td>STU Q (7)</td>
</tr>
<tr>
<td>STU F (2)</td>
<td>STU I (6)</td>
<td>STU Z (7)</td>
</tr>
<tr>
<td>STU G (2)</td>
<td>STU N (6)</td>
<td>STU A (8)</td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU U (6)</td>
<td>STU M (8)</td>
</tr>
<tr>
<td>STU C (3)</td>
<td>STU V (8)</td>
<td></td>
</tr>
<tr>
<td>STU W (3)</td>
<td>STU Y (8)</td>
<td></td>
</tr>
<tr>
<td>STU L (4)</td>
<td>STU J (10)</td>
<td></td>
</tr>
<tr>
<td>STU R (4)</td>
<td>STU K (10)</td>
<td></td>
</tr>
<tr>
<td>STU O (11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the 14 students in the weak or moderate levels, consider the following reflection questions:

4. Do students understand how to independently use reading strategies to tackle a new text?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.

Synthesis

In this section, students demonstrate their ability to express their understanding of key knowledge gained in the unit by writing an extended response to a prompt. Write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit. The maximum number of points was 11 in this section, based on 4 points from the RCWE rubric times 2 plus 3 points from the LC rubric.

### Synthesis Levels

<table>
<thead>
<tr>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU L (0)</td>
<td>STU E (3)</td>
<td>STU D (6)</td>
</tr>
<tr>
<td>STU N (0)</td>
<td>STU J (3)</td>
<td>STU M (6)</td>
</tr>
<tr>
<td>STU T (0)</td>
<td>STU P (3)</td>
<td>STU R (6)</td>
</tr>
<tr>
<td>STU W (1)</td>
<td>STU A (4)</td>
<td>STU V (6)</td>
</tr>
<tr>
<td>STU C (2)</td>
<td>STU B (4)</td>
<td>STU O (7)</td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU F (4)</td>
<td></td>
</tr>
<tr>
<td>STU G (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU I (4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For the 19 students in the weak or moderate levels, consider the following reflection questions:

5. Do students understand how to independently unpack the demands of a writing task?
6. Do students understand what writing genre and organizational structure to use in response to the prompt (e.g., opinion, informational, or narrative)?
7. Do students understand how to connect ideas using transition words and complex sentences?
8. Do students have knowledge of grade-level appropriate spelling, grammar, and conventions?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.

Profiles

Classroom Profiles

The table below summarizes the patterns of Strong, Moderate and Weak across the the Knowledge, Application and Synthesis sections for the 28 students who had a score on the Word Play unit assessment in window 2.

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Application</th>
<th>Synthesis</th>
<th>Number of Students</th>
<th>Student Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>3</td>
<td>STU L, STU T, STU X</td>
</tr>
<tr>
<td>Weak</td>
<td>Weak</td>
<td>Moderate</td>
<td>3</td>
<td>STU E, STU F, STU G</td>
</tr>
<tr>
<td>Weak</td>
<td>Moderate</td>
<td>Moderate</td>
<td>2</td>
<td>STU P, STU U</td>
</tr>
<tr>
<td>Weak</td>
<td>Strong</td>
<td>Moderate</td>
<td>1</td>
<td>STU Y, STU NA</td>
</tr>
<tr>
<td>Moderate</td>
<td>Strong</td>
<td>Moderate</td>
<td>1</td>
<td>STU I</td>
</tr>
<tr>
<td>Moderate</td>
<td>Strong</td>
<td>Moderate</td>
<td>2</td>
<td>STU Q, STU Z</td>
</tr>
<tr>
<td>Strong</td>
<td>Weak</td>
<td>Weak</td>
<td>2</td>
<td>STU C, STU W</td>
</tr>
<tr>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
<td>1</td>
<td>STU R</td>
</tr>
<tr>
<td>Strong</td>
<td>Moderate</td>
<td>Weak</td>
<td>1</td>
<td>STU N</td>
</tr>
<tr>
<td>Strong</td>
<td>Moderate</td>
<td>Strong</td>
<td>1</td>
<td>STU D</td>
</tr>
<tr>
<td>Strong</td>
<td>Strong</td>
<td>Moderate</td>
<td>4</td>
<td>STU A, STU B, STU J, STU K, STU NA</td>
</tr>
<tr>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
<td>3</td>
<td>STU M, STU O, STU V, STU NA</td>
</tr>
</tbody>
</table>
Resources

- Parent Guide
- Parent Overviews
  - Guidebooks: 6, 7 and 8
- Score Report Guidance
- Grades 6-8 Operational Assessment Guide
- Supports Flow Chart
Introduction

The purpose of this report is to summarize results from a series of four focus groups that examined prototype score reports for the Innovative Assessment Project (IAP) during spring 2023. The IAP is meant to align state assessment with both curriculum and standards in English Language Arts. This program is operating under an Innovative Assessment Demonstration Authority (IADA) waiver that was initially granted in April 2018. The program launched operationally in the 2021-22 school year with grade 7 for a single curriculum, ELA Guidebooks. For school year 2022-23, grades 6-8 are operational, again just for ELA Guidebooks. Other grade levels are piloting this approach for the ELA Guidelines and for other curricula.

The IAP is a through-year assessment model that is curriculum Aligned, places a strong emphasis on writing, and provides districts with a choice of unit assessments to select. Figure 1 illustrates the timeline for the Grade X IAP associated with the ELA Guidebook Curriculum.

Figure 1
Timeline for the Through Course Assessments and Reporting
In the sections that follow, we describe the data collection and participants, analysis approach and summarize key findings.

**Data Collection**

A series of four focus groups were completed in March 2023, each one facilitated following a common, semi-structured protocol. The participants were recruited by the Louisiana Department of Education to be part of IAP Collaboration groups. The majority were classroom teachers, but groups also included a few non-teaching staff such as a principal or other district leaders.

Each session began with a brief introduction of the context and purpose of the meeting: as part of the Louisiana Innovative Assessment Program the Department of Education is committed to a development process that seeks feedback from Louisiana teachers and leaders. The purpose of the focus groups was to specifically collect feedback to make improvements to the score reports that are provided at the end of each unit assessment.

Participants were provided time to review several different prototype reports during the meeting, using data from their classes, including an overall class summary, a set of student profiles to show performance levels for each section, a table that showed scores and levels for each student, and links to additional resources. They were given ten to fifteen minutes to review them and note any questions, ideas or comments that they had before participating in a whole group discussion.

The whole-group discussion focused on which reports or views the teachers found most useful and why, whether results were surprising in any way, how the teachers might use the information to inform their instruction, and what else they wished was included in the reports. Each discussion was recorded and detailed notes taken to capture the feedback. Finally, participants were asked to complete an exit survey that contained two open-ended questions that asked for any additional feedback on the score reports and any additional feedback on the IAP.

Table 1 shows the number of participants by site and data source.

| Table 1
<table>
<thead>
<tr>
<th>Number of Participants for each Focus Group and Exit Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
</tr>
</tbody>
</table>
Analysis

Each set of meeting notes and exit survey comments were reviewed. All of the comments were entered into a spreadsheet to facilitate coding and sorting. There were a total of 72 separate comments, 55 (76%) of which came from the focus groups and the remaining 17 (24%) from the exit surveys. The comments were first reviewed to remove ones that were about the existing LEAP assessments, general questions, or comments made by observers rather than the participants. The remaining 60 comments were classified in a set of themes by one author and then reviewed by the second. They then met to resolve any areas of disagreement. Themes from this feedback are summarized below.

Overall Positive Reactions

Overall the participants were supportive of the information that they received in the reports, and identified ways in which they could use the information.

“We love these assessments, they are FAIR to our kids and our teachers and we are excited about the future of ELA instruction with these in place. The score reporting we have been given so far has been helpful to us in informing instructional choices and guiding student and teacher reflection.”

Common themes mentioned by the participants include the following:

- the value of the reflection questions and links to teaching strategies (N=8)
  “The reflection questions and action plan is a great starting point to help remediate students”
- being able to use the reports to inform lesson planning (N=7)
  “teachers were getting together tomorrow to talk about test results and this report
would be beneficial. They’d be able to talk about what to do with kids low in knowledge vs kids low in application.”

● the value of the student groupings (N=6)
  “thought the profiles would be beneficial for grouping students and that, yes, she could use it in her classroom”

● how they would share information with students (N=5).
  “the raw score reports are extremely powerful for her students. She passed out the reports and they discussed it whole group. She would tell them “don’t look at 2 out of 8”, look at how their score compares to the state average. Students could compare themselves to the average. Because of this, students would come to her and ask her to work with them on their writing.”
  “The students need to see more information on their reports, not less. Students use these reports to reflect on the information that they have learned and to set goals for the future, even if the tests are not scored with the same point rubric, they still deserve to see the information if they are going to be required to take the test. The move in your current direction takes away students' ability to self motivate and self reflect and pushes them to take a more removed or hands off approach to their own learning.”

Areas of Push-Back

There were a few areas where the participants made comments about the kind of additional information they would like. The most common requests are summarized below:

● There were 13 comments supporting the continued use of the raw scores or raw scores in addition to the achievement levels (or a scaled score) and only one participant who preferred the achievement levels rather than the raw scores.
  “My teachers would love for us to figure out a way to combine the score reporting to include the number score and weak/moderate/strong.”
  “Strong/Moderate/Weak by standard would be helpful.; Please keep raw score data, it helps inform instruction and students understand their scores better in comparison to state, school, and district peers.”

● There were nine comments about various kinds of comparative information that could be
useful for the teacher, students or parents. Comparisons included being able to compare an individual student’s score to a school, district or state average, but also included comparing scores across assessments within a year in order to see if students were on track for the end of the year.

“David said providing percentages in each reporting category (e.g. in state 40% were moderate in knowledge) could be a possibility. Willa said this comparative information would be helpful. When sitting down with a parent, it’s helpful to show students comparatively.”

“She would like to see how students performed compared to last year, e.g. trending right on their growth line or lower/higher than expected.”

- There were five comments about reporting student performance at the standards level. Note that the IAP assessments were not designed with sufficient items to support this level of granular reporting.

“The reflection questions and action plan is a great starting point to help remediate students but if there were standards attached would be really great.”

- Five comments were related to student performance on the writing task, which included wanting more information about what a score of zero could mean, additional information justifying a particular score, or state averages to support comparisons.

“A teacher asked if there was a way to be able to differentiate which kids need more help with addressing the prompt versus a deficit of understanding of the text itself. She said when she sees a zero, she doesn’t know if they wrote a five paragraph summary and need to now work on moving out of that, or what is the real problem.”

Conclusion

The focus groups largely provided support for the nature of the score reports. They also revealed some areas where supports may be needed to clarify why certain requests cannot be met (e.g., reporting by standard). Considering what additional information can be provided to help students use the reports and to support deeper understanding of the essay scoring may be important.
Empirical Results for Comparability of Louisiana’s Through-Year Assessment

Working Paper

Audra Kosh\textsuperscript{1}, Nathan Dadey\textsuperscript{2}, David Hopkins\textsuperscript{3}, Xiangdong Liu\textsuperscript{3}, Ruth Caillouet\textsuperscript{3}, Leslie Mugan\textsuperscript{1},

Leslie Keng\textsuperscript{4}

\textsuperscript{1}NWEA

\textsuperscript{2}National Center for Improvement of Educational Assessment

\textsuperscript{3}Louisiana Department of Education

\textsuperscript{4}National Board of Medical Examiners

National Council on Measurement in Education Annual Conference

April, 2023
Empirical Results for Comparability of Louisiana’s Through-Year Assessment

In 2018 and 2019, under the Innovative Assessment Demonstration Authority (IADA), the Office of Elementary and Secondary Education invited states to submit applications to establish and operate innovative assessment programs as alternatives to existing end-of-year state summative assessments. Four states received approvals: Georgia, North Carolina, New Hampshire, and Louisiana. An additional state, Massachusetts, was approved in 2020.

In Louisiana, students historically have taken the Louisiana Educational Assessment Program (LEAP) test. Under the IADA waiver, however, Louisiana began developing and piloting an innovative through-year, curriculum-aligned English Language Arts (ELA) assessment, known as the Innovative Assessment Program (IAP). Louisiana was one of the first states to report summative scores of record from an IADA program, beginning with grade seven students that participated in the IAP pilot in school year 2021-22. This reporting was also one of the first for a through-year program meant for a general population\(^1\). With grade seven students across Louisiana receiving scores from either the IAP assessment or the statewide LEAP assessment, it is critical to understand the degree to which scores from each assessment are comparable. This paper addresses the overarching research question: what evidence exists to demonstrate the degree of comparability of IAP and LEAP scale scores? The papers focus on empirical results of student scores and educator ratings of alignment between IAP content and LEAP achievement levels. Readers are encouraged to first read Dadey and Kosh (2023) for an overview of the IAP test design and scaling methodology prior to reading this paper.

The scale scores and achievement levels produced by IAP are meant to be comparable to those of the statewide LEAP ELA assessment program. As described by Lyons and Marion

\(^1\) Dynamic Learning Maps, an alternate assessment of alternate academic achievement standards, has offered a through-year option, referred to as the “instructionally embedded model” since 2016 (see Clark & Karvonen, 2021).
“Comparability is a judgment based on an accumulation of evidence to support claims about the meaning of test scores and whether scores from two or more tests or assessment conditions can be used to support the same interpretations and uses” (p. 8). Lyons and Marion go on to recommend that comparability be considered in terms of (1) proficiency or achievement level classifications, and (2) the alignment of each system to the content standards. They also provide a guideline for evaluating comparability – that evidence from statewide programs can serve as a best case scenario for comparison. The degree of comparability established between multiple forms within and across years or across modes (i.e., through equating) provides an empirical criterion. The comparability of the IADA program to the statewide program can then be evaluated in relation to this, although additional judgment is needed to determine the specific metric and how close is “close enough”. Essentially, Lyons and Marion advocate for an evaluation of student-level achievement level classifications that involve not only analysis based on student performance but also based on content experts' judgments of alignment.

Defining comparability as a judgment implies that there are different types of comparability claims. In this work we have aimed for comparability for both scale scores and achievement levels, as a conservative starting point for the program. The comparability claim is then that students would have similar scale scores and achievement levels, regardless of which program they took.

Following Lyon and Marion’s recommendation to investigate comparability from multiple perspectives, including in terms of alignment, the comparability of the resulting TSI and CSI classifications is supported by three bodies of evidence: (a) the quality of the linking to LEAP via a common item non-equivalent group design; (b) empirical analyses that compare performance between IAP and non-IAP students, and; (c) the judgments of teachers about the
alignment of the IAP items to LEAP achievement level descriptors (ALDs) through an ALD alignment workshop. This paper describes evidence for (b) and (c); the supplementary paper Dadey and Kosh (2023) provides evidence for (a). The specific research questions for the empirical analyses and alignment workshop are:

1. How similar, or divergent, are the aggregate trends in individual-level data in terms of:
   a. Patterns of achievement level transitions from grade six in 2020-21 to grade seven in 2021-22 across the two programs (i.e., grade six LEAP to grade seven IAP compared to grade six LEAP to grade seven IAP)?
   b. Correlations of scale score from grades six and seven in 2020-2021 and 2021-22 across the two programs?
   c. Percent of students in each achievement level from IAP in grade seven in 2021-22 as compared to historical trends of LEAP achievement level distributions for grade seven ELA from 2015-16 to 2020-21?

2. How do educators rate the alignment of IAP content and LEAP ALDs?

The next two sections of the paper address these research questions, starting with comparisons of student performance (RQs 1a, 1b, 1c) and moving to educator judgements of alignment (RQ 2) 

**Comparisons of Students Performance**

Using data from the operational administration of IAP to grade seven students in school year 2021-22, as well as LEAP data from 2021-22 and prior years, this section presents evidence based on empirical analyses that demonstrate the degree of comparability between IAP and LEAP.

**Method**
Because students participating in IAP represent a sample of Louisiana’s statewide population, propensity score matching was used to create a matched sample of students that participated in IAP with similar demographic composition and prior ELA achievement as the LEAP statewide sample. The sets of students used for matching were students with two years of ELA test scores, namely LEAP ELA scores in grade six for school year 2020-21 and either LEAP ELA or IAP scores in grade seven for school year 2021-22. Students that took IAP in grade seven are known as the “IAP Sample” whereas students that took LEAP in grade seven are known as the “LEAP Comparison Sample”. Demographic variables present in the data were: gender, race/ethnicity, migrant status (“yes” if the primary funding source for the student’s instructional language services is identified as English Language Learner), homeless status, foster care status, military affiliation status (“yes” if students’ parent(s) are active, retired, or reserve military), and economically- disadvantaged status (“yes” if the student is eligible for Louisiana’s food assistance program for low-income families, disaster food assistance program, or assistance to needy families with children to assistant parents in becoming self-sufficient, respectively known as SNAP, DSNAP, and TANF). In some cases, students were missing race/ethnicity data or economically-disadvantaged status and were removed from the data set. Students in the IAP sample were also removed if they did not have LEAP grade six scores in school year 2020-21. Table 1 shows sample counts before and after data filtering.
Table 1

Sample Size Counts Before and After Data Filtering

<table>
<thead>
<tr>
<th>Sample</th>
<th>N Before Filtering</th>
<th>Missing Race/Ethnicity</th>
<th>Missing Spring 2021 LEAP Scale Score</th>
<th>Missing Economically-Disadvantaged Status</th>
<th>N After Filtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEAP Comparison Sample</td>
<td>42,279</td>
<td>20</td>
<td>0</td>
<td>77</td>
<td>42,182</td>
</tr>
<tr>
<td>IAP Sample</td>
<td>4,733</td>
<td>23</td>
<td>286</td>
<td>0</td>
<td>4,424</td>
</tr>
</tbody>
</table>

Note. Counts after filtering may not equal the difference of counts before filtering and the sum of missing data due to some students having missing data in multiple categories.

Propensity score matching was performed with covariates including all available aforementioned demographic variables as well prior year LEAP ELA scale scores (i.e., grade six scores from SY 20-21). Nearest neighbor matching was used without replacement via the R “MatchIt” package. Table 2 shows the distribution of covariates before and after matching.
### Table 2

**Frequencies of Covariates Before and After Propensity Score Matching**

<table>
<thead>
<tr>
<th></th>
<th>Before Matching</th>
<th>After Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEAP Comparison Sample</td>
<td>IAP Sample</td>
</tr>
<tr>
<td>N</td>
<td>42,182</td>
<td>4,424</td>
</tr>
<tr>
<td>LEAP Scale Score from Grade 6 in 2021, Mean (SD)</td>
<td>737.00 (31.17)</td>
<td>738.58 (28.82)</td>
</tr>
<tr>
<td>Gender, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>20,814 (49.3)</td>
<td>2,208 (49.9)</td>
</tr>
<tr>
<td>Male</td>
<td>21,368 (50.7)</td>
<td>2,216 (50.1)</td>
</tr>
<tr>
<td>Race/Ethnicity, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>264 (0.6)</td>
<td>7 (0.2)</td>
</tr>
<tr>
<td>Asian</td>
<td>638 (1.5)</td>
<td>49 (1.1)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>18,296 (43.4)</td>
<td>1,868 (42.2)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>3,970 (9.4)</td>
<td>153 (3.5)</td>
</tr>
<tr>
<td>Native Hawaiian/Other Pacific Islander</td>
<td>35 (0.1)</td>
<td>1 (0.0)</td>
</tr>
<tr>
<td>Two or more races</td>
<td>1,449 (3.4)</td>
<td>71 (1.6)</td>
</tr>
</tbody>
</table>
## COMPARABILITY OF LOUISIANA’S THROUGH-YEAR ASSESSMENT

<table>
<thead>
<tr>
<th></th>
<th>Before Matching</th>
<th>After Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEAP Comparison Sample</td>
<td>IAP Sample</td>
</tr>
<tr>
<td><strong>Student is migrant, N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42,141 (99.9)</td>
<td>4,416 (99.8)</td>
</tr>
<tr>
<td>Yes</td>
<td>41 (0.1)</td>
<td>8 (0.2)</td>
</tr>
<tr>
<td><strong>Student is homeless, N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>41,248 (97.8)</td>
<td>4,398 (99.4)</td>
</tr>
<tr>
<td>Yes</td>
<td>934 (2.2)</td>
<td>26 (0.6)</td>
</tr>
<tr>
<td><strong>Student is in foster care, N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42,097 (99.8)</td>
<td>4,407 (99.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>85 (0.2)</td>
<td>17 (0.4)</td>
</tr>
<tr>
<td><strong>Student has military affiliation, N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>41,447 (98.3)</td>
<td>4,413 (99.8)</td>
</tr>
<tr>
<td>Yes</td>
<td>7,35 (1.7)</td>
<td>11 (0.2)</td>
</tr>
<tr>
<td><strong>Student is economically disadvantaged, N (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12,516 (29.7)</td>
<td>1,542 (34.9)</td>
</tr>
<tr>
<td>Yes</td>
<td>29,666 (70.3)</td>
<td>2,882 (65.1)</td>
</tr>
</tbody>
</table>
Using the matched sample, a variety of analyses were conducted and then compared across IAP students and LEAP students. These analyses included the degree of agreement between students’ performance levels from grade six in school year 2020-21 to their performance levels from grade seven in school year 2021-22 and the correlation between students’ scale scores from those same years of testing. Additionally, using the full unmatched sample of IAP students and historical trends of LEAP data for districts that participated in IAP, analyses examined distributions of achievement levels between 2016 and 2021 for LEAP compared to the distribution of 2022 IAP achievement levels.

**Results**

Below we provide the results by each research subquestion.

**RQ 1a: Comparison of Achievement Level Transitions Across Programs**

First, the level of agreement between students’ achievement levels from 2021 to 2022 was extremely similar for students in the 2022 IAP pilot and those not in the pilot (e.g., students taking LEAP). Specifically, for the matched sample of students in the grade seven IAP pilot that also took LEAP as sixth graders in 2021, the weighted kappa (between grade six LEAP achievement level in 2021 and grade seven IAP achievement level in 2022) was 0.567 ($p < .001$). For comparison, for students that were not in the IAP pilot but rather took LEAP in both grade six (2021) and grade seven (2022), the weighted kappa was 0.558 ($p < .001$). Tables 3 through 6 show the counts and relative frequencies of achievement levels from grade six and grade seven, for the matched IAP sample and LEAP comparison sample.
### Table 3

*Frequencies of Achievement Level Classifications From Grade Six LEAP and Grade Seven IAP*

<table>
<thead>
<tr>
<th>LEAP Achievement Level in Grade Six 2021</th>
<th>Un satisfactory</th>
<th>Approaching Basic</th>
<th>Basic</th>
<th>Mastery</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>221</td>
<td>137</td>
<td>42</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Approaching Basic</td>
<td>181</td>
<td>423</td>
<td>387</td>
<td>82</td>
<td>2</td>
</tr>
<tr>
<td>Basic</td>
<td>16</td>
<td>170</td>
<td>604</td>
<td>528</td>
<td>29</td>
</tr>
<tr>
<td>Mastery</td>
<td>0</td>
<td>8</td>
<td>163</td>
<td>829</td>
<td>420</td>
</tr>
<tr>
<td>Advanced</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>149</td>
</tr>
</tbody>
</table>

### Table 4

*Frequencies of Achievement Level Classifications From Grade Six LEAP and Grade Seven*

*LEAP*

<table>
<thead>
<tr>
<th>LEAP Achievement Level in Grade Six 2021</th>
<th>Un satisfactory</th>
<th>Approaching Basic</th>
<th>Basic</th>
<th>Mastery</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>263</td>
<td>110</td>
<td>25</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Approaching Basic</td>
<td>234</td>
<td>394</td>
<td>360</td>
<td>84</td>
<td>3</td>
</tr>
<tr>
<td>Basic</td>
<td>34</td>
<td>180</td>
<td>538</td>
<td>559</td>
<td>36</td>
</tr>
<tr>
<td>Mastery</td>
<td>1</td>
<td>19</td>
<td>160</td>
<td>769</td>
<td>472</td>
</tr>
<tr>
<td>Advanced</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>24</td>
<td>155</td>
</tr>
</tbody>
</table>
### Table 5

*Relative Frequencies of Achievement Level Classifications From Grade Six LEAP and Grade Seven IAP*

<table>
<thead>
<tr>
<th>LEAP Achievement Level in Grade Six 2021</th>
<th>Unsatisfactory</th>
<th>Approaching Basic</th>
<th>Basic</th>
<th>Mastery</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>0.05</td>
<td>0.031</td>
<td>0.009</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Approaching Basic</td>
<td>0.041</td>
<td>0.096</td>
<td>0.087</td>
<td>0.019</td>
<td>0</td>
</tr>
<tr>
<td>Basic</td>
<td>0.004</td>
<td>0.038</td>
<td>0.137</td>
<td>0.119</td>
<td>0.007</td>
</tr>
<tr>
<td>Mastery</td>
<td>0</td>
<td>0.002</td>
<td>0.037</td>
<td>0.187</td>
<td>0.095</td>
</tr>
<tr>
<td>Advanced</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.007</td>
<td>0.034</td>
</tr>
</tbody>
</table>

### Table 6

*Relative Frequencies of Achievement Level Classifications From Grade Six LEAP and Grade Seven LEAP*

<table>
<thead>
<tr>
<th>LEAP Achievement Level in Grade Six 2021</th>
<th>Unsatisfactory</th>
<th>Approaching Basic</th>
<th>Basic</th>
<th>Mastery</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>0.059</td>
<td>0.025</td>
<td>0.006</td>
<td>0.001</td>
<td>0</td>
</tr>
<tr>
<td>Approaching Basic</td>
<td>0.053</td>
<td>0.089</td>
<td>0.081</td>
<td>0.019</td>
<td>0.001</td>
</tr>
<tr>
<td>Basic</td>
<td>0.008</td>
<td>0.041</td>
<td>0.122</td>
<td>0.126</td>
<td>0.008</td>
</tr>
<tr>
<td>Mastery</td>
<td>0</td>
<td>0.004</td>
<td>0.036</td>
<td>0.174</td>
<td>0.107</td>
</tr>
<tr>
<td>Advanced</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.005</td>
<td>0.035</td>
</tr>
</tbody>
</table>
**RQ 1b: Comparison of Correlations Across Programs**

Second, again using the matched samples, the correlation between IAP students’ LEAP scores in grade six and their IAP score in grade seven \((r = .84)\) was similar to the correlation between grade six and grade seven LEAP scores for non-pilot students \((r = .82)\). Figures 1 and 2 show the scatter plots for these relationships.

**Figure 1**

*Relationship Between Students’ LEAP Grade Six Scale Scores in 2021 to IAP Grade Seven Scale Scores in 2022*
Figure 2

*Relationship Between Students’ LEAP Grade Six Scale Scores in 2021 to LEAP Grade Seven Scale Scores in 2022*

![Matched LEAP Comparison Sample](image)

\[ R = 0.82, p < 2.2e^{-16} \]

**RQ 1c: Comparing Historical Distributions of LEAP 2016-2021 and IAP 2022**

Third, based on historical aggregate trends for school districts that participated in the IAP pilot, Table 6 shows the distribution of students across achievement levels between 2016 and 2021 for LEAP compared to the distribution of 2022 IAP achievement levels. The data in this table are based on the full IAP sample for 2022 (i.e., no propensity score matching). For 2016-2021 comparisons, the table includes all students that were historically enrolled in districts that participated in IAP. In other words, if a student took LEAP in 2016 as a seventh grader in a district that later joined the IAP pilot when it became available in 2022 for seventh graders, that student’s data are included in this table. Thus, the table provides the most direct comparison of historical data by filtering to only IAP participating districts. With the exception of LEAP in 2021, which was likely impacted by pandemic-related school closures, the distribution of IAP
performance levels is generally similar to historical LEAP trends, especially when compared to the most recent school year that was not impacted by the pandemic (i.e., 2019).

Table 6

*Historical Trends of Percent of Students in Each Achievement Level (Grade Seven), for Students in Districts that Participated in IAP*

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Unsatisfactory</th>
<th>Approaching</th>
<th>Basic</th>
<th>Mastery</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>Basic %</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>2016 LEAP</td>
<td>5,211</td>
<td>11</td>
<td>19</td>
<td>29</td>
<td>29</td>
<td>11</td>
</tr>
<tr>
<td>2017 LEAP</td>
<td>5,326</td>
<td>14</td>
<td>19</td>
<td>26</td>
<td>30</td>
<td>11</td>
</tr>
<tr>
<td>2018 LEAP</td>
<td>5,166</td>
<td>10</td>
<td>18</td>
<td>29</td>
<td>32</td>
<td>11</td>
</tr>
<tr>
<td>2019 LEAP</td>
<td>5,123</td>
<td>10</td>
<td>16</td>
<td>25</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td>2021 LEAP</td>
<td>5,059</td>
<td>10</td>
<td>26</td>
<td>30</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>2022 IAP</td>
<td>4,677</td>
<td>9</td>
<td>17</td>
<td>27</td>
<td>33</td>
<td>14</td>
</tr>
</tbody>
</table>

*Note.* Data include eight out of nine districts participating in IAP. No statewide assessment was administered in 2020 due to the COVID-19 pandemic.

**ALD Alignment Workshop**

The comparisons of student performance form one dimension of the overall comparability argument. In addition to the student comparison analyses, subject matter experts were enlisted to make alignment judgements. Drawing on the ALD alignment method originally developed and implemented for the National Assessment of Educational Progress (NAEP; Donahue, Pitoniak, & Beaulieu, 2010; National Assessment Governing Board, 2021), educators participated in a multi-day workshop to make qualitative judgements regarding alignment between IAP and LEAP. On May 16, 17, and 18, 2022 a group of six educators met virtually
with representatives from the National Center for the Improvement of Educational Assessment, NWEA, and the Louisiana Department of Education to determine the degree of alignment between the grade seven IAP and the ALDs used for LEAP assessments. The ALD alignment method is a multistep process that involves the creation of descriptions of the knowledge and skills needed to respond to the IAP items, which are then used to develop “summary statements” that are finally compared to the current LEAP ALDs.

**ALD Review Panels**

Panelists were drawn from school and districts participating in the IAP pilot, including existing committees formed for item bias and sensitivity reviews during the test development process. Committee members represented five districts; most of these districts participated in the operational grade seven assessment, and one district participated in field and pilot testing. The majority of panelists were classroom teachers that had more than six years of teaching experience, and all panelists had significant knowledge of the grade seven units of instruction. Demographics and other information regarding the committee members is summarized in Table 7.
Table 7

Panelists’ Demographic Information

<table>
<thead>
<tr>
<th>Total Number Participants</th>
<th>% White</th>
<th>% Black</th>
<th>% with 20+ years Experience</th>
<th>% with 10 or less Years Experience</th>
<th>% with Experience Teaching ELA Guidebook Unitsa</th>
<th>% with Experience Teaching Students with an IEP</th>
<th>% with Experience Teaching English Learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>83%</td>
<td>17%</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
<td>83%</td>
<td>66%</td>
</tr>
</tbody>
</table>

a ELA Guidebooks is the curriculum for which IAP is aligned. For more information see Dadey and Kosh (2023).

**ALD Alignment Materials**

Materials for the workshop were organized into a Google Drive folder which included folders for each panelist as well as a folder of facilitator materials. In advance of the workshop, items were sorted into achievement levels following the same methods as described by the National Assessment Governing Board (2021) for NAEP. Because the ALD alignment workshop was held in May 2022, prior to the third window administration window of IAP, only window one and window two items and data were used, meaning items from one of five unit assessments was not included. The process of sorting items into achievement levels was as follows:

1. Calculate students’ theta estimates from window one and two data using the two-parameter item response theory (IRT) model based on the pooling approach (see Dadey & Kosh, 2023).
2. Classify students into LEAP achievement levels based on the pre-established cut scores for LEAP. The LEAP achievement levels are labeled as Unsatisfactory, Approaching Basic, Basic, Mastery, Advanced.

3. For all students in a given LEAP achievement level, calculate those students’ probability of a correct response to each item (or each score point for multi-point items), again using the two-parameter IRT model. Average the probability across all students in the achievement level. This results in five conditional p-values for every item: one p-value per achievement level that is conditioned upon students in the particular achievement level. The NAEP achievement level alignment process refers to this averaged p-value as the anchoring probability.

4. For each item, locate the lowest achievement level for which the anchoring probability is greater than or equal to 0.67, and classify the item with that achievement level. For example, hypothetically, if an item had anchoring probability for Unsatisfactory at .25, Approaching Basic at .36, Basic at .56, Mastery at .68, and Advanced at .79, the item would be classified as Mastery. In the case were an item did not have a conditional p-value greater than 0.67 for any level, a cut off of greater than or equal to 0.50 was used, again looking for the lowest achievement level. This secondary cut resulted in placing all items into LEAP achievement levels.

The result of this process is that every item or score point is associated with a LEAP achievement level. Items were then provided to panelists grouped by achievement level as part of the first step, as described in the next section. Within each achievement level, items were provided from easiest to most difficult, based on the conditional p-value.
ALD Alignment Review Tasks and Procedures

After introductions, committee members were reminded about security and confidentiality, and were given an overview of the materials and how to access them. Each participant was given access to a Google folder\(^2\), which housed all of the materials for that participant, including multiple scoring spreadsheets. Secure materials, including item content, were housed separately on the test platform in forms that matched the item order of the scoring spreadsheets. The committee members were also provided with an orientation of the process steps to be completed over the course of the three days. Committee members developed individual summaries for each innovative assessment item, which would be used to create whole group judgment of the alignment between the innovative assessment and the achievement level descriptors. Figure 3 provides an overview of the process in addition to the text below for each step.

The committee members were also instructed that:

- The process is entirely dependent on each participant’s professional judgment.
- There are no right or wrong answers.
- The research team/facilitators cannot provide specific direction. This high-level guidance is by design, to ensure that the results reflect just the group’s expert judgment.

\(^2\) The Google docs suite provided a simple solution to coordinating all of the multiple materials used the workshop. Instead of standing up a website or application based portal, as often typical in similar processes like standards setting, we used a series of linked spreadsheets (using the IMPORTRANGE function to reference the url of the various spreadsheets) and docs to facilitate the workshop. In addition, many educators are familiar with Google docs.
Figure 3

*Overview of ALD Alignment Workshop Process*

The general structure of the three day workshop was:

- **Day 1:**
  - Step 1. Creation of individual item statements
  - Step 2. Individual summary statements
  - Step 3. Small group summary statements

- **Day 2:**
  - Step 4. Alignment judgments of small group summary statements
  - Step 5. The creation of whole group summary statements
  - Step 6. Alignment judgments of whole group summary statements

- **Day 3:**
  - Step 7. Group discussion regarding recommendations for increased alignment
  - Step 8. Development of performance level descriptors for the reporting categories (not discussed in this paper)
Step 9. Initial reporting category cut score creation (not discussed in this paper)

Step 1: Item Statements

During step one, panelists received access to the assessment’s online platform in order to review the items in the sequence they occurred for each form. Participants were instructed to describe the knowledge and skills required to answer the item correctly or receive a specific score point. Participants were also instructed that the statements did not need to be complete sentences and, instead, could be notes that were written for their benefit. The participants were also instructed to think beyond the standards – to focus on the specific knowledge and skills a student needs to correctly answer the item.

Prior to starting the item statement writing, participants were asked to develop practice item statements from an unused test form developed for the ELA guidebook unit, Memoir. The purpose of this practice was to both orient them to the tools being used, as well as to provide them an opportunity to practice writing an item statement. Once the practice was complete, participants transitioned to writing item statements individually. Participants were given a Google workbook that had individual tabs for each assessment form, allowing participants to generate item statements while experiencing the test items and associate text passages as students would.

Step 2: Summary Statements

After developing the item statements, the participants were brought together to review the tools and the process that would be used for this step. Although similar to how the items were set up for the creation of the item statements, items and their corresponding item statements were arranged differently. The key differences were the following:
Items were grouped into tabs by achievement levels, not by form, beginning with the lowest achievement level (Unsatisfactory) and increasing to the highest (Advanced), based on the method described above.

Within each achievement level, items were ordered by difficulty (easiest to hardest).

Items included a space for participants to provide an overall summary statement.

Once oriented with the materials, participants were reminded that the summary statements are meant to both capture the knowledge and skills involved in answering the items in the specific level, as well as to generalize across the individual item statements.

**Step 3: Small Group Summary Statements**

The next step of the process involved the creation of small group summary statements. During this step in the process, the committee members were placed into two groups made up of three members. They were asked to develop a small group summary statement for each achievement level by examining, and building on, the individual summary statements. In order to support this work, each small group was provided a document containing the individual summary statements for the committee members in that group.
Step 4: Alignment Judgments – Small Group Summary Statements

First working as individuals, participants were asked to make an overall judgment of each small group summary statement as compared to the LEAP ELA ALD. They were asked to determine if the level of alignment at each achievement level was considered to be weak, moderate, or strong, and they were encouraged to add comments related to their judgments, particularly for judgments of moderate and weak. The facilitator reminded the participants that alignment judgments are meant to be holistic, and when considering the classification of weak, moderate or strong alignment judgments, they should look across the summary statements and the LEAP ALDs and focus on the bigger picture.

Step 5: Alignment Judgment Review – Small Group Summary Statements

During step 5, the participants reconvened in their small groups. As a group they reviewed a summary of their ratings at each level and considered whether the summary statements should be revised to clarify or better explain the knowledge and skills involved in answering the items in the specific level. Revisions were then discussed and agreed to within each group, and the summary statements were revised.

Step 6: Whole Group Summary Statements

For this step in the process, the full panel convened together again and worked together to create a whole group summary statement. To do this, they reviewed both small group summary statements, and then developed a whole group statement.

Step 7: Individual Alignment Judgment - Whole Group Summary Statements
Finally, participants were asked to make individual judgments between the whole group summary statements and the LEAP ALDs. After each round, participants discussed why they judged the alignment as weak, moderate, or strong, and then went back to consider the discussions and comments as they approached the next round. This continued for three rounds. Results of these judgments can be found in the Results and Conclusions section.

*Step 8: Alignment Recommendations*

Finally, the committee was asked to provide their recommendations regarding how to increase alignment between IAP and the LEAP achievement level descriptors. It was noted that these recommendations could be in terms of:

- Items
- Standards Coverage
- Text(s)
- Other aspects of the IAP Assessment

**Results**

The participants generally found strong alignment between the items in the innovative assessment and the LEAP ALDs. As further noted in Table 8, while there was some judgment of weak alignment, in all cases at least 80% of the participants rated the alignment as Moderate or Strong.
Table 8

*Summary of Alignment Judgments*

<table>
<thead>
<tr>
<th>Level</th>
<th>Achievement Area</th>
<th>Counts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weak (1)</td>
<td>Moderate (2)</td>
<td>Strong (3)</td>
</tr>
<tr>
<td>2</td>
<td>Reading</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Writing</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Reading</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Writing*</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Reading</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Writing</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Reading</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Writing</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*One participant did not rate this category.

After the judgments were completed, the panel members were asked to consider recommendations for increasing the alignment between the innovative assessment and the LEAP achievement level descriptors. Recommendations were focused on both specific and general areas of improvements across Levels 2-5 of the LEAP ALDs.

*Level 2:*

Recommendations for Improved Alignment based on the Whole Group Summary Statement:

- Add more segments of the anchor text for students to analyze before attempting the questions to remind students of parts of the novel.
- Add more questions to the warm read on the assessment
- Lower complexity level on the cold reads
• Possibly create an essay topic that does not require most items looked for in the state rubric/different expectations/scaling for each test across test windows.
• Scale the essays so that Window 1 has an essay on the anchor text, Window 2 has an essay on two texts, and Window 3 has an essay on the anchor, supplemental, and warm read.
• Include vocabulary that is directly tied to levels (complexity of the vocabulary skills).

General Recommendations:
• Consider expectations in each window, especially the writing.
• Consider vocabulary as they move from level to level (a progression in how they acquire vocabulary skills); more about preparation in how teachers teach vocabulary.
• More generally, consider progression across the units, in terms of complexity.
• Align the complexity of the culminating writing task to the complexity of the essay.

Level 3:

Recommendations for Improved Alignment based on the Whole Group Summary Statement
• Use vocabulary that is directly tied to levels (complexity of the vocabulary skills).
• Include more decisions around listing specific skills (in descriptors).
• Include more mention of citing evidence in the summary descriptors to better align with LEAP ALDs.
• Consider text complexity on warm read; more readily accessible texts.
• Address style and organization in summary descriptors to better align with LEAP ALDs.

General Recommendations:
● Include more opportunities to show proficiency in identifying evidence in various ways and item types (add TEIs).

● Include more text and context in hot-read section (not just mention a chapter), especially at level 2 and 3.

● Provide more text for students to respond to.

● Consider how to match text complexity and higher interest to increase ability to make connections.

● Include more support around selected-response questions.

● At every level, provide students with the texts they are asked to write about. Essays might be based on warm reads.


\textit{Level 4:}

Recommendations for Improved Alignment based on the Whole Group Summary Statement

● In SR questions based on the novel, need more text/context to jog students’ memory and support understanding of what the question is asking.

● Based essays on the warm reads and not the novel.

● Add more items that are directly tied to level 4 (vocabulary, theme, etc.)

● Address style and organization in summary descriptors to better align with LEAP ALDs.

General Recommendations:

● On the essay, just paraphrasing evidence now; opportunity to include direct evidence.

● Adding more questions attached to the skill of paraphrasing and summarizing (indirect quotes); choose best paraphrase.
● Like to see a better spread of items across the levels; a better range across complexity levels (add more questions to get a better determination of student understanding); more like an end-of-unit test than a section quiz; more item types.
● Add more questions about the supplemental texts in the unit, with enough of the text included as part of the items (BTS, Carol supplemental texts: If People Could Fly, etc.); that’s the GBs: would add more alignment to the approach of the GBs.
● Focus on how well students learned the unit, not just the anchor texts.
● Provide access to at least excerpts from the text when answering specific questions about that text (to provide more context).
● Rely less on the students’ memory.

Level 5:

Recommendations for Improved Alignment based on the Whole Group Summary Statement

● Address style and organization in summary descriptors to better align with LEAP ALDs.

Panelists’ Feedback

As part of the ALD alignment workshop, panelists were asked to participate in two surveys; one after Day 1, and one after Day 3. Day 1 feedback was used to inform any modifications the facilitators may want to consider for the remainder of the workshop, and Day 3 feedback was used to understand panelists’ overall satisfaction with the meeting. Five out of six panelists completed the surveys. Generally, panelists understood the purpose of the meeting and felt satisfied with the item statements and summary statements they developed. Open-ended comments indicated that panelists felt the meeting was well run. Full survey results are available in Appendix A.
Summary and Discussion

Results provide strong evidence of comparability of IAP scale scores and achievement levels to those of LEAP via multiple methods. First, using propensity matching to create a matched sample of IAP students and LEAP students based on demographic characteristics and prior ELA achievement, all empirical analyses had nearly identical results for both the IAP sample and the LEAP comparison sample. Second, educators that participated in a multiday workshop to review IAP content for the degree of alignment to LEAP ALDs generally rated alignment of IAP to LEAP ALDs as strong or moderate.

Future work should consider how these results change or stabilize over time as the IAP program matures and expands to additional grade levels and/or curriculums. In addition, future work could explore the development of specific criteria for comparability to operationalize how close is “close enough” to support the claim that students would have similar scale scores and achievement levels, regardless of which program they participated in. In relation to this, the current work assumed a fairly conservative stance towards comparability. Future work may consider more liberal definitions, potentially involving whether or not schools are identified, or can exit, Comprehensive Support and Improvement (CSI) and Targeted Support and Improvement (TSI) statuses.
References


Appendix A

Day One Survey and Responses

The ALD Alignment workshop goals and process were clearly explained.
5 responses

- 40% Strongly Agree
- 60% Agree

Item Statements

I understood the tools to create the Item Statements.
5 responses

- 40% Strongly Agree
- 60% Agree
I understood the purpose for creating the Item Statements.
5 responses

How satisfied are you with the Individual Item Statements you created today?
5 responses
Summary Statements

I understood the purpose of creating the Summary Statements.
5 responses

- Strongly Agree: 60%
- Agree: 20%
- Disagree: 20%
- Strongly Disagree: 0%

How satisfied are you with the Item Summary Statements you created today?
5 responses

- Very Satisfied: 40%
- Satisfied: 40%
- Neutral: 20%
- Dissatisfied: 0%
- Very Dissatisfied: 0%
Days Two and Three Survey and Responses

I understood the purpose of the Group Summary Statements.

- Strongly disagree: 20%
- Disagree: 20%
- Agree: 60%
- Strongly Agree: 20%
I understood how to use the tools to create the Group Summary Statements.
5 responses

How satisfied are you with the Group Summary Statements written by your group?
5 responses
I understood the purpose of the individual alignment judgements.
5 responses

80% Strongly Agree
20% Agree

I understood how to use the tools to create the individual alignment judgements.
5 responses

80% Strongly Agree
20% Agree
I understood the purpose of creating the Whole Group Summary Statements.
4 responses

How satisfied are you with the final judgements of the Whole Group Summary Statements?
5 responses
I understood the purpose of the Benchmarking activity.
5 responses

- Strongly Agree: 80%
- Agree: 20%

How satisfied are you with the outcome of the Benchmarking activity?
5 responses

- Very Satisfied: 40%
- Satisfied: 40%
- Neutral: 20%
Based on your participation on Days 2 and 3, is there anything that can be improved about the tools used or the process? Other thoughts you would like to share? Thank you so much for your participation!

5 responses

<table>
<thead>
<tr>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thank you for this opportunity!</td>
</tr>
<tr>
<td>Very well run meeting. Working on the Benchmark activity was a little confusing at first with accessing the different documents, etc., but the hosts were extremely helpful.</td>
</tr>
<tr>
<td>Everything was well thought out!</td>
</tr>
<tr>
<td>I think that once the test(s) are revised with suggestions given, we may be able to take a look at all of this again and make better judgements on the benchmarking.</td>
</tr>
<tr>
<td>I think everyone did fabulous job of participating and facilitating. This was hard but fun work. I enjoyed it and would love to be included for more. The students in Bossier Parish enjoyed the IAP.</td>
</tr>
</tbody>
</table>
Scaling and Score Creation for
Lousianna’s Innovative Assessment Demonstration Authority Program

Working Paper
Nathan Dadey\textsuperscript{1} and Audra Kosh\textsuperscript{2}

\textsuperscript{1} National Center for Improvement of Educational Assessment

\textsuperscript{2} NWEA

National Council on Measurement in Education Annual Conference
April, 2023
Abstract

The Louisiana Innovative Assessment Program (IAP) is designed to integrate state assessment and accountability with curriculum and instruction through the use of multiple assessments that are based on English Language Arts curricular units. Under a waiver provided through the Innovative Assessment Demonstration Authority (IADA), results from these unit based assessments are used to produce annual student determinations - scale scores and achievement levels - used within statewide accountability. In this paper we provide a description of the scaling and scoring methods used within the IAP program, as well as context for this unique program, as it is an operational, curriculum-based, through-year program operating under an IADA waiver.
Introduction

Since late 2018 and 2019, a number of states and their partners have been working on assessment programs under waivers granted by the Innovative Assessment Demonstration Authority (IADA) authorized in section 1204 of the Every Student Succeeds Act. The first of five¹ states to be approved under the waiver (USDOE, n.d.), Louisiana’s approach is to use a set of end of unit assessments within each grade to produce annual student determinations - achievement level classifications as well as the scale scores they are based on - used within statewide accountability. This approach also means that the Louisiana program is amongst a newly emerging set of through-year programs, which have recently garnered much attention (e.g., Dadey, Evans & Lorié; 2023; Education First, 2022). Even amongst these novel and emerging programs, Louisiana’s program is unique as it is directly tied to English Language Arts (ELA) curricular units.

Within the first section of this paper, we provide background on Louisiana's program and the state context, as well as background on the IADA and through-year assessment programs. In the second section we detail the assessment design and the approach taken to scale and score the results of the IAP during the 2021-2022 school year. The third second presents results, and the fourth section provides conclusions and considerations.

Background

The Louisiana Innovative Assessment Program

The Louisiana program, known as the Innovative Assessment Program (IAP), is designed to assess students in relation to both the state academic content standards and the curriculum.

¹ Four states, Georgia, Louisiana, New Hampshire and North Carolina were approved in 2018 and 2019. An additional state Massachusetts, was approved in 2020. This state is not listed on the USDOE’s IADA webpage (USDOE, n.d.).
During the 2021-2022 school year, the program was operationally administered in seventh grade to students whose instruction was based on the ELA Guidebooks curriculum, which is an open source ELA curriculum developed and supported by the Louisiana Department of Education. ELA Guidebooks unit assessments are being administered in grades six to eighth during the 2022-23 school year, with additional grades being added in following years. The majority of schools and districts within Louisiana use ELA Guidebooks. Unit assessments for another widely used curriculum, Wit & Wisdom, are being field tested in 2022-23. Finally, additional research is being conducted to determine what other curricula are used within the state as well as the approaches to develop assessments for these less widely used curriculum.

All of this work is enabled by Louisiana's approach to curriculum. In addition to providing an open source curriculum, the department also rates available curricular offerings, designating each curriculum into one of three tiers. This extensive support curriculum is enabled through state legislation, specifically Act 389 of 2015, and the state’s educational board regulations, Bulletin 741 §1703 (LDOE, 2020).

The ELA Guidebooks reflects the principles\(^2\) of the highest tier rating and has been in continuous development since 2013. Use of the ELA Guidebooks is completely voluntary, as local school boards of education can choose to adopt any of the curricular materials rated as Tier I by the department, conduct a local review of materials that have not yet been reviewed, or adopt a combination of Tier I and locally reviewed materials. As noted previously, the majority of districts within Louisiana use the Guidebooks. This level of adoption represents an extraordinary level of commitment from the state house to local school boards, and all parties in between.

---

To build on this unique and enabling context, the department submitted for and received an IADA waiver in late 2018. The IADA pilot program allows an approved state to develop and implement an innovative assessment system with a subset of schools and districts, while all other schools and districts continue to participate in the state’s already existing ESSA compliant assessment system. Critically, the IADA waiver allows states to use results from the innovative assessment system within their current accountability system, removing the need for double testing. The intent of the program is to explore innovations that can be considered for statewide implementation. Louisiana’s IADA waiver’s end date was extended from 2022–2023 to 2024-2025 by the US Department of Education, and the goal of the project is to have a full suite of unit assessments that cover the major curriculum within the state in grades three to eight by the close of the project.

By being curriculum aligned, Louisiana’s innovative assessment program focuses on assessing reading and writing in ways that reflect, and is sequenced with, knowledge-rich classroom instruction. In doing so, the assessment is meant to measure the degree to which students are able to understand and build knowledge from the texts they have read and express that knowledge and understanding in writing. This approach differs from more traditional approaches to assessment, which are designed around passages that include experiences and vocabulary that students may have never encountered, potentially placing those who have fewer life experiences at a disadvantage and focusing educators on skill-building more than knowledge-building (e.g., Bransford, Brown, & Cocking, 2000; Willingham, 2017). Students are better able to comprehend what they read when they possess knowledge of the subject: “Once kids can decode fluently, reading comprehension depends heavily on knowledge. By failing to provide a solid grounding in basic subjects we inadvertently hobble children's ability in reading
comprehension” (Willingham, 2012). Similarly, “whether or not readers understand a text depends far more on how much background knowledge and vocabulary they have relating to the topic than on how much they’ve practiced comprehension skills” (Wexler, 2018).

Ultimately, the IAP is meant to support teachers in enacting the instructional shifts embodied in the curriculum and implementing the curriculum with fidelity. In this small, but important way, the IAP is meant to be one part of a larger set of activities that focus on instruction on the activities, process and approaches within high quality curricular materials.

The Innovative Assessment Demonstration Authority

The IADA allows a second assessment system to operate within a state on a temporary basis subject to specific guide rails. These guide rails are: assessment quality, comparability, statewide scalability, and demographic diversity and similarity (these guide rails are based on the requirements found in §1111(b)(2)(B) of the Every Student Succeeds Act; ESSA).

The first guide rail, assessment quality, requires that the innovative assessment be of adequate technical quality, essentially meaning it must meet all of the requirements that any other ESSA compliant statewide summative would meet. The second guide rail, comparability, requires that the assessment results between the two programs must be comparable, without specific guidance on what comparability is nor how it should be evaluated. The third guide rail, statewide scalability, requires that the innovative program be implemented statewide within five years, with the potential for a two year extension. The fourth and final guardrail, demographic diversity and similarity, requires that the innovative assessment system be piloted with schools and districts that serve a diverse set of students that reasonably approximate the diversity of the state. These guide rails mean that innovation must build on, and take place within, the constraints of current large state summative practice.
The guide rail of comparability poses one of the greatest challenges to the development of assessment programs under the IADA. This guide rail is in place in order to ensure that results from the IADA assessment program can be used within the state’s current accountability and system, and in doing so, do not result in a decrease in the expectations for student achievement established under the current statewide system (e.g., Lyons & Marion, 2016). As defined by Lyons and Marion (2016), “comparability is a judgment based on an accumulation of evidence to support claims about the meaning of test scores and whether scores from two or more tests or assessment conditions can be used to support the same interpretations and uses” (p. 8). Strictly interpreted, comparability could be understood in the same light as interchangeability across forms. Under this strict interpretation, the IADA essentially asks states to develop a different and innovative assessment program, but get the same results. However, this type of strict interpretation may also limit innovation - that “perfect agreement would be an indication of failure” (ibid, p. 7).

In contrast to this strict interpretation, comparability can also be understood as comparability of not annual student determinations, but comparability of school identifications. Instead of focusing on the key measure of the statewide accountability system, this liberal interpretation focuses on the key measure of the accountability system, school identifications. The question of comparability then becomes, “would schools whose students are participating, in part or in whole, in the innovative program, receive the same identifications if their students had participated in the statewide assessment”. This definition is much coarser, as school identification is meant to flag schools with the lowest performance as a whole or for individual

---

3 In practice, only one other state besides Louisiana has reported operational scores from an IADA program - New Hampshire for the 2018-19 school year (New Hampshire Department of Education, 2019). This may be, in part, due to challenges in addressing the comparability requirement. Currently, only three of the five states have programs active under the IADA waiver.
subgroups. Comparability analysis would then be concerned with understanding how similar identifications are under the innovative system and the counterfactual statewide system.

The IAP has taken an approach that falls in between these two extremes, but is closer to the strict interpretation. As explained later, the IAP uses a non-equivalent groups common item design to place IAP items onto the scale of the current statewide assessment, the Louisiana Educational Assessment Program (LEAP). In addition, the IAP uses the same achievement level descriptors as LEAP, and is similar or identical in terms of various aspects of its methodological implementation: the IAP uses the same item response theory (IRT) model as LEAP, the same scoring rubric, the same raw to scale score method, and similar item types. These similarities are meant to help support the comparability claim that students would have similar scale scores and achievement levels, regardless of which program they took. These similarities also are meant to help support possible statewide implementation. In keeping the many aspects of the programs the same, the hope is that the key and most substantial difference between the programs - that the IAP is rooted in curriculum - can receive the most technical and implementation support.

**Through-Year Assessment**

The IAP is also a through-year assessment program. Following Dadey, Gong, Lorie, and Marion, (2021) we define a through-year assessment program as:

1. Administered in multiple distinct sessions during a school year, and
2. Intended to support (a) the production and use of a summative determination, and (b) one additional aim.

Through-year assessment is a blanket term that was introduced to encapsulate a variety of distinct approaches to summative assessment that have become more prominent over the last

---

4 This rubric is also used in the ELA Guidebooks curriculum.
5 How similar is similar enough is addressed in a companion paper, Kosh & Dadey (2023).
twenty or so years (see Dadey, Gong, Lorié, & Marion, 2021; Gianopulos, 2019). The term is meant to be broad, and includes through-course assessment⁶ (e.g., Jerald, Doorey & Forgione, 2011), as well as approaches based on:

- Already existing interim assessment programs (e.g. Dadey & Gong, 2017).
- A few relatively unique programs that have a number of short assessments of individual standards (e.g., Clark, & Karvonen, 2021; Georgia Department of Education, 2022, see p. 4).
- New state lead programs, or substantial revision of previous statewide programs (e.g., Texas Department of Education, n.d.; North Carolina Department of Education, n.d.).

These points are grouped based on historical development and are given to illustrate the varying types of programs that fall under the term through-year. Each of these points reflects a certain development history, and each of these approaches is currently being considered by one or more states. Finally, these points overlap and are not meant to be a comprehensive summary of through-year assessment programs to date.

For a comprehensive summary of through-year assessment programs as of late 2022, see Education First (2022). In addition, Dadey, Evans & Lorié have grouped all of these emerging programs into a few key models based on (i) whether each assessment measures all content standards or a subset and (ii) whether within-year information is used in the creation of annual determinations of proficiency (2023, see Appendix A). These two features help describe the variation in through-year assessment programs. Dadey & Gong (2023) add an additional feature to these two that helps distinguish among programs - (iii) how the assessments are administered and what flexibility is afforded at varying levels of control.

---

⁶ We view through-course as more limited than through-year, and primarily concerned with dividing up assessed content throughout the year and, similarly, better allocating performance assessments during the year.
In terms of the IAP, (i) each unit assessment measures a parallel set of the state content standards, but the genre of text as well as the curriculum unit changes across each of three administrations, (ii) each unit based assessment is used within the creation of annual determinations, and (iii) districts select the sequence of unit assessments given several options of ELA guidebooks curricular units.

**Scaling and Scoring**

Within the remainder of this section, and paper, we focus on issues related to (ii) annual determinations. The production of annual determinations, i.e., achievement level classifications and supporting scale scores, has been a topic of limited research, although this appears to be changing as the field takes more interest in through-year assessment. This paper is one contribution to that body of research, which as of now consists primarily of the works of Wise (2011), Gianopulos (2019), and Dadey & Gong (2017 & 2023). Here we suggest that the production of annual determinations within a through-year assessment program can be understood in terms of a familiar framing - scaling and scoring.

First, scaling is generally concerned with the estimation of the item parameters from measurement models such as IRT or diagnostic classification models, Second, scoring is concerned with the estimation of student trait(s) such as theta values or membership in attribute profiles. These activities are unremarkable to the typical measurement practitioner. However, they do take on additional complications within a through-year program.

In terms of scaling, a key question is: when? Given that assessments are administered throughout the year, what data is used to estimate item parameters? Intentional data collection is important, and pilot or field testing designs need to pay far more attention to the complexities of administration. For example, one design for piloting or field testing is to administer all
assessments at the same time under a common person design. The parameters based on this data could then be applied across the year, assuming invariance holds. However, in some cases, like the IAP, such an approach does not work. Within the IAP, the assessments are based on instructional units. It would be a mismatch between scaling and intended use to, say, administer all of the unit assessments at one point in time, as the units are meant to come immediately after instruction. Instead, the IAP uses a set of items and associated passages as external anchors across windows, almost akin to an anchor test design used in vertical scaling applications. This allows unit assessments to be given across the year and linked.

In terms of scoring, a key question is: what? What matters? Is it the end-of-year performance? Is it within-year performance? A combination of the two? A key distinction in through-year assessment programs is that within-year performance can be factored into annual determinations. Some motivations for including within-year information include: (1) to increase measurement precision, (2) to follow a policy priority, like giving multiple opportunities to achieve proficiency or signaling the importance of the assessments, or (3) to address the depth and breadth of the standards when parts of the content domain are only assessed on one occasion. Each of these motivations are connected to various through-year assessment programs that are currently being considered. The IAP draws on all three of these motivations to some degree, but most importantly (2), that value is being placed on each of the unit assessments as an important, that each unit is critical to student learning and thus is factored into annual determinations. The approach to the creation of annual determinations, as explained later, is to pool data across units and estimate score scales using an item response theory model (IRT), essentially producing an average theta across assessment administrations. In doing so, the IAP intentionally shifts the
claim away from performance at the end of the year. For comparison, Table 1 provides a summary of other approaches that are currently being considered within the field.

Table 1. *A Summary of Possible Aggregation Approaches.*

<table>
<thead>
<tr>
<th>Score</th>
<th>Approach to Score Creation</th>
<th>Example Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRT Based Scale Score</td>
<td>Score from last assessment is used with precision improved by (1) across assessment adaptive routing or (2) using prior assessments as priors in score estimation</td>
<td>Student proficiency at the end of the year</td>
</tr>
<tr>
<td>Weighted Average of IRT Base Scale Scores</td>
<td>An average or weighted average of the scale scores from multiple assessments across the year</td>
<td>Typical student proficiency at multiple points during the year</td>
</tr>
<tr>
<td>IRT Based Scale Score based on Pooled Data</td>
<td>Treat all assessments as if they were all one single assessment, scale using IRT to produce a “composite” or average theta and thus scale score</td>
<td>Typical student proficiency at multiple points during the year</td>
</tr>
<tr>
<td>Best of end of year assessment scale score</td>
<td>The single summative score is the best of either (1) the IRT based scale score from the last assessment or (2) the weighted average across assessments</td>
<td>Best of end of the year proficiency or typical student proficiency</td>
</tr>
<tr>
<td>Best of end of year assessment scale score or weighted average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum of mastered attributes from CDM</td>
<td>Count the number of mastered attributes from CDM based mastery profiles based on standards or skill aligned assessments administered on demand</td>
<td>Number of mastered skills, potentially immediately after instruction</td>
</tr>
</tbody>
</table>

Finally, the question of what matters cannot be directly answered without considering the intended inference, or claim, to be made about students. For the IAP, this claim is that the student *typically* works efficiently with a range of fiction and nonfiction texts from, and related to, the ELA Guidebook curriculum units. The student is typically able to (1) read the texts fluently, (2) use their knowledge to make sense of them, and (3) write about them in order to gain insight into the world as well as to explore human nature and identity.

**Methods**
In this section we outline the assessment and linking design, sample and approach to scaling and scoring. One research question that is addressed within the overall approach is whether the item parameters from a pooling approach, which combines data across windows, are similar to those from separate calibrations using just from a single window. If the item parameters differ substantially, then item invariance may not hold - leading to questions about which parameters to use or even the viability of the item response theory scaling approach (see Appendix A for a list of contegenies developed for the program). This research question was investigated using data from just windows 1 and 2, to allow for course correction if needed during the operational implementation.

**Assessment Design**

The Grade 7 ELA Innovative Assessment Program consists of unit based assessments administered in three test windows throughout the school year. Window 1 was open 10/18/21 - 11/5/21, Window 2 was open 1/24/22 - 2/11/22 and Window 3 was open 4/25/22 - 5/25/22. There were five test forms for four instructional units: The Giver (Forms A and B), Written in Bone, A Christmas Carol, and Behind the Scenes. In administration window 1, school districts chose to have their teachers administer either The Giver or Written in Bone. In administration window 2, school districts administered either The Giver or A Christmas Carol. In window 3, all school districts administered Behind the Scenes. Thus, there were three resulting patterns of possible form combinations. These patterns, along with the numbers of students taking each pattern, are summarized in Table 2 below. Patterns 1 and 3 were chosen by participating districts; a small number of students (N=3) had an alternate pattern (Pattern 2) due to transferring schools mid-year. These unit based assessments were communicated to the field as end-of-unit assessments, with the intention that educators would finish instruction on each unit at or near the
assessment window. These patterns essentially act as test forms from a traditional summative state assessment program.

Table 2. *Administration Patterns.*

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Window 1</th>
<th>Window 2</th>
<th>Window 3</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Giver, Form B</td>
<td>A Christmas Carol</td>
<td>Behind the Scenes</td>
<td>2920</td>
</tr>
<tr>
<td>2</td>
<td>Written in Bone</td>
<td>A Christmas Carol</td>
<td>Behind the Scenes</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Written in Bone</td>
<td>The Giver, Form A</td>
<td>Behind the Scenes</td>
<td>1810</td>
</tr>
</tbody>
</table>

Each unit assessment has the same blueprint. Each assessment is designed to test the state content standards with the same emphasis and levels of complexity while also aligning to the content of the instructional unit. In addition, the IAP introduced new reporting categories, named knowledge, application, and synthesis. Each unit assessment is also structured by these reporting categories with students answering items in the knowledge category first, followed by application and then finally synthesis. Within the knowledge section students answer questions about the texts they read in class to show their understanding of key knowledge and skills taught in the unit. Within the application section students read a new text or texts related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit. Finally, within the synthesis section, students wrote an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit. The first two sections, knowledge and application, are administered together in one sixty minute test session. The third section, synthesis, is administered as a second sixty minute test section. In addition, there is one final section that contains a passage and item set from the LEAP 2025 assessment that serves as an external anchor, which is discussed in the next section. All unit assessments were worth 23 points and were made up of 10 items, with the exclusion of the final unit, Behind the Scenes, which was
worth 25 points and made up of 11 items. Detail is provided in Table 3 below as well as Table 1B in Appendix B.

Table 3. Overall Design of End-of-Unit Assessments.

<table>
<thead>
<tr>
<th>Reporting Category</th>
<th>Description</th>
<th>Number of Items</th>
<th>Number of Score Points</th>
<th>Item Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Knowledge and understanding of the unit text(s) based on “hot read” text¹</td>
<td>4-5</td>
<td>7-8</td>
<td>• 4-5 Multiple Choice, Multiple Select or Evidence Based Selected Response Items</td>
</tr>
<tr>
<td>Application</td>
<td>Application of knowledge of the unit topic to new, but related information and ideas based on a “warm read” text²</td>
<td>3-5</td>
<td>7-11</td>
<td>• 2 or 4 Multiple Select or Evidence Based Selected Response Items • 1 Constructed Response Item (4 Rubric Points)</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Synthesize information from the Unit Text(s), the Unit-Related Text, and/or any additional supporting text(s)</td>
<td>1</td>
<td>7</td>
<td>• 1 Essay Item (4 Rubric Points for Reading Comprehension and Written Expression, 3 Points for Language Conventions)</td>
</tr>
</tbody>
</table>

Notes: ¹A hot read is a text that students have experience with before, e.g., the novel students read during instruction. ²A warm read is a text that is related, often thematically and topically, to a hot read text. For example, the text The Giver is a hot read for students who had the Giver unit, and a warm read might be a new story about a dystopian society. ³Much of the variation in item and score points is due to the design of the final unit, which was intentionally designed with one less item in knowledge and an additional item in application.

Scaling Design

To support comparability, each IAP form contained a linking set containing items previously calibrated to the LEAP scale. In administration window 1, students took either linking set A or B; sets were administered with spiraling across students. In administration window 2, students took whichever of those two linking sets they did not take in window 1. In administration window 3, students took one of four linking sets (C, D, E, or F), which were spiraled at the student level. This linking design is shown in Figure 1 below.
All linking items were two-point evidence-based selected response (EBSR) items. EBSR items include two parts: part A in which the student must select the correct answer (1 point), and part B in which the student must select the evidence to support the correct answer in part A (1 point). If the student answers part A incorrectly, then they automatically receive a zero on part B. The number of items ranges from 7-8 and the number of score points range from 14-16. See Table 3B in Appendix B for more information on the linking sets. Linking sets were selected from a larger pool of items based on (1) the highest possible test information and (2) to achieve a balance of literary and informational passages. Linking passage sets A and B were literary and informational genres, respectively, achieving a balance of genres across the first two windows.

**Sample**

A total of 4733 students were in the final calibration sample. Because students had to complete all three forms to be included in calibration and receive scores, the N counts for
separate testing windows are equivalent to N counts by pattern, as shown by Table 2 above.

Table 4 below provides a comparison between the statewide LEAP Sample and the IAP Sample.

Table 4. Demographics of IAP Sample as Compared to Statewide LEAP Sample.

<table>
<thead>
<tr>
<th></th>
<th>Statewide LEAP Sample</th>
<th>IAP Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>42,279</td>
<td>4,733</td>
</tr>
<tr>
<td>Gender, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>20,870 (49.4)</td>
<td>2,347 (49.6)</td>
</tr>
<tr>
<td>Male</td>
<td>21,409 (50.6)</td>
<td>2,386 (50.4)</td>
</tr>
<tr>
<td>Race/Ethnicity, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>264 (0.6)</td>
<td>7 (0.1)</td>
</tr>
<tr>
<td>Asian</td>
<td>638 (1.5)</td>
<td>53 (1.1)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>18,342 (43.4)</td>
<td>1,990 (42.0)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>3,974 (9.4)</td>
<td>169 (3.6)</td>
</tr>
<tr>
<td>Native Hawaiian/ Other Pacific Islander</td>
<td>35 (0.1)</td>
<td>1 (0.0)</td>
</tr>
<tr>
<td>Two or more races</td>
<td>1,450 (3.4)</td>
<td>76 (1.6)</td>
</tr>
<tr>
<td>White</td>
<td>17,556 (41.5)</td>
<td>2,414 (51.0)</td>
</tr>
<tr>
<td>Missing</td>
<td>20 (0.0)</td>
<td>23 (0.5)</td>
</tr>
<tr>
<td>Student is migrant, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42,238 (99.9)</td>
<td>4,724 (99.8)</td>
</tr>
<tr>
<td>Yes</td>
<td>41 (0.1)</td>
<td>9 (0.2)</td>
</tr>
<tr>
<td>Student is homeless, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>41,342 (97.8)</td>
<td>4,706 (99.4)</td>
</tr>
<tr>
<td>Yes</td>
<td>937 (2.2)</td>
<td>27 (0.6)</td>
</tr>
<tr>
<td>Student is in foster care, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42,194 (99.8)</td>
<td>4,715 (99.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>85 (0.2)</td>
<td>18 (0.4)</td>
</tr>
<tr>
<td>Student has military affiliation, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>41,542 (98.3)</td>
<td>4,722 (99.8)</td>
</tr>
<tr>
<td>Yes</td>
<td>737 (1.7)</td>
<td>11 (0.2)</td>
</tr>
<tr>
<td>Student is economically disadvantaged, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>12,516 (29.6)</td>
<td>1,639 (34.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>29,686 (70.2)</td>
<td>3,094 (65.4)</td>
</tr>
</tbody>
</table>

17
Scaling and Scoring

For calibration, we pooled item response data across all forms and test administration windows into a single sparse item response matrix. In this matrix missing data corresponds to the unit assessments a student did not take during windows one and two. For example, if a student took a The Giver (Form A) in window one, his or her item response matrix would have missing values for Written in Bone (Form A) in window one. We used this sparse matrix in a single concurrent calibration in which the parameters for the LEAP linking items were fixed to their values from the LEAP program (i.e., fixed parameter calibration). Following LEAP, we used the two parameter logistic model for dichotomous items and the generalized partial credit model for polytomous items. The specification of these item response theory models can be found in Appendix C.

We implemented this scaling approach multiple times over the course of the Spring 2022 semester, starting in late February after the data from both window 1 and window 2 became available. This early run let us test the viability of the approach, and if needed, implement one of the alternative approaches outlined in Appendix A. Key in this early run was comparisons of the item parameters of the pooled approach to runs within window, allowing us to explore the invariance, or lack thereof, of the item parameters. We then ran this analysis with an early pull of window 3 data, as well as the final data. With the final data, we then examined item performance based on the differences between fixed and free calibrations as another check on invariance.

Results

Examination of Invariance of IAP Items based on Pooled vs. Separate Approaches

For the examination of pooled vs. separate approaches based on the window 1 and 2 dataset, the key outcome was the item parameters themselves. Figures 2 and 3 show the slope and intercept

| Missing | 77 (0.2) | 0 (0.0) |

---

1 flexMIRT is based on the nominal response model, whereas LEAP uses the generalized partial credit model. To conduct the fixed item parameter calibration, we converted the LEAP generalized partial credit model results to the flexMIRT nominal response parameters following the process outlined in Appendix C.
parameters, in gamma format (see Appendix C), from 40 items from the four unit assessment forms administered in Windows 1 and 2. The y-axis has the values from the within window, or separate, calibrations and the x-axis has the values from the pooled calibration. For the sake of simplicity, we have stacked the multiple item intercept parameters into a single graph. From visual inspection, the intercept parameters look markedly similar, and the slopes look quite similar, except for a few outliers with high values under both pooled and separate approaches. Additional detail on this analysis is provided in Appendix D.

Figure 2. Item Slope Comparisons.
Figure 3. *Item Intercept Comparisons.*

Ultimately, it appears that both approaches have produced similar item parameters.

**Examination of Linking Items Based on the Final Dataset**

The LEAP linking items were evaluated for inclusion into the final linking set based on comparing their parameters from LEAP 2025 to their parameters when freely concurrently calibrated with IAP items. Three LEAP linking items were removed from the anchor set calibration because their parameters are extreme or outliers when compared to their LEAP parameters, as shown in Figure 4 below.

Figure 4. *Plots Comparing LEAP Linking Item Parameters in Free and Fixed RCWE Runs*
IAP items were also evaluated. One IAP item was removed for poor IRT a and b parameters, and then removed from scoring following subject matter expert review.
Implications and Next Steps

The IAP has been designed with comparability in mind. Methodologically, the program is quite similar to LEAP in a number of respects. The linking approach performed well in terms of item parameter invariance. Ultimately, the scaling and scoring has been done in service of producing scale scores and achievement levels that both reflect the unique values reflected in the IAP program, as well as sufficiently comparable to the LEAP program (see Kosh & Dadey, 2023). There are a great number of additional investigations that can be done to better understand the implications of score creation not only at the population level, but also for subgroups.
References


Louisiana Department of Education. (n.d.). Instructional Materials Evaluation Tool for Alignment in ELA Grades K–12 (IMET) Available Online at:


Willingham, D. (2017, November 25) How to get your mind read. Available Online at:

Appendix A

Options for reporting based on potential outcomes of model fit and parameter invariance.
Appendix B

Table 1B and Table 2B show the number of items and points possible for each reporting category, based on the original test design and after item reviews (e.g., for poor model fit), respectively.

**Table 1B. Number of Items and Points by Form in Original Test Design**

<table>
<thead>
<tr>
<th>Form</th>
<th>Knowledge</th>
<th></th>
<th>Application</th>
<th></th>
<th>Synthesis</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
</tr>
<tr>
<td>The Giver, Form B</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>The Giver, Form A</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Written in Bone</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>A Christmas Carol</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Behind the Scenes</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>11</td>
<td>2</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

1. The essay item is scored on two separate rubrics (one for language conventions, and one for reading comprehension and written expression), and hence is counted as two separate items in these counts.
2. The reading comprehension and written expression item is weighted as double in the raw score total.

**Table 2b. Number of IAP Items and Points by Form (After Item Removal Decisions)**

<table>
<thead>
<tr>
<th>Form</th>
<th>Knowledge</th>
<th></th>
<th>Application</th>
<th></th>
<th>Synthesis</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
<td>Points</td>
<td>Items</td>
</tr>
<tr>
<td>The Giver, Form B</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>The Giver, Form A</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Written in Bone</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>A Christmas Carol</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Behind the Scenes</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>11</td>
<td>2</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>
Table 3B. *Number of Linking Items and Points by Linking Item Set (Prior to Item Removal Decisions)*

<table>
<thead>
<tr>
<th>Linking Set</th>
<th>Number of Items</th>
<th>Number of Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>F</td>
<td>8</td>
<td>16</td>
</tr>
</tbody>
</table>
Appendix C

Using the sparse time response matrix, we estimated item parameters using the 2-parameter logistic (2PL) model for dichotomous items and the generalized partial credit model (GPCM; Muraki, 1992) for polytomous items. These are the models used for current LEAP ELA assessments. Under the 2PL model, the probability that a student with a trait or scale score of \( \theta \) will respond correctly to MC item \( j \) is

\[
P(x_i = 1|\theta_p, a_i, b_i) = \frac{\exp [Da_i (\theta_p - b_i)]}{1 + \exp [Da_i (\theta_p - b_i)]}
\]

(1)

where \( P(x_i = 1|\theta_p, a_i, b_i) \) is the conditional probability of a correct response for examinee \( p \) on item \( i \) given \( \theta_p \), and \( \theta_p \) is examinee ability, \( b_i \) is location parameter (item difficulty), \( a_i \) is the slope parameter (item discrimination), and D is a scaling constant set to 1.7 to approximate the normal ogive.

The GPCM model produces the probability of choosing response level \( k \) over \( k - 1 \) for an examinee at a given ability level, given by,

\[
P(x_i = 1|\theta_p, a_i, b_i, d_{iv}) = \frac{\exp [\sum_{v=0}^{h} Da_i (\theta_p - b_i - d_{iv})]}{\sum_{c=0}^{m_i} \exp [\sum_{v=0}^{c} Da_i (\theta_p - b_i - d_{iv})]}
\]

(2)

where \( P(x_i = 1|\theta_p, a_i, b_i, d_{iv}) \) is the probability of examinee \( p \) obtaining a score of \( h \) on item \( i \); \( m_i \) is the number of item score categories; \( b_i \) is the item location parameter; \( d_{iv} \) is the category parameter for item \( i \), category \( v \); and D is a scaling constant given previously.

To conduct the fixed item parameter calibration, we converted the LEAP item parameters to the format expected by flexMIRT. The LEAP item parameters are:

- Parameterized in terms of the Generalized Partial Credit (GPC) model
- In the normal ogive metric
- In the slope difficulty format
- Without a location parameter

To use these item parameters in flexMIRT, we must first be converted into the format expected by flexMIRT, which is based on a constrained form of the nominal response model. That is, flexMIRT implements the GPC model by constraining the nominal response model, so the LEAP parameters must be transformed to match the nominal response model specification (i.e., gamma values). The underlying logic of this conversion is provided in Thissen, Ci & Bock (2010, see p. 59-61 for the nominal response model formula as well as relationship between the GPC model parameters and the nominal response model).
To do so we:

- Converted the LEAP threshold parameters into deviation format
- Converted the deviation format parameters to follow Bock’s original c format (i.e., the matrix c)
- Compute the Fourier T matrix following 3.35 and 3.36 from Thissen, Ci and Bock (2010, p. 59)
- Computed intermediary matrices $T'$, $\text{inv}(T'T)$, and $T'c$
- Computed the final set of intercept parameters, referred to as gamma within the spreadsheet.
Appendix D

Figure 1D. Pooled vs. Window Thetas.

Figure 2D. Window 1 vs. 2.
Innovative Assessment Program (IAP) Reporting Focus Group Grade 7

October 6, 2022
3:30 - 4:30 PM CT
Introductions

Nathan Dadey, Ph.D.
Facilitator
Senior Associate
The Center for Assessment

Ruth Caillouet, Ph.D.
Observer
Innovative Assessment Program Coordinator
Division of Assessment Content
Office of Assessments, Accountability, & Analytics
Louisiana Department of Education

Audra Kosh, Ph.D.
Notetaker
Senior Research Scientist
NWEA

David Hopkins
Observer
Assessment Research Manager
Office of Assessments, Accountability, & Analytics
Louisiana Department of Education
Introductions, Continued

Your Turn!

- Name
- School & District
- Experience with the Innovative Assessment Program
Agenda

1. Introduction & Overview (5 minutes)
2. Considering the 2021 End-of-Unit Score Report (15 minutes)
3. Considering the 2022 End-of-Unit Score Report (20 minutes)
4. Recommendations for Transition (20 minutes)

Goal: To hear from you about the IAP reports for the upcoming Grade 7 Operational Administrations.
So please, keep in mind that everything we are going to discuss is only for the Grade 7 operational administrations.
1. Introduction & Overview
Norms

Shared Norms

- Be Present
- Assume Positive Intentions
- Take Risks
- Maintain Confidentiality
Norms

Shared Norms

- Be Present
- Assume Positive Intentions
- Take Risks
- Maintain Confidentiality

Facilitator Norms

- Facilitate conversation, not lead it
- Act as a learner and a listener
Norms

**Shared Norms**
- Be Present
- Assume Positive Intentions
- Take Risks
- Maintain Confidentiality

**Facilitator Norms**
- Facilitate conversation, not lead it
- Act as a learner and a listener

**Focus Group Norms**
- No personal attributions
- Aggregated results
- Shared with key stakeholders
Recording & Notetaking

- We will be recording this session as well as using Zoom’s automatic transcription feature.
- If you are uncomfortable being recorded, please message me privately using the chat and I will suspend the recording and rely on transcription alone.
On Materials

- Today you’ll have access to draft materials that are currently under development
  - We ask that you not share these materials outside of this meeting.
  - We are working hard to make the program the best it can be, so things can and will change!
IAP: Continuing Improvement

- The work of the IAP is highly iterative - we are always trying to do better.
- Part of this iterative work involves score reporting.

2021 Grade 7 End-of-Unit Score Report

2022 Grade 7 End-of-Unit Score Report
2. 2021 End-of-Unit Score Report (15 Minutes)
Take a few minutes to review the **2021 report**. Give a thumbs up when you are ready for discussion.
Reflecting on the 2021 Report

What did you learn?
- If you administered the Grade 7 assessment operationally, what, if any, takeaways did you draw from the End-of-Unit reports?
- If not, what takeaways might you draw?
Reflecting on the 2021 Report

What did you do?
- If you administered the Grade 7 assessment operationally, what, if anything, did you do based on these takeaways?
- If not, what *might* you have done?

<table>
<thead>
<tr>
<th>Unit Title and Points Earned</th>
<th>Window 1</th>
<th>Window 2</th>
<th>Window 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of Unit Texts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this section, students answer questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit.</td>
<td>5 out of 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0 Class Average</td>
<td>4.8 School Average</td>
<td>4.8 School System Average</td>
</tr>
<tr>
<td>Application of Unit Knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this section, students read a new text(s) related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit.</td>
<td>3 out of 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0 Class Average</td>
<td>2.4 School Average</td>
<td>2.8 School System Average</td>
</tr>
<tr>
<td>Synthesis and Expression of Knowledge Across Texts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this section, students write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit.</td>
<td>0 out of 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5 Class Average</td>
<td>0.3 School Average</td>
<td>0.5 School System Average</td>
</tr>
<tr>
<td>Knowledge and Use of Language Conventions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension &amp; Written Expression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 out of 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge and Use of Language Conventions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension &amp; Written Expression</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This report has been suppressed to protect student privacy. Lowest and highest averages will be reported as a range of values rather than a specific value. If there are fewer than 10 students in a group, the average score will not be reported (NR).
3. 2022 End-of-Unit Score Report (20 Minutes)
What the report looks like after window 1.

What the report looks like after window 2.

The text for Meets/Not Meets for each Reporting Category.

- **Knowledge of Unit Texts**
  - A student who meets expectations is typically able to show general knowledge of the unit text being identified, clear ideas, main points, main ideas, and main points of the unit text in the report. The student shows strong evidence in identifying main ideas, and main points of the unit text, and provides evidence for the ideas written in the report.
  - A student who is working toward expectations needs to show more support in the knowledge of the unit text being identified, clear ideas, main points, main ideas, and main points of the unit text in the report. The student needs support in identifying main ideas, main points of the unit text, and providing evidence for the ideas written in the report.

- **Synthesis and Expression of Knowledge Across Texts**
  - A student who meets expectations typically is able to explain an essay that addresses a given prompt and synthesizes knowledge gained in the unit across multiple texts. The student’s essay includes a focus, some development and organization of ideas with supporting evidence, and a general command of the grade-level conventions of Standard English, including mechanics, usage, and grammar.
  - A student who is working toward expectations needs more support in explaining an essay that addresses a given prompt and synthesizes knowledge gained in the unit across multiple texts. The student’s essay includes a focus, some development and organization of ideas with supporting evidence, and a general command of the grade-level conventions of Standard English, including mechanics, usage, and grammar.
Take a few minutes to review the 2022 report. Give a thumbs up when you are ready for discussion.
Considering the 2022 Report

What might you learn?

- What takeaways might you draw from the End-of-Unit reports?
Considering the 2022 Report

What might you do?
- What, if anything, might you do based on these takeaways?
4. Recommendations for Transition (20 Minutes)
Each End-of-Unit test measures the knowledge and skills taught during the instructional unit. This test is just one of the many ways to measure gains in learning and academic growth. Other information, such as classroom assignments and scores on other assessments, will help determine each student’s academic strengths and needs.

**Knowledge of Unit Texts**
In this section, students answer questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit.

- **5 out of 9**
  - 5.0 Class Average
  - 5.0 School Average
  - 5.0 State Average

**Application of Unit Knowledge**
In this section, students read a new text(s) related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit.

- **3 out of 7**
  - 3.0 Class Average
  - 3.0 School Average
  - 3.0 State Average

**Synthesis and Expression of Knowledge Across Texts**
In this section, students write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit.

- **0 out of 8**
  - 0.5 Class Average
  - 0.5 School Average
  - 0.5 State Average

**Knowledge and Use of Language Conventions**
In this section, students write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit by developing their ideas with information from the unit texts and a new text.

- **1 out of 3**
  - 1.0 Class Average
  - 1.0 School Average
  - 1.0 State Average
Each End-of-Unit test measures the knowledge and skills taught during the instructional unit. This test is just one measure of how well a student is performing academically. Other information, such as classroom assignments and scores on other assessments, will help determine each student's academic strengths and needs.

<table>
<thead>
<tr>
<th>Knowledge of Unit Texts</th>
<th>Un</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>out of 9</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>Class Average</td>
</tr>
<tr>
<td></td>
<td>4.8</td>
<td>School Average</td>
</tr>
<tr>
<td></td>
<td>4.8</td>
<td>School System Average</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>IAP Average</td>
</tr>
</tbody>
</table>

Knowledge of Unit Texts
In this section, students answer questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit.

Window 1: Unit Title
Meets Expectations

A student who meets expectations is typically able to show general knowledge of the unit text(s) by identifying central ideas, main plot details, overall structure and author's purpose, important vocabulary words, and interactions between characters or ideas. The student shows some ability to provide evidence in supporting their understanding of the unit text(s).

% Meeting Expectations: Class 48% School 53% School System 58% IAP 62%
Take a few minutes to review the **2021 to 2022 Comparisons**. Give a thumbs up when you are ready for discussion.
Recommendations for Transition

The IAP program will be providing the 2022 End-of-Unit report for those involved in the Grade 7 22-23 operational administrations*, not the 2021 End-of-Unit report.

- What pros, if any, do you see in shifting to the 2022 End-of-Unit report? What cons, if any, do you see in shifting to the 2022 End-of-Unit report?
- What recommendations do you have to support the implementation of this new report?
- What else do we need to know? Questions we didn’t ask?

*Note that Grades 6 and 8 will still receive the 2021 style report.
Recommendations for Transition

The IAP program will be providing the 2022 End-of-Unit report for those involved in the Grade 7 22-23 operational administrations, not the 2021 End-of-Unit report.

● What pros, if any, do you see in shifting to the 2022 End-of-Unit report? What cons, if any, do you see in shifting to the 2022 End-of-Unit report?
● What recommendations do you have to support the implementation of this new report?
● What else do we need to know? Questions we didn’t ask?
Recommendations for Transition

The IAP program will be providing the 2022 End-of-Unit report for those involved in the Grade 7 22-23 operational administrations, not the 2021 End-of-Unit report.

- What pros, if any, do you see in shifting to the 2022 End-of-Unit report? What cons, if any, do you see in shifting to the 2022 End-of-Unit report?
- What recommendations do you have to support the implementation of this new report?
- What else do we need to know? Questions we didn’t ask?
IAP Reporting Focus Group Notes
October 6, 2022

1. Participants
   a. Teacher 1
      i. ELA curriculum specialist (first year in that role)
      ii. “New to IAP” (unclear if this meant new as of last school year or new this school year)
   b. Teacher 2
      i. Principal of junior high last year
      ii. Administered IAP
   c. Teacher 3
      i. Teachers 6th and 7th grade ELA
      ii. Administered operational 7th grade IAP last year, will be administering this year as well
   d. Teacher 4
      i. Instructional coordinator for ELA grades 6-12
      ii. Experience with IAP as teacher and administrator
   e. Teacher 5
      i. Previous English teacher
      ii. Piloted IAP in 6th and 7th grade “years back”
      iii. Now school test coordinator

2. 2021-22 Unit Report
   a. What did you learn / what were the takeaways from the report?
      i. Student report got breakdown of writing scores by LC and RCWE. The teachers report did not have that breakdown of information (just aggregated synthesis score LC+RCWE). Teachers don’t have time to go through individually and hand-import from student report.
      ii. As a teacher, the report didn’t assist in making instructional changes or figuring out weaknesses. Lacking in actionable information. Follow up Q: What would be actionable? A: the “meets expectations” in the 2022 report; very excited when saw this.
      iii. As a principal, we were able to identify that our students do better in application texts than knowledge texts. We were able to adjust our instruction to build on what we were seeing from those reports.
      iv. Participant asked if performance levels will be available faster this year or be available on unit reports? David shared current plans about trying to get faster. Ruth shared timeline of EOU report within about a month and we recognize we need to speed this up. EOY report targeted around same time LEAP scores go out.
      v. Parents have been “conditioned” to look for trigger words (e.g. 5/9, 5/7). Parents then come back to teachers to ask “is that good?” “is that bad?”
      vi. An old LEAP ELA report breaks down into literary text, informational, vocabulary - this gives more guidance regarding actionable data, concrete
skills students can continue to work on. A second participant agreed with this.

b. What did you do with the report (if administered G7 operationally) / what might you do with the report?
   i. Looked at written expression and reading comprehension and saw students were struggling with writing. Built a school-wide initiative around writing. Able to see writing improve by end of year.
   ii. One participant said nothing was actionable. Her teachers were vocal around the fact that they weren’t sure what to do with data. Because reports were so late while moving on to next unit, they went by the wayside. Didn’t find a way to have good conversations about the scores. A lot of students got scores 0/8 which left a lot of questions when comparing to work seeing in class. Previously asked a lot of questions in advisory group about what constitutes a 0 but still unclear on where zeros coming from.
   iii. By the time got report, were almost finished with next unit. Not did not much with the information on the report.
   iv. Focused more heavily on writing after seeing many scores of 0 on writing. Also thought about what changes to make to instruction next year (e.g., she knew The Giver would happen next year) but not so much the next unit.

3. Proposed 2022-23 Unit Report
   a. What might you learn from this report / what takeaways might you draw?
      i. This report gives me less information than the old report. I like that it gives “meets” and “working toward” but was under the impression there would still be a score and also the performance level. If a teacher were to see this, we are missing a lot of information. I would rather have the old information than this information.
      ii. From looking at interpretive guide, was under impression would have both score and performance level - that is useful. But if we have the descriptions without the score, we have to ambiguously guess where they are at.
      iii. Agree on above comments on scores - as a teacher I feel like I am receiving less information.
      iv. Hoping there would be three performance levels versus two.A second participant agreed (there could be a lot of kids very far or very high; trying to figure out who’s in between or close).
      v. Performance level along with score would be beneficial rather than one or the other.
      vi. For a parent, might be fine just to get the performance level. But for a teacher to make adjustments, are they a 750 mastery or a 797 mastery etc? How close is the kid from the next level?
      vii. It’s a lot of information to piece together what equals what to get the end of year score, whereas if you just use numbers it is less ambiguous.
viii. At this point David jumped in and explained why we are not giving scale scores (not precise/reliable, only based on a few items). David also clarified two kinds of scores (raw and scale) and that raw scores are not comparable. Scale scores are what David hears the participants asking for. Nathan asked David to table this conversation and shift to educator reactions (not every educator is going to be able to talk to David). David also explained unit reports don’t have an overall performance level.

ix. Participant clarified it is the raw score she would like to see, not the scale score David explained. She understood why we don’t have scale scores. David said we worry raw scores are not comparable across units.

x. Some discussion around if raw scores would be on school report (student roster report). One person said she would be ok with taking the raw scores off the student report if they were retained on school report. David said even if we provided that information it would not be a valid comparison.

xi. Participant asked for a “scale” for each window, e.g. 1-3 is below, 4-6 is meets etc. “for the adults in the room” to explain to their teachers. They understand that raw scores ranges for LEAP performance levels change year to year.

xii. Nathan summarized conversation to clarify that teachers want a scale of some sort, some number range.

xiii. Teachers said that even if the raw score ranges change each year (range of raw score that corresponds to performance level) the teachers still want to see it.

b. What might you do with takeaways from the new report?

i. If I have a student in the working toward expectations range on any category, as we worked through the next unit, I would be very conscious of my assessment of those skills. E.g. if student was “working towards” on application, would work with the student to focus on that.

ii. Use to change RTI grouping for next unit. Second and third participant agreed with this. Compare unit reports to “section quizzes” to determine RTI grouping.

4. Recommendations for Transitions

a. Pros/Cons in shift

i. Last year’s report was easier to read and had more information.

ii. Would love the two reports married: numerical values and performance levels. Second participant agreed.

iii. Pro: like that the report says some of the skills the students were assessed on during that session of the test

iv. Q from Nathan: does anyone disagree with the group recommendation to combine the reports? Recommendation was unanimous to combine the old and new reports.

b. Recommendations to support implementation (if we had to move forward with the 2022-23 design as is)
i. Have more than two performance levels, e.g. weak, moderate, strong, to relate to things they already know

c. What else do we need to know?
i. Might be more of an ADAM request, but when we got the end of year reports the kids were already in new classes and new grades. To print off new scores was a hassle; would like them just alpha ordered to print at the end of the year to send out.

5. Closing
   a. Nathan explained how item difficulty varies and impact on raw scores. Would the field have an understanding of why the score points move around (e.g. meets expectations is 4 points on unit 1 then 5 points on unit 2)?
      i. Response: it won’t be confusing to teachers; at least have the raw scores for the teacher report (teachers know this happens every year with LEAP test). Parents might not understand though.
   ii. Second participant agreed and added that you will still have some adults that don’t understand, but you can educate teachers on how to read the reports just like you educate anytime something is new.
Open Feedback Session: IAP Collaboration Groups
Score Reporting Prototypes
March 20, 2023
# Introductions

## Assessment Development Partners
- Nathan Dadey, Center for Assessment, Facilitator
- Audra Kosh, NWEA, Note Taker

## State Assessment & Accountability
- David Hopkins, LDOE
- Ruth Caillouet, LDOE
- Shantell Lee, LDOE
- Erin Hughes, LDOE
- Kathy Judy, LDOE
- Jennifer Baird, LDOE
Recording & Notetaking

● We will be recording this session as well as using Zoom’s automatic transcription feature.
● If you are uncomfortable being recorded, please message me privately using the chat, and I will suspend the recording and rely on transcription alone.
The Louisiana Department of Education (LDOE) is committed to developing the IAP in partnership with Louisiana teachers and leaders. This meeting is a part of that commitment.

Today’s Meeting

- LDOE and its partners are working to improve score reporting at the end of each unit.
- We’ve developed a number of “prototype” score reports that are wrapped up in a single html file that you can open with your browser.
- Goal:
  - Introduce you to the reports and get an initial round of feedback.
  - Find the areas we need to do more work.
Caveats

These reports reflect early and experimental thinking.

- The reports have not been subjected to the extensive quality assurance procedures that all other program are.
  - You may see some areas for improvement.
  - These reports are based on the class codes provided within the ADAM platform.
  - While they should reflect the students in your classroom, they may not be perfect. Let us know via direct chat message if something doesn’t look right.
Caveats, Part 2

These reports reflect early and experimental thinking.

- What is shown in these reports **may not be in future reports**.
- Please use the reports and supports to inform your upcoming instruction, but don’t share the actual reports.
Technology Check

- You should have received an email invite from a website called sharefile inviting you to set up a password to login and download html files.
- There should be one html file per classroom.
- Each file can be opened in your browser of choice and explored just like a webpage.

If you have issues, please chat, or direct chat, Nathan.
Meeting Structure

**Individual Review.** Take about 10-15 minutes to explore the reports.

- Write down questions, ideas, comments - whatever comes to mind.
- Raise your hand on zoom when done.

**Whole Group Conversation & Feedback.**

- What questions do you have? What did you find confusing, unclear, etc.
  - Likely, there are a number of areas that need improvement and we need to know them - the goal of this meeting is to find out what needs to be improved.
- Review each view/report.
Meeting Structure

Whole Group Conversation & Feedback.

- Which report or views do you find useful? Why?
- Are any of the results surprising? Are any of the results in line with your expectations?
- Can any of these results inform your instruction in the behind the scenes unit? If so how?
- What do you wish you had in these reports? How would it be useful?
Next Steps

We will continue to improve the score reports and hope to have a follow up with you at the end of your instruction on the current unit.

Please take a few minutes to fill out the exit survey so we can follow up with you:

(link removed)
Classroom Score Reports

Introduction

The example unit assessment, like all unit assessments in the IAP, measures how well students understood what they read in during the unit and how well they can apply and build on that knowledge. Each of the three sections of the test - (1) knowledge, (2) application, and (3) synthesis - has a specific purpose, and each is meant to help you better understand where students are at in their learning. In each section below, student scores are provided in terms of strong, moderate and weak ratings:

- Strong rating requires similar knowledge and ability of at least the Mastery achievement level on the LEAP 2025 ELA Assessment;
- Moderate rating requires similar knowledge and ability as the Basic achievement level on the LEAP 2025 ELA Assessment; and
- Weak rating is comparable to the knowledge and ability required below the Basic achievement levels on the LEAP 2025 ELA Assessment.

Note that the score needed to obtain each performance rating within a category or subcategory can vary from unit to unit.

Guidance

In the classroom with the classroom code Example Classroom within the ADAM platform, there were 30 students, of which 28 had a score on the example unit assessment in window 2. The table below summarizes performance by each section. Each row is a section, and the counts within each row are the number of students who are weak, moderate or strong within that section. So in the Knowledge section, 9 student(s) had scores at the Weak level, 3 student(s) had scores at the Moderate level, and 12 student(s) had scores at the Strong level.

<table>
<thead>
<tr>
<th>Section</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Application</td>
<td>9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Synthesis</td>
<td>6</td>
<td>13</td>
<td>5</td>
</tr>
</tbody>
</table>

For students in the moderate or weak score ranges, use the reflection questions in each section to help design or select follow-up supports. For students who are weak or moderate in multiple sections, start with Knowledge, then Application and finally Synthesis. Within each section the reflection questions should be considered in order. In this classroom, there are 12 students were in the Weak or Moderate Levels for Knowledge, 14 students for Application, and 19 for Synthesis.
Knowledge

In this section, students answered questions about the anchor texts they read in class to show their understanding of key knowledge and skills taught in the unit. The table below provides a list of students by level, along with the points each student scored within parentheses. The maximum number of points was 9 in this section.

<table>
<thead>
<tr>
<th>Knowledge Levels</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU E (1)</td>
<td>STU I (6)</td>
<td>STU B (7)</td>
<td></td>
</tr>
<tr>
<td>STU L (2)</td>
<td>STU Q (6)</td>
<td>STU M (7)</td>
<td></td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU Z (6)</td>
<td>STU N (7)</td>
<td></td>
</tr>
<tr>
<td>STU G (3)</td>
<td></td>
<td>STU V (7)</td>
<td></td>
</tr>
<tr>
<td>STU U (3)</td>
<td></td>
<td>STU W (7)</td>
<td></td>
</tr>
<tr>
<td>STU P (4)</td>
<td></td>
<td>STU C (8)</td>
<td></td>
</tr>
<tr>
<td>STU T (4)</td>
<td></td>
<td>STU K (8)</td>
<td></td>
</tr>
<tr>
<td>STU Y (4)</td>
<td></td>
<td>STU R (8)</td>
<td></td>
</tr>
<tr>
<td>STU F (5)</td>
<td></td>
<td>STU A (9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>STU D (9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>STU J (9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>STU O (9)</td>
<td></td>
</tr>
</tbody>
</table>

For the 12 students in the weak or moderate levels, consider the following reflection questions:

1. Do the students show gaps in foundational reading skills or struggle to read the grade-level content?
2. Do the students lack vocabulary knowledge or skills in deciphering word meaning?
3. Do the students have enough background knowledge about the unit topics?
4. Do the students struggle to answer higher-order questions about grade-level texts?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.
Application

In this section, students read a new text(s) related to the unit content and respond to questions and a writing prompt that measures their ability to apply the key knowledge and skills taught in the unit. The maximum number of points was 13 in this section.

<table>
<thead>
<tr>
<th>Application Levels</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU T (0)</td>
<td>STU P (5)</td>
<td>STU B (7)</td>
<td></td>
</tr>
<tr>
<td>STU E (2)</td>
<td>STU D (6)</td>
<td>STU Q (7)</td>
<td></td>
</tr>
<tr>
<td>STU F (2)</td>
<td>STU I (6)</td>
<td>STU Z (7)</td>
<td></td>
</tr>
<tr>
<td>STU G (2)</td>
<td>STU N (6)</td>
<td>STU A (8)</td>
<td></td>
</tr>
<tr>
<td>STU X (2)</td>
<td>STU U (6)</td>
<td>STU M (8)</td>
<td></td>
</tr>
<tr>
<td>STU C (3)</td>
<td>STU V (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU W (3)</td>
<td>STU Y (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU L (4)</td>
<td>STU J (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU R (4)</td>
<td>STU K (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STU O (11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the 14 students in the weak or moderate levels, consider the following reflection questions:

1. Do the students struggle with certain aspects of text complexity (central ideas, meaning of words and phrases, text structure, background knowledge needed for comprehension)?
2. Do the students have trouble applying their learning to independent tasks?

If you answered yes to any of these questions, consider the example next steps in this supporting document. This guidance is also available in a more complete document for all sections.
Synthesis

In this section, students write an extended response that demonstrates their ability to express their overall understanding of the key knowledge they gained in the unit. The maximum number of points was 11 in this section, based on 4 points from the RCWE rubric times 2 plus 3 points from the LC rubric.

<table>
<thead>
<tr>
<th>Synthesis Levels</th>
<th>Weak</th>
<th>Moderate</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>STU L (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU N (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU T (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU W (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU C (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU X (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU G (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU I (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU K (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU Q (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU U (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU Y (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STU Z (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the 19 students in the weak or moderate levels, consider the following reflection questions:

1. Do the students have trouble at the sentence level of writing?
2. Do the students have trouble understanding the expectations of writing tasks?

If you answered yes to any of these questions, consider the example next steps in [this supporting document](#). This guidance is also available in a more [complete document for all sections](#).
IAP Score Report Guidance

For students scoring in the moderate or weak score ranges within each domain, use the reflection questions to help design follow-up supports. Note that the domains and each set of reflection questions are ordered intentionally. Weaknesses in multiple domains should be addressed in order of 1) Knowledge, 2) Application, and 3) Synthesis. Additionally, reflection questions should be asked and addressed in the order provided within each domain. For example, if a student scores weak in Knowledge, you must first assess whether or not they have gaps in foundational reading skills before working on building capacity for answering higher-order questions about grade-level texts.

*If you answer “yes” to any questions, click the domain title for examples of next steps in supporting students in each area.*

| I. **Knowledge** | 1. Do the students show gaps in foundational reading skills or struggle to read the grade-level content?  
2. Do the students lack vocabulary knowledge or skills in deciphering word meaning?  
3. Do the students have enough background knowledge about the unit topics?  
4. Do the students struggle to answer higher-order questions about grade-level texts? |
| II. **Application** | 1. Do the students struggle with certain aspects of text complexity (central ideas, meaning of words and phrases, text structure, background knowledge needed for comprehension)?  
2. Do the students have trouble applying their learning to independent tasks? |
| III. **Synthesis** | 1. Do the students have trouble at the sentence level of writing?  
2. Do the students have trouble understanding the expectations of writing tasks? |
<table>
<thead>
<tr>
<th><strong>Reflection Questions</strong></th>
<th><strong>Next Steps and Resources for Supporting Students</strong></th>
</tr>
</thead>
</table>
| Do the students show gaps in foundational reading skills or struggle to read the grade-level content? | Ensure students are proficient at decoding and fluency. If issues are suspected in these areas, consult the following resources:  
  - LIFT  
  - FIRE                                                                                                                                                                                                                                    |
| Do the students lack vocabulary knowledge or skills in deciphering word meaning?       | Ensure students are proficient at meaning-making of complex, grade-level text. If issues are suspected in these areas, consult the following resources:  
  - Vocabulary Protocol and Supports for Diverse Learners embedded in teaching notes of 2018 Guidebook units  
  - Teaching Morphological Awareness                                                                                                                                                                                                             |
| Do the students have enough background knowledge about the unit topics?                 | Help students with building background Knowledge of grade-level texts.  
  - Use the Let’s Set the Context lessons for the unit.  
  - Provide independent reading opportunities of texts related to unit topics. Consult with your librarian to develop a list of available related texts. Create an independent reading time protocol for your classroom with simple reading progress tracking for students. |
| Do the students struggle to answer higher-order questions about grade-level texts?     | • Expose students to grade-level standards through both Guidebook lesson activities and an awareness of the grade-level standards; support the specific language of the grade-level standards in all parts of instruction. Click here for additional support for this area.  
  • Have students make connections between the Culminating Task and the daily lesson objectives.  
  • Use additional questions in Supports for Diverse Learners embedded in teaching notes of 2018 Guidebook units to support students in accessing the knowledge needed to comprehend the text. |
### Application

<table>
<thead>
<tr>
<th>Reflection Questions</th>
<th>Next Steps and Resources for Supporting Students</th>
</tr>
</thead>
</table>
| Do the students struggle with certain aspects of text complexity (central ideas, meaning of words and phrases, text structure, background knowledge needed for comprehension)? | ● Analyze lessons to anticipate student struggles based on data.  
● Embed appropriate scaffolds for complex text embedded in teaching notes of 2018 Guidebook units. Click [here](#) for additional support for this area. |
| Do the students have trouble applying their learning to independent tasks?            | ● Intentionally plan when release of responsibility to students will occur. (When during each lesson/section/unit will you move students from “I do,” to “We do,” to “You do?”)  
● Ensure students have ample practice with “You do.”  
  ○ For example, note when the “You do” would likely be appropriate in your upcoming unit:  
    ■ Section_; Lesson_; Slides_  
● If students struggle with the “You do,” add in additional scaffolding and assess progress through more “You do” opportunities. |
### Synthesis

<table>
<thead>
<tr>
<th>Reflection Questions</th>
<th>Next Steps and Resources for Supporting Students</th>
</tr>
</thead>
</table>
| Do the students have trouble at the sentence level of writing? | Provide students with transferable writing strategies that can be applied during independent practice.  
- Find examples embedded into the [Supports Flow Chart](#).  
- Provide sentence and paragraph frames.  
- Use the [Language Links and Language Tasks](#): strategies from The Writing Revolution or Mentor Sentence lessons.  
- Design additional grade-appropriate grammar instruction using the [Grammar Guide](#). |
| Do the students have trouble understanding the expectations of writing tasks? | - Review and ensure student understanding of [success criteria](#) for each type of writing and each “level” of success within each type.  
- Facilitate [activities](#) and provide detailed anchor charts and other visuals to support students and reinforce writing strategies and success criteria throughout each unit.  
  - Success criteria anchor chart example ([Watch from 18:30-19:54](#)) |
Grade-level ELA Standard Language and Lesson Supports

The wording of the ELA standards are designed to increase in rigor throughout the grade levels. In fact, literary techniques such as theme, characterization, word choice, structure, and point of view are asked in very specific ways from grade to grade. Let’s look at one example of how this comes to life through one ELA standard and text.

Here is an example progression of standard RL.2 (Theme/Central Idea/Summarizing) and aligned questions based on *A Christmas Carol*.

If *A Christmas Carol* is a Grade 7 unit, a teacher of this unit may reflect on how frequently this standard is asked at the correct grade level’s rigor during instruction.

1. What’s the theme or central idea? How do the details support the message? (Grades 3, 4, 6)
   - What is the theme of A Christmas Carol? How does Scrooge’s gift of a turkey and the raise he gives to Bob support the theme?

2. How do the characters respond to challenges? How does the speaker of the poem reflect about the topic? (Grade 5)
   - How does Scrooge respond when he is shown his own headstone in the churchyard?

3. How does the theme or central idea develop throughout the text? (Grade 7)
   - How does the theme of relationships being more valuable than possessions develop throughout the text?

4. How does the theme or central idea develop throughout the text? What is the relationship of the theme to the characters/setting/plot? (Grade 8)
   - How do Scrooge’s experiences help develop the theme?
   - How does the description of the Cratchit’s home help develop the theme?
   - How does Scrooge show the theme through his development in the story?

For students who struggle with gaining knowledge from the unit texts, use the possible support questions under the lesson slides. Therefore, if students were failing to understand the theme of the text in this lesson on *A Christmas Carol*, the questions below provided within the lesson are designed as in the moment scaffolds for students when dealing with theme (RL.2).
Possible Supports During the Lesson:

- If students are not providing answers similar to the Student Look-Fors:
  - Ask: “Given the lessons that Scrooge learns, what does this suggest we should value most? Why?”
  - Ask: “How does Scrooge perceive himself differently than others perceive him? Why? Does Scrooge realize how much he is missing in his life before his experiences? What does this suggest about how others may see us vs. how we see ourselves?”
  - Ask: “How do relationships sustain and support the different characters in the text? Consider the Cratchit family, Fred’s family and friends, and Scrooge’s relationship with Marley and the Ghosts.”

Next Steps

1. Use the Vertical Progressions section of the ELA Standards to review standards RL. 2-6 and RI. 2-6.
2. Make notes about words and phrases within your grade-level standards that are specific as well as distinct from previous grade-levels.
3. As you plan your lessons, anticipate additional questions you may need to ask students to help them understand the grade-level questions embedded into the slides. Ensure these questions are aligned to specific grade-level language of the standards.
4. Make note of any supporting questions already embedded into the teaching notes of the lesson slides.
Open Feedback Session:
IAP Collaboration Groups

February, 2023
Introduction & Overview
## Introductions

### Assessment Development Partners
- Nathan Dadey, Center for Assessment, Facilitator
- Audra Kosh, NWEA, Notetaker

### State Assessment & Accountability
- Ruth Caillouet, LDOE
- Shantell Lee, LDOE
- David Hopkins, LDOE
- Jennifer Baird, LDOE
Introductions

Hello

my name is

Your Turn!

- Name
- Grade level you teach
- Experience with the Innovative Assessment Program
Norms

Shared Norms

- Be present
- Assume positive intentions
- Take risks
- Maintain confidentiality

Facilitator Norms

- Facilitate conversation, not lead it
- Act as a learner and a listener

Focus Group Norms

- No personal attributions
- Aggregated results
- Shared with key stakeholders
Recording & Notetaking

- We will be recording this session as well as using Zoom’s automatic transcription feature.
- If you are uncomfortable being recorded, please message me privately using the chat, and I will suspend the recording and rely on transcription alone.
The Louisiana Department of Education (LDOE) is committed to developing the IAP in partnership with Louisiana teachers and leaders.
  ○ This meeting is a part of that commitment.
LDOE would like to partner with you in a series of ongoing feedback sessions to think through the IAP.
  ○ We want to hear from you about what is important (which is the focus of today’s session and we can schedule additional sessions as needed).
  ○ We also have topics we want to think through with you in future meetings.
Today’s Meeting

● To talk through *your* feedback on the IAP.
● Some topics that we could walk through are:
  ○ Implications of the results for instruction (e.g., essay scores)
  ○ Score reporting feedback
  ○ Instructional supports
  ○ Engaging with parents and teachers
● Where should we start?
Implications of the Results for Instruction
Implications of the Results for Instruction

- Has participating in the IAP changed the way you teach? If so, how?
- What, if anything, did you do different in your instruction based on the IAP assessment results?
  - Is there other information or results that you wish you had to inform your instructional decision making? What else could the IAP provide to help you?
- What reactions did students or their parents or caregivers have to the results?
  - What can the IAP do better in terms of communicating results to students, parents or caregivers?
Feedback on Score Reporting
Score Reporting

● What information about student performance on IAP would be most helpful to you? How would you use that information?
● What else would you like us to know about score reporting?
Score Reporting

- What information from score reports have you found most useful? How have you used that information?
- What other information would you like to see on score reports?
- What else would you like us to know about score reporting?
Instructional Supports
Instructional Supports

- The IAP provides a number of instructional resources (e.g., Reporting Guidance, Assessment Guide). Are there other resources you wish you had?
  - What, if anything, could be added to better support your instruction?
- Which areas of ELA instruction would you benefit the most from in terms of having additional instructional supports?
- What format of resources appeals to you most (e.g., lesson plans, videos of exemplar instruction, etc)?
Engaging with Parents and Teachers
Engaging with Parents and Teachers

● How have you been receiving information about the IAP thus far, and are there other ways that would be more effective to reach teachers?
● What are the best ways to share information about student performance with parents?
● How can the IAP program better reach other teachers and parents to receive feedback on the program?
We appreciate your time today and welcome more input on additional topics in further collaboration sessions.

If you have not yet completed an exit survey, Audra will message you with a link.