IAR Score Interpretation Webinar Script

Slide 1

Hello. On behalf of the Illinois State Board of Education, welcome.

This webinar focuses on *Unpacking the Illinois Assessment of Readiness Results to Inform Instructional Practices*.

Slide 2

The purpose of this webinar is to provide Illinois educators with information about how to interpret and use the results from the IAR to inform *future* instructional practices.

The goal of the webinar is to help participants understand:

- the purpose of the assessment program, along with the intended uses;
- the resources available to support the interpretation of the results; and
- how to best unpack the results to inform instructional practices.

Along the way, we'll work to answer several frequently asked questions.

Much of the information we cover may be familiar to you. Other information may be new. Hopefully, you'll discover a few new ways of thinking about the results that prove useful.

Slide 3

The first thing to understand is that the IAR is a summative assessment. It measures student learning at the end of the school year in grades three through eight, in both English language arts and mathematics.

The English Language Arts Assessments focus on how well students comprehend and write effectively when analyzing text.

The mathematics assessments focus on how well students apply skills, concepts, work through multi-step problems using abstract reasoning, and

model real world problems. Students are expected to increase their precision, perseverance, and strategic use of tools that they've learned.

Slide 4

The primary purpose of the assessment is to measure what students have learned in each grade level – that is, what they know and can do. The information accumulated from the assessment is intended to:

- assist educators and administrators in supporting student learning,
- provide information on readiness for the next level of learning, and ultimately college and career, and
- inform state and federal accountability.

The identified uses of the IAR include:

- summarizing student achievement,
- describing student performance relative to meeting the state content standards – that is, the Illinois Learning Standards, with the intended purpose of
    - supporting improvement planning, and
    - understanding what worked in instructional programs and perhaps what did not.

In short, the assessment program is meant to help illuminate trends in student achievement and helps educators reflect on the effectiveness of various school initiatives.

Slide 5

Because the IAR is a summative assessment, it's important to understand that the results provide a snapshot of how well students have mastered the state content Standards. It's designed to shine a light on trends in student achievement to inform *future* instructional efforts.

Summative accountability assessments, like the IAR, take place in the spring, and the results are often not available until the close of the school year. This, coupled with the fact that the test is a snapshot, makes diagnosing individual student strengths and weaknesses difficult. To best understand an individual

student's strengths and weaknesses we need a lot of information on that student.

Summative assessments, however, provide information on a lot of students. This means that the summary group statistics provide the richest information.

Slide 6

IAR results can help district and/or school teams evaluate which instructional programs, strategies, lessons, assignments, and tasks appear to have worked well and should be continued or expanded—and whether trends across years support that conclusion.

The results can also help teams identify approaches that may need to be adjusted or replaced, as well as students who may benefit from enrichment or additional support.

These questions should be explored with multiple sources of evidence. The IAR provides an important signal, but it should not be the only information used to make instructional decisions.

In summary, summative assessments, such as the IAR, help us reflect on the instructional opportunities provided to students and help us plan for future students.

Slide 7

Let's now review the resources available to support interpretation of the IAR.

Slide 8

There are a variety of resources available to help us make sense of the results.

The primary resource is the Illinois Learning Standards. The ILS lays the foundation for the assessment program.

Another important resource includes the Evidence Statements. The statements are derived from the Illinois Learning Standards – and we'll talk more about them shortly.

The test blueprints outline the specific design of each grade level test. There are two types of blueprints available – structural and content.

Performance Level Descriptors, also referred to as PLDs, describe what students know and can do at each performance level. PLDs provide important insight into how well a student has mastered the grade-level standards and help us better understand what the state expects in terms of mastery of the ILS.

The range of scale scores associated with each performance level, for each grade and content area, is also available.

The scoring rubrics for writing tasks are also available. For each grade level, there are two rubrics – one rubric for the Literary Analysis and Research Simulation Tasks, and another for the Narrative Writing Task.

In addition, there are two new resources include a writing guide and a checklist for the prose constructed response items, which are the ELA writing items.

The task models, for ELA, and the task types, for mathematics, provide information about the types of activities and items students may encounter when taking the test. And for ELA, there is also a text complexity infographic available.

Information about how the constructed-response items in math are scored is provided.

Items and tasks, released from previous administrations, are available through the TestNav platform, providing an opportunity for students to experience the variety of items they may encounter when testing. This also promotes familiarity with navigating the platform.

Finally, a Score Interpretation Guide provides an overview of the types of scores and score reports.

All of these resources are posted on ISBE's website for the IAR.

Slide 9

Key ELA resources for each grade include the Evidence Statements, the content blueprint, Performance Level Descriptors, and Samples to Success, which illustrate the general expectations across each performance level.

The ELA Assessment Structure resource provides additional context about the format of the assessment, including the types of passages and items, the number of items and test sections.

Slide 10

A parallel set of grade-specific resources is available for mathematics.

The Mathematics Assessment Structure provides additional information about how the assessment is organized.

For both ELA and math, these resources connect the standards, assessment design, and examples of student performance.

Slide 11

Let's briefly talk about several of these key resources, starting with the Illinois Learning standards, something Illinois educators are very familiar with. The ILS are the state's content standards and detail what students in Illinois should know, understand, and be able to do at each grade level.

The ILS serves as the primary resource for curriculum, instruction, and assessment, and as such, the framework for the assessment. All items are designed to measure student mastery of one or more content standards, and therefore understanding what is covered in the standards is essential for interpreting IAR scores.

Slide 12

Evidence Statements are another important tool that help us understand the assessment. The statements unpack the content standards to further outline the specific knowledge and skills students should demonstrate to be proficient.

The Evidence Statements guide item and task development and are designed to clarify what mastery of a standard or a set of standards looks like. They outline the knowledge and skills that an item or task should elicit, so that students can show their understanding of the content and skills inherent in that standard.

Some standards have multiple Evidence Statements. In math, some Evidence Statements reflect a combination of multiple standards. These Evidence Statements are denoted with an INT that stands for integrated.

Slide 13

Performance Level Descriptors, or PLDs, describe the *range* of knowledge and skills students typically demonstrate within each performance level.

Beginning with the updated performance levels, established in the Summer of 2025, the IAR has four performance levels: Level 4, Above Proficient; Level 3, Proficient; Level 2, Approaching Proficient; and Level 1, Below Proficient.

The PLDs increase in sophistication across levels from level one to level four. The PLDs are performance standards and help us understand how well the student performed on the assessment – they provide meaning to the scale scores. The PLDs help us understand what a student knows and is able to do relative to the content standards.

Like the Evidence Statements, the PLDs are derived from the Illinois Learning Standards.

Slide 14

Let's focus a little bit more closely on the Evidence Statements. As I mentioned earlier, the Evidence Statements are derived from the Illinois Learning Standards. They provide a description of the competencies and knowledge students are expected to demonstrate for each standard or group of standards.

The items are developed to elicit evidence of student mastery – hence, the term Evidence Statement. The statements are a key resource but, importantly, the evidence statements should not replace the ILS. Rather, they

should serve as a companion resource to augment our understanding of the expectations in the standards.

If the ILS are "what" students should know and be able to do, the Evidence Statements help us understand "the degree to which" students must demonstrate the knowledge and skills. The PLDs are also useful in this regard.

Slide 15

Evidence Statements are organized by grade level and claim.
Note that for English Language Arts, Evidence Statements have been developed for both Reading and Writing.

In ELA, under the Reading claim there are three sub-claims: Reading Literature, Reading Information, and Vocabulary Interpretation and Use.

Under the Writing claim there are two sub-claims: Written Expression, and Knowledge of Language and Conventions.

In mathematics, the claims are Major Content, Additional and Supporting Content, Mathematical Reasoning, and Mathematical Modeling.

The content and skills within each claim differ by grade. Page 4 of the Evidence Statements for each grade identifies the grade-specific content and skills associated with the claims.

Items are often aligned to multiple Evidence Statements to reflect the rich problem solving students must work through to answer each test item.

Slide 16

Here is an example from the Grade 6 Evidence Statements for Reading Literature. We see along the left-hand side that Reading Literature standards are listed, RL-one through three. And along the right-hand side we have the Evidence Statements with a number in parentheses at the end of each statement. These numbers identify each of the statements. Notice that RL 1 has one evidence statement. RL 2 and 3 have multiple Evidence Statements.

If we look at RL 3, which states: “describe how a particular story or drama's plot unfolds in a series of episodes, as well as how the characters respond or change as the plot moves toward a resolution,” we can see how that standard is unpacked via the Evidence Statements.

Items on the assessment will ask students to:

Provide a description of how a particular story or drama's plot unfolds in a series of episodes toward a resolution, or

Provide a description of how characters respond or change as the plot moves toward a resolution.

Slide 17

The mathematics Evidence Statements are also organized by grade level and claim. The subclaims are denoted with an A, B, C, or D.

A represents Major Content,

B represents Additional and Supporting Content,

C represents Mathematical Reasoning, and

D represents Mathematical Modeling.

The Evidence Statement designation includes the specific standard that it reflects, while the text outlines what students must demonstrate to show mastery of that standard.

Clarifications and content limits specific to that statement are outlined to help clarify the expectations of each standard.

Finally, the Mathematical Practices associated with each Evidence Statement are listed. Students are expected to know not only content and skills but also apply Mathematical Practices appropriately when warranted.

Slide 18

The PLDs are the key resource to help us understand the state’s expectation for mastery of each standard as well as help us interpret the results of the

IAR. PLDs provide meaning to a scale score by describing what students within each performance level typically know and can do. Across levels, they represent a progression in understanding, thinking, and reasoning. Each higher performance level describes an increased level of sophistication in knowledge and skills that students demonstrate.

Illinois provides multiple types of PLDs. Policy PLDs are high-level statements defining the state's general expectations at each performance level.

Summary PLDs describe typical performance at the IAR claim level.

Detailed PLDs describe typical performance at the standard level.

Each type serves a different purpose, but all help educators understand the expectations inherent in the ILS and how students performed.

We can and *should* use the PLDs for more than simply understanding student performance once the test results are available. PLDs should be used throughout the school year as we design or review lesson plans and tasks, activities, and assignments.

Slide 19

The Policy PLDs provide the broadest description of the four performance levels. Read across the levels from Below Proficient to Above Proficient to see the progression in the command of grade-level knowledge, skills, and reasoning.

These policy statements apply across content areas.

The Summary and Detailed PLDs add content-specific information that helps educators understand what the learning progression for a claim and standard looks like in ELA and mathematics.

Slide 20

This example shows the detailed ELA PLDs for grade 5 Reading Literary Text. The descriptors are organized by standard and performance level.

Notice how the expectations increase across the four levels. The progression reflects changes in text complexity, use of evidence, depth of explanation, and sophistication of analysis. It's important to remember that PLDs are cumulative in nature. This means that students at a high level demonstrate the skill described at that level but have also demonstrated the skills described at all lower levels.

Interpretation should also consider the focus of the standards, including Key Ideas and Details, Craft and Structure, Integration of Knowledge and Skills, Vocabulary Acquisition and Use, Knowledge of Language and Conventions, and Written Expression.

Slide 21

The mathematics PLDs are also organized by grade, content area, standard, and performance level. This example shows how performance progresses within Operations and Algebraic Thinking.

As you read across the levels, focus on the increasing complexity of the concepts, procedures, reasoning, and application students are expected to demonstrate.

The PLDs help us understand the typical progression of understanding and skill as students build and extend their expertise.

Slide 22

Some resources help inform curriculum and instruction by defining the WHAT students should know and be able to do. They inform scope and sequence, and include the Illinois Learning Standards, along with the Evidence Statements.

Other resources inform instruction and assessment and help calibrate expectations. They help define HOW MUCH and work to ensure our expectations match the intention of the standard in terms of student achievement. These resources help us design instructional tasks and activities that allow students to demonstrate mastery of the standards. They include the Evidence Statements, the PLDs, the task models, the released

items, rubrics, and student exemplars. They help us calibrate our expectations for student learning.

Slide 23

The blueprints are another important resource available to help us understand the assessment and make sense of the results. The blueprints communicate the overall test design. Two types of blueprints are available.

The structural blueprints detail the number of points by type of item for each claim, and they're arranged by section of the test.

The content blueprints detail the percentage of points that contribute to each claim and list the eligible standards and Evidence Statements to be assessed under each claim.

It's important to note that for ELA, there are two blueprints, as two forms are administered each year. One form includes a Literary Analysis Task, and the other form includes a Narrative Writing Task. Both forms include the Research Simulation Task.

Students only take one ELA test form. These two test forms are spiraled in each Illinois classroom. Some students will receive the test form with the Literary Analysis Task, while others will receive the test form with the Narrative Writing Task.

Slide 24

Here's an example of the content blueprint for the Grade 5 test form that includes the Literary Analysis Task. You'll note that we have the Illinois Learning Standard strands along the columns at the top. Then along the rows we see the subclaims, which also serve as reporting categories for the test. For the reading claim, the subclaims include Reading Literary Text, Reading Informational Text, and Vocabulary.

For the writing claim, the subclaims include Written Expression and Knowledge of Language and Conventions.

The blueprints also provide the approximate percentage of points associated with each subclaim. For instance, for Vocabulary we can expect approximately 14% of the points to come from items measuring the standards and Evidence Statements noted.

And remember that for ELA there are two blueprints – the blueprint for the form that contains the narrative writing task can be found on the IAR website. Both blueprints include the Research Simulation Task that all students encounter.

Slide 25

For mathematics, we have an example of the Grade 7 blueprint. The subclaims, which also serve as reporting categories, and the approximate percentage of points associated with each are listed as column headers. The domains from the Illinois Learning Standards for grade 7 are listed in the rows. The Mathematical Practices associated with each subclaim are also presented.

The specific content and skills associated with Major and Minor Content differ by grade level, as the standards differ by grade. Page 4 of the Evidence Statements for each grade provides a summary of the content and skills included for each sub-claim for the grade.

Slide 26

We'll now turn to IAR scores and reports.

Slide 27

Four types of scores and indicators are central to interpreting IAR results.

The primary score is the student's scale score. Scale scores are standardized scores that account for the difficulty of the collection of items on each test form. Scale scores allow for comparisons to be made across *the same grade level* in a content area, regardless of the test form or the year the test was taken. For example, in ELA, whether it was the Literary Analysis Task or the Narrative Writing Task test form, or the year in which a student took the test, that is whether they took the test in 2025 or 2026, the scale score has the same meaning. A scale score of 739 on the fifth-grade ELA test is necessary to

reach the Proficient performance level. Any fifth-grade student who achieves a scale score of 739 on the grade 5 ELA test, regardless of test form or year taken, has demonstrated the same level of performance.

Scale scores allow us to identify trends in student achievement both within a grade and across school years for each grade level. Every year, the grade level test forms are built very carefully and intentionally to ensure that what it takes, in terms of student achievement, to achieve each performance level is consistent.

IAR scale scores range from 650 to 850 for both content areas – ELA and math. The scale score needed to reach the Proficient performance level differs slightly across each grade and content area. In other words, while 739 is the cut score for Proficient in grade 5 ELA, other grade levels may have a different cut score for Proficient. The scores associated with each performance level for each grade and content area are posted on the IAR webpage, as well as included in the Score Interpretation Guide, also found on the IAR webpage.

Because each grade level measures standards with expectations unique to that grade level, what students must demonstrate to achieve each performance level may differ slightly from one grade to the next (for example from grade 6 to grade 7). The Student Growth Percentile, or SGP, is provided to describe changes in student achievement across grade levels. We'll revisit the SGP in a minute.

In short, we can compare scale scores from the same grade level, such as grade 5, within each subject across school years (2025 to 2026 to 2027). However, we should ***not*** compare scale scores across grade levels (grade 5 to grade 6 to grade 7).

Performance levels classify performance based on the scale score. The four levels are Above Proficient, Proficient, Approaching Proficient, and Below Proficient. Performance levels provide meaning to the scale score – they allow us to understand what the typical student at each level knows and is able to do based on their command of the grade level standards.

Each performance level represents a range of student achievement. The scale score ranges associated with each performance level are called cut scores and are available on the IAR website as previously mentioned. These ranges were established by Illinois educators during a Summer 2025 standard setting and will not change. Given that achievement expectations were revisited in Summer 2025, comparison to IAR results earlier than 2025 should not be made. Rather, the 2025 IAR results established a new baseline.

Students who achieve a lower performance level demonstrated less mastery of the grade level standards than those who achieve the higher performance levels. The performance level descriptors help us understand the differences across levels.

In ELA, two additional scale scores are provided for each claim – Reading and Writing. Reading scale scores range from 10 to 90, and Writing scale scores range from 10 to 60. Note that no performance levels are associated with these scale scores.

Readiness Indicators classify performance for each subclaim relative to students who were proficient on the overall content-area assessment. The three readiness levels are High, Middle, and Low.

Students who demonstrate higher readiness performed well on the items measuring the sub-claim. Students who performed lower were less successful with the items measuring that sub-claim.

The Readiness Indicators provide information about student achievement of the knowledge and skills in each sub-claim. The sub-claims can be considered the building blocks of each content area test.

The Student Growth Percentile, or SGP, is a measure of how much improvement a student has made in a content area from one year to the next. Student Growth Percentiles result from comparison of a student's scale score to other academically similar students from across the state. Academically similar students are those students with similar prior scale scores.

Percentiles range from 1 to 99. Generally, the 50th percentile is considered average. To calculate the Student Growth Percentile, a student must have a scale score for a minimum of two consecutive years in each content area. In other words, the student must have a scale score for the current year as well as the previous year to calculate the SGP. If more years are available, the Student Growth Percentile will use those additional years.

The Student Growth Percentile provides an indication of how a student is progressing in learning compared to other academically similar students from across Illinois.

Slide 28

Four reports are available at the district level. These reports are designed to help district administrators understand the trends in student achievement.

The intended audience for each report is provided, as is a description of each report. It's important to remember that any report containing individual student information is secure. This includes the District Performance Level Summary Report, as the data are reported by student subgroup and may reflect small numbers of students, making it possible to identify some students.

The piece of information that is unique and only available at the district level is the School Content Standards Roster File. This is a very nuanced data file and provides the average percent of points earned for the collection of standards under each strand, for ELA, or domain, for math, of the Illinois Learning Standards for each student who took the assessment. Caution should be used when reviewing these data as they are based on percent correct and do not take into consideration the difficulty of the individual questions for each strand or domain.

Slide 29

There's a variety of reports produced at the school level. The intended audience for each report is provided, as is a description of each report. It's important to remember that any report containing individual student information is secure. This includes the School Performance Level Summary

Report, as the data are reported by student subgroup and may reflect small numbers of students, making it possible to identify some students.

While all reports include the state results as a benchmark, including the Individual Student Report, several reports also include the district results.

The series of reports at both levels – the school and district – are designed to help teams drill down to encourage reflection on the instructional opportunities provided to students. We'll spend some time discussing a protocol for reviewing the reports.

Slide 30

Let's begin with the Individual Student Report, or ISR.

Slide 31

Page 1 displays the student's overall scale score and performance level, together with the standard error of measurement. The standard error, referred to on the report as the margin of error, reminds us that every test score is an estimate rather than a perfectly precise value. We would expect the student's true score – if we could sample all knowledge and skill, to fall with the range depicted. On the day this student took the test, the student's score was 737, which falls at the top end of the Approaching Proficient performance level.

The state, district, and school average scale scores are provided as points of comparison. This helps a parent understand how their student performed relative to other Illinois students in the same grade.

For students in grades 4 through 8, the report also provides the Student Growth Percentile when sufficient prior-year information is available. As previously mentioned, at least two consecutive years of student scores are necessary to calculate the SGP. This student's SGP is 17. This means that the student's academic progress was the same or better than 17% of Illinois fifth graders who have the same academic history on previous administrations of the IAR.

While the student's scale score places the student at the top of the Approaching Proficient performance level, nearing Proficiency, the student's

SGP indicates that student made less academic progress than 83% of their academically similar peers.

Note that all reports shown in this webinar have been redacted to protect the identities of students, schools, and districts.

Slide 32

Page 2 displays the High, Middle, or Low Readiness Indicator for each subclaim. These indicators help identify broad patterns within the content area. The Readiness Indicators can be considered signals but should not be considered diagnostic.

The Readiness Indicator compares how the student performed on the set of items related to the subclaim, such as Informational Text, to how students, who achieved proficiency on the overall test, performed. In the example report displayed, the student performed similarly to students who were Proficient on the test on the items measuring the comprehension of Informational Text. The student's Readiness Indicator for this claim is indicated as High. The same is true for the claims of Vocabulary and Knowledge and Use of Language Conventions. For Literary Text, the student's performance on the items measuring this claim was close to students who were Proficient on the test, while for Written Expression, the student's performance was lower than students who achieved Proficiency in ELA.

The report also provides suggested next steps and resources and displays the student's predicted Lexile measure for ELA or Quantile measure for mathematics, along with the associated range – because no score is ever perfectly precise.

As always, interpret these results together with classroom evidence and other measures of student learning.

Slide 33

We'll now review the district-level reports.

Slide 34

The District Performance Level Summary provides a districtwide view for the selected grade and content area. It includes the number of students tested, the average scale score, the percentage of students at each of the four performance levels, the percentage Proficient and Above, and the average SGP (for grade levels 4 – 8).

The report also presents results for student groups. Use caution when interpreting results based on small numbers of students and protect the confidentiality of student-level information.

This district assessed 137 fifth-grade students, who had an average scale score of 727, slightly lower than the state average of 729. Both of these scores fall in the Approaching Proficient range; and we can see the highest percent of students (35%) at the school achieved this performance level. Overall, 36.5% of the school's fifth-grade students achieved Proficient or Above Proficient (31.4% and 5.1%, respectively). The average SGP for the school was 47, which was slightly below the state average SGP of 50.

Slide 35

The District Summary of Schools allows district leaders to review the performance of the state, district, and individual schools within the district. It displays, both graphically and numerically, the percent of students achieving each performance level. The number of students tested and the average scale score is presented, along with the percent of students scoring at each Readiness Indicator level (high – middle – low) for each subclaim.

The report can be used to identify patterns that merit further investigation but should not be used to compare schools without considering differences in context, student populations, and other evidence.

This report may be helpful to Curriculum and School Improvement staff, as the distribution across the Readiness Indicators may help shed light on the effectiveness of instructional programs for each subclaim. The standards and Evidence Statements eligible to be assessed under each subclaim are outlined in the test blueprint.

Slide 36

The District Evidence Statement Analysis displays the average percent correct for each assessed Evidence Statement at the state and district levels. Evidence Statements are ordered from more difficult on the left to less difficult on the right. Remember that it is important to review the student count associated with each Evidence Statement, found on page 2 of the report, before drawing any conclusions.

Differences between district and state performance may suggest areas to explore at both the district and school levels, but the results must be interpreted considering the student count and the broader instructional context.

Slide 37

Beginning on page 2, the District Evidence Statement Analysis provides the student count for each Evidence Statement, along with the related standard, domain or strand, and item type.

Student counts matter because not every student receives an item aligned to every Evidence Statement. Focus interpretation on statements supported by larger student counts and avoid drawing conclusions from very small samples.

For example, for the first Evidence Statement (5.C.7-3), we see only 2 students in the district took a form that included items that measured this standard. This is insufficient information to make any statements about how well students in the district mastered the content and skills inherent in this Evidence Statement.

Slide 38

We'll now turn to the school-level reports.

Slide 39

The School Performance Level Summary presents the number of students tested at the state, district, and school levels, along with the average scale score, performance-level distribution, percentage Proficient and Above, and average SGP for grades 4 through 8.

Results are also disaggregated by student group. Because subgroup counts may be small, treat the report as confidential and use appropriate caution when interpreting patterns. Remember that summary statistics are more reliable when there are more students. A general rule of thumb is that care and caution should be used when interpreting data on less than 15 students.

Slide 40

The Student Roster provides the average scale score for the state, district, and school as well as the percentage of students by Readiness Indicator for each subclaim. Note that the average scale score for Reading and Writing is also displayed. Finally, the average SGP is provided.

Overall, students in the school demonstrated stronger achievement in reading than in writing. This can be discerned by noting that more students demonstrated lower readiness in writing than in reading.

The readiness display uses shades of purple. Darker purple indicates stronger readiness, and lighter purple indicates lower readiness. For the district and school, the percent of students achieving each readiness indicator is presented, followed by student-level results. Student level results also include the Lexile Measure and SGP. For math reports, the student's Quantile Measure is provided.

Slide 41

The Evidence Statement Report is presented by order of difficulty for all students across the state of Illinois. The average percent correct for items measuring each Evidence Statement are graphed.

The first Evidence Statement, on the left-hand side, proved to be the most difficult for Illinois students.

Statements listed on the right-hand side represent those on which students across the state performed more successfully.

State performance on the items measuring each Evidence Statement is represented by the blue boxes. The performance of students at the school is

represented by the orange triangle, and district-level performance is represented by the green circle.

Note the information is also given for Writing (Expression and Knowledge) as well as by the type of Prose Constructed Response.

Look for areas in which the school's performance appears relatively strong or weak, but before drawing any conclusions, check the student count for the identified Evidence Statements on page 2 of the report. This will help ensure the pattern identified is meaningful.

Slide 42

The information presented on page 2 of the report is important in helping to identify specific Evidence Statements to consider. This table provides the number of students who experienced items associated with each Evidence Statement. Any time we interpret data, we want to look at the number of students associated with those data.

Notice that for the most difficult Evidence Statement, which measured RI 5.9.1, no students at the school experienced the item. This means that the forms administered at the school did not contain any items measuring that Evidence Statement. Use the district and state information as a proxy for such an Evidence Statement. Future students in the school will likely see items measuring this Evidence Statement in future years.

We'll use this table to identify those Evidence Statements with data based on larger numbers of students. We know that 112 students in 5th grade in this school took the assessment. Any Evidence Statement where all or at least half of the students in the school took items associated with that Evidence Statement is a good area to focus on.

Statements with higher student counts have sufficient data to allow a reliable, or stable, indication of student achievement related to that statement. Always use caution when interpreting data with small numbers.
Finally, the Illinois Learning Standard and its associated domain and item type are also provided for each Evidence Statement in the table. This information can aid interpretation. For example, we know that students at the

school demonstrated weaker performance in Writing. Notice that the Evidence Statements that proved to be the most difficult for students are the prose constructed response, or PCR. These are the items that measure student writing skills.

Slide 43

Now, let's turn to unpacking the IAR results.

Slide 44

Several principles are important when reviewing IAR results.

Comparisons across forms and years are appropriate within the same grade and content area. The IAR is not vertically scaled, so comparing scale scores across grade levels does not yield meaningful information. Use the Student Growth Percentile to describe year-to-year progress relative to academically similar students.

Each performance level covers a range of achievement, so the scale score helps indicate where performance falls within the assigned level.

Readiness Indicators compare subclaim performance with the performance of students who achieved Proficient or Above Proficient on the overall assessment. They are useful signals, but they should be interpreted with the complete set of available evidence.

Slide 45

As we begin to unpack the IAR results, we want to use the score reports to identify areas where students performed well and areas where they performed less well. We want to look for patterns and trends in student performance to help guide interpretation.

Remember that summative data is limited in what it can tell us about individual students. The strongest information comes from the summary data where we have many students as opposed to a single student.
It's also important to remember that all data sends a signal, which has to be interpreted. Best practice will use other sources of information about student achievement to help inform interpretations.

The main purpose of the assessment is to identify trends in student achievement and to help us reflect on the instructional opportunities we provided students throughout the previous school year. As we review those instructional opportunities, the PLDs can help us.

Slide 46

There are 7 steps to follow in the unpacking protocol. One way to think of the process is to consider a funnel. We'll start broadly and then try to narrow the information to identify a few areas to further investigate and reflect upon.

Let's quickly review the 7 unpacking steps and then we'll use actual reports as an example to illustrate the steps.

The first step is to review the District or School Performance Level Summary. Note the distribution across the four performance levels for all students and student groups and identify broad areas of success and opportunity.

Next, review the School Roster, if analyzing district reports, or the Student Roster, if analyzing school level reports. Examine the distribution across the three readiness levels for each claim or subclaim – at the group level (district or school). Darker purple represents stronger readiness, while lighter purple represents lower readiness.

Select one claim or subclaim for deeper examination and review the same grade-level reports from prior years to determine whether the pattern persists.

Slide 47

Step 3 is to review the District or School Evidence Statement Analysis Report.

For areas of success, note the Evidence Statements toward the right side of the report, where students generally earned a higher percentage of points. For areas of opportunity, examine statements toward the left, where the items were more difficult.

Before drawing any conclusions, check the student count beginning on page 2. The count represents the number of students who encountered items aligned to each Evidence Statement. Use caution with low counts and focus on statements supported by larger numbers of students. A good rule of thumb is to void Evidence Statements where fewer than approximately half the students in the district or school encountered items measuring those statements.

Slide 48

Step 4 is to reflect on the instructional opportunities associated with the selected Evidence Statements and standards. Consider when and how the standards were taught and the assignments and tasks students completed.

Focus on both the summary and detailed PLDs for the identified Evidence Statements. Review the expectations across the four performance levels and consider where lessons, assignments and tasks fell along that continuum. You can also use the Samples to Success, released items, rubrics, and student exemplars to examine whether local expectations were calibrated to the state's expectations.

As you complete this step, consider what worked well and what may need adjustment.

Steps 5 and 6 involve looking across two or three years, across relevant grade levels, and across other evidence of student performance to determine whether a pattern is consistent.

Step 7 is to decide what adjustments may benefit future students and to develop a plan for implementation and monitoring.

Slide 49

Now, let's see the unpacking steps in action using school-level reports for grade 5 ELA.

Slide 50

For Step 1, we begin with the School Performance Level Summary for Grade 5 ELA. (Note that this same process can be followed using the District report and can be used with math.)

Review the number of students tested (112) and the distribution across Above Proficient (about 9%), Proficient (about 20%), Approaching Proficient (about 38%), and Below Proficient (33%).

Key data points for the school are circled. Note that the average scale score for the school is 722, which is lower than both the state (742) and district (731). This scale score falls in the Approaching Proficient range for Grade 5 ELA. The scale score range for this performance level runs from 700 to 739. Remember, each performance level reflects a range of achievement. A scale score of 739 is the threshold score needed to be classified as Proficient in this grade and content area.

Next review the distribution of students across each of the performance levels, including the percent of students who achieved Proficient or Above: 28.6%. Note that the largest group of students (38%) reached the Approaching Proficient performance level. Adding the 33% of the school's student's scoring Below Proficient to the percentage scoring Approaching Proficient, a total of 71.4% of the students are not proficient. We need to take a deeper look at what might be going on.

Also review the average SGP for the school, which is 50. This matches the state and is slightly higher than the district average SGP of 49.

Remember that summary results provide the starting point for inquiry. We are not yet drawing conclusions; we are identifying patterns that warrant a closer look.

Slide 51

For Step 2, we examine the School Student Roster. At the district level, Step 2 involves reviewing the District Summary of Schools.

In ELA, scale scores are provided for the two major areas of focus: Reading and Writing. These two scores serve as a barometer of performance, along with the more granular readiness indicators for each subclaim. These scale scores are derived from the items that contribute to each claim and reflect how students in the district or school performed on the sets of items related to either Reading or Writing. For Reading, the range of the scale is 10 to 90, while for Writing, the range is 10 to 60.

For this school, the average scale score for Reading was 40, slightly below the district and state. For Writing, the average scale score was 20, also slightly below the district and state average. In both cases, the school's average scale score falls in the middle of the scale range, which is 80 points for Reading, and 50 points for Writing. The Writing scale score of 20 is a bit lower than the middle of the distribution.

The circles and arrows draw attention to readiness patterns across the ELA subclaims.

Darker purple signals stronger readiness, while lighter purple signals lower levels of readiness. In this example, the pattern suggests that writing-related performance may warrant deeper investigation. 66% of the school's students demonstrated a low level of readiness in both Written Expression and Writing Conventions compared to their peers, statewide, who achieved Proficiency on the grade 5 ELA IAR.

We also note that fewer students demonstrated a low level of readiness in Reading Information (23%), than in Reading Literary and Vocabulary (30% and 33%, respectively). Overall, students in the school demonstrated stronger readiness in Reading than in Writing.

We'll carry these observations into the Evidence Statement Analysis.

Slide 52

For Step 3, we turn to the School Evidence Statement Analysis.

We see the state, school, and district performance indicated by the blue boxes, orange triangles, and green circles, respectively.

We know that because the report is ordered by difficulty (or average percent correct) at the state level, the Evidence Statements at the left were generally more difficult for students across the state, while those toward the right were less difficult – students across the state achieved a higher percent correct for items measuring the Evidence statements on the right.

For instance, for second item in the list (on the left), circled here in red, it appears from the position of the orange triangle that students at this school achieved an average percent correct of approximately 10% for items measuring RI5.7.1.

Conversely, for the 24th item, circled here in blue, we see that students at the school achieved an average percent correct of approximately 58% for items measuring RL5.3.1.

Given the fact that ELA has two blueprints that are spiraled across the state and within schools and classrooms, before going too far, we should review the student counts on the next page.

Slide 53

Page 2 of the Evidence Statement Analysis report provides more detailed information about the items. For each Evidence Statement, the corresponding Illinois Learning Standard is listed, along with the domain, item type, and student count.

Looking at the column furthest to the right, we note that there are several Evidence Statements for which no students at the school encountered items so the student count is 0. Some Evidence Statements may only have a very small student count, perhaps 3 or 4. Often this occurs when students need an accommodated test form.

For Evidence Statements with few, if any students, we can use the state and district as proxies. While we don't have information about how students at this school performed, we know the likelihood that these Evidence Statements will be assessed in the future is high.

Because student counts are so crucial to interpretation of the report, we will use the table on this page to identify those Evidence Statements with data based on larger numbers of students. We know from the School Performance Summary that there are 112 5th grade students in this school. We also know that we want to focus on Evidence Statements where at least half of the students in the school encountered items associated with that Evidence Statement.

For the second Evidence Statement in the list, RI 5.7.1, we note that it related to the domain of Reading Informational Text , and that all 112 students in the school encountered items associated with this evidence statement. Remember that page 1 showed students had an average percent correct of approximately 10% for this Evidence Statement. This Evidence Statement requires students to answer a question or solve a problem using information from multiple print or digital sources.

We note a couple of other Evidence Statements to consider – RL 5.9.1 (which asks students to provide a comparison and contrast of stories in the same genre on their approach to similar themes and topics) and RI 5.6.2 (which asks students to provide an analysis of multiple accounts of the same topic, noting similarities and/differences in points of view). Both of these Evidence Statements have a sufficient student count (59 and 112 students, respectively) and together, these represent the three most difficult Evidence Statements for the school.

Interestingly, as we can see in the Item Type column, all three of these statements were measured, in part, by the prose-constructed response item type. These are the Writing prompts and we have already noted a higher percent of students who performed at the lower level of readiness in Writing.

Slide 54

The Evidence Statement Analysis also provides information on the different opportunities students have to write. We see from this example (the red circle on the left) that performance in Written Expression and Knowledge of Language and Conventions is low for the school.

Across the three writing opportunities, performance on the Narrative Writing Task and Literary Analysis Task is similar, while performance on the Research Simulation Task is somewhat weaker, shown in the red circle on the right.

Because every student completes the Research Simulation Task, it is a reasonable area to investigate further.

Slide 55

Next, we connect the report with the ELA Task Models, which can be found on ISBE's website for the IAR. Now that we have decided to focus on Evidence Statements related to the Research Simulation Task prose constructed response that have a high student count, we identify evidence statements fitting these criteria, which include RI 5.7.1, RI 5.6.2, and RI 5.1.1.

Together, these expectations call for students to draw from multiple sources, analyze similarities and differences in accounts or points of view, quote accurately, explain explicit meaning, and draw inferences.

The report indicates that some of these statements were among the more difficult for students at this school (most difficult, third most difficult, and eighth most difficult, respectively), which helps us refine the focus for instructional reflection.

Slide 56

As the next portion of this step, we want to take a look at the PLDs to review those expectations across the achievement continuum for our identified Evidence Statements and the standards they are associated with.

For expectations at the claim level, we can review the Summary PLDs. For the Research Simulation Task specifically, the claims to review include Reading Informational Text, Written Expression, and Knowledge of Written Language and Conventions.

For expectations related to the identified evidence statements and their associated standards, we can review the Detailed PLDs.

Slide 57

In reviewing the Summary PLDs, we want to focus on the claims associated with the Research Simulation Task. In this case, we have pulled out the PLDs for Claim 2: Reading Informational Text; Claim 4: Written Expression; and Claim 5: Knowledge of Written Language and Conventions.

Note that the Summary PLDs for these specific claims have been recreated here for illustrative purposes, and the formatting will look slightly different in the Summary PLDs on the IAR website.

We will start by looking at the Summary PLDs for Claim 2: Reading Informational Text, as the three Evidence Statements we identified earlier all fall under this claim. We see that the Summary Proficient PLD for Claim 2 states: Explain the main idea and how the author's use of facts, examples, and text features support it. Describe relationships such as cause and effect or how ideas build across a section. Compare two texts for similarities and differences in how information is presented.

We also want to review the Summary PLDs for Claim 4: Written Expression. At the Proficient level, students write with a clear focus, supporting their ideas with reasons, facts, and descriptive examples. They use transitions and organize their writing.

Slide 58

Next, we can turn to the Detailed PLDs, which show us the expectations at the standard level for the three Evidence Statements we identified. In our example, the PLDs associated with all three are found in the Reading Informational Text section.

As with the Summary PLDs, the Detailed PLDs for these specific examples have been recreated here for illustrative purposes, and the formatting will look slightly different in the Detailed PLDs on the IAR website.

We see that the Detailed PLDs at the Proficient level state that students should be able to draw on key information from multiple sources to efficiently answer questions (RI 5.7), analyze the way points of view influence multiple

accounts of the same topic of event (RI 5.6), and use quotes to express explicit ideas and draw logical inferences (RI 5.1).

Slide 59

Because we identified the Writing claims as an area of opportunity, we also want to review the Detailed PLDs associated with them.

In the Detailed PLDs for Claim 4, We see that at the Proficient level, students should be able to state and support an opinion with logically ordered reasons and relevant information and explain a topic with organized ideas, facts, and definitions.

Slide 60

We then want to repeat the process for the other Writing claim, Claim 5: Knowledge of Written Language and Conventions.

Here, we see that at the Proficient level, students should be able to use appropriate grammar and varied sentence structure; use grade-level conventions to support meaning and clarity; and more
We can use the Detailed PLDs for all three claims in developing opportunities for students to practice the skills listed.

We can use the Detailed PLDs for all three claims in developing opportunities for students to practice the skills listed.

For example, we could consider providing opportunities for students to practice writing tasks where they must incorporate these claims together, such as drawing on information from multiple sources or analyzing points of view, developing their points with clear examples and details, showing their command of language (for example through word choice and sentence structure) and using appropriate grammar.

Slide 61

We now move on to Step 4. We are going to think about the instruction provided to students in the areas covered by the Research Simulation Task.

We can use the grade 5 Task Models on the IAR website to help us reflect on the types of work students may be asked to do in a Research Simulation Task: analyzing relationships among concepts, analyzing the role of illustrations, analyzing multiple accounts, and analyzing an author's use of evidence.

With these task focus in mind, our grade level teams can review the opportunities students received to use informational texts, make claims, and support those claims with evidence.

Slide 62

Next up, we will look at the released grade 5 Research Simulation Tasks, along with the associated rubrics and scored student exemplars to provide another lens on the expectations.

Our goal here is to compare local lessons, tasks, assignments, and feedback with the Evidence Statements, task models, released items, rubrics, and exemplars, asking ourselves whether students had sufficient opportunities to demonstrate the targeted skills and whether local expectations were calibrated appropriately.

This review is not about teaching to the test. It is about understanding the state's expectations and combining that information with professional knowledge of students and instruction.

Slide 63

For Steps 5 and 6, we want to look for corroborating evidence. How did prior grade 5 cohorts perform? Is there a trend? If so, we have additional evidence that adjustments in instructional opportunities may be warranted.

We also want to look at how students in adjacent grades (3 and 4) performed on related Research Simulation Task expectations to see if we spot similar trends.

Finally, we want to look at what other formative, interim, classroom, or student-work evidence are available to determine if that evidence supports the results we are seeing.

When we have completed our review, we should discuss the findings with our colleagues. A pattern that appears across years, grades, and sources of evidence provides a stronger basis for action than a result from one report alone.

Slide 64

For Step 7, we want to identify which instructional plans and student assignments worked well and which adjustments might benefit future grade 5 students.

We will develop a plan of action for the upcoming school year and identify how student learning will be monitored, specifying the checks that will be used, when they will occur, and how the team will respond to the evidence. These steps are best completed collaboratively by district, school, and grade-level teams, together with individual reflection.

Slide 65

In closing, here are a few concluding thoughts.
While there are many factors that influence student achievement, one factor that has significant influence is the opportunity to learn that we give our students.

It's important to remember that summative assessments such as the IAR are designed to illuminate trends in student achievement. While the students on which the data are based have moved on to the next grade level, if we reflect on the instructional opportunities we provided those students, we can be better prepared for future students.

And no single assessment can encapsulate everything a student has learned in a school year. Summative assessments are a necessary part of the continuous improvement cycle, and they can prove useful when used to consider trends in student achievement.

Slide 66

Thank you. We hope you found this webinar helpful.

If you have any questions, please don't hesitate to contact the ISBE Assessment Department at the email address given.

And always remember, that the primary role of any assessment is to inform teaching and learning. Thank you again.