

# ACCELERATE LEARNING WITH BETTER FORMATIVE ASSESSMENT

Reliably assess academic  
progress and address learning  
needs with Mastery View  
Formative Assessments





## Checklist: Selecting a Reliable Classroom Assessment

Over the course of the pandemic, educators and students alike had to adapt to extraordinary changes in teaching and learning. Unfortunately, for many students who struggled with hybrid or remote instruction, the abrupt shift has led to gaps in their learning. Addressing those gaps is top of mind for educators, but pinpointing areas where learning may be unfinished can be a struggle.

The pandemic amplified the issue, but it's something many teachers face in any school year. When it comes to identifying students' learning needs, they may not have a complete understanding of what their students know and don't know, and whether they have mastered grade-level learning standards. State summative assessments don't provide teachers with timely data to address learning needs in the moment. While formative assessments can, many teachers don't have the time or simply don't have the skillset to develop assessments that can reliably determine mastery.

On top of that, while measuring student progress may be more important than ever, there's a perennial clamor about test anxiety and over-testing students. So, the question remains; how can we be sure that teachers are efficiently gathering and using assessment data—in the moment—to drive learning forward?

As a result of the pandemic, federal stimulus investments have been made to address learning loss among students, including the administration and use of "high-quality assessments that are valid and reliable, to accurately assess students' academic progress and assist educators in meeting students' academic needs, including through differentiating instruction."

The result is an unprecedented opportunity to put highly effective tools into the hands of educators and allow them to gauge student mastery quickly and accurately and adjust classroom instruction accordingly.

To meet each student's academic needs, today's teachers require effective and efficient assessments that provide the insights needed to determine where students are in their learning. Here are five things to consider when evaluating formative assessments for classroom use.

- ✓ **Are the assessments aligned to our state learning standards?**
- ✓ **Have the assessment items been tested and used to collect score data to support their reliability?**
- ✓ **Do the assessments allow us to take periodic checks for understanding, in line with the concepts being presented in the classroom?**
- ✓ **Are the assessments based on a research-backed, statistically driven assessment model?**
- ✓ **Do the assessments meet stimulus funding requirements for high-quality assessments?**





## Mastery View Formative Assessments: Designed for High Reliability with Fewer Questions →

To address and meet this need, Instructure has introduced **Mastery View Formative Assessments**—short, expertly-developed assessments that reliably assess academic progress. Mastery View Formatives are standards-aligned assessments that use a unique psychometric algorithm—called the Diagnostic Classification Model (DCM)—to deliver highly reliable indicators of student mastery with less testing time. While appearing to students like a familiar online quiz, Mastery View Formative Assessments reliably determine students' mastery of learning standards with the fewest questions required, delivering meaningful, actionable data that teachers can use with minimal impact on instructional time. That means teachers have more time to teach, students have more time to learn, and assessment becomes a seamless part of the instructional cycle.

In this guide, we'll dive into the three characteristics that make the Mastery View Formative Assessments unique:

1. An innovative, statistically-driven psychometric algorithm
2. Expert item development rooted in a proven methodology
3. Delivered within an assessment management system that gives educators the immediate, actionable data they can use.

## What Do We Mean by Highly Reliable Assessments? →

A test is typically administered because a decision needs to be made, and the results of the testing help inform that decision. Assessment results need to be interpreted and a case needs to be made that those interpretations are valid for that situation. Validity, therefore, is an argument that is made about our assumptions, based on test scores. The case is made that the instrument used does, in fact, measure what we hoped to measure. One kind of support for the validity of the interpretation is that the test is consistent. This is known as the reliability of the test. Reliability is a quantification of how much error there is in a test score. Reliability is a necessary condition for validity. A reliable test may or may not be valid, but an unreliable test can never be valid. Reliable assessments have consistent—and thus, dependable—results. Given there is error in all assessments and the conclusions about achievement need to be made with confidence, it becomes very important to minimize error and increase reliability. In an unreliable test, students' scores consist largely of measurement error. To ensure test scores reflect more than just random error, assessments must have solid measures of reliability.



# The Basics of Psychometric Models →

An assessment is meant to provide students and teachers information about how students are performing. Psychometric or test theory models help explain the relationship that exists between actual test scores and estimated performance. Additionally, test theory models are necessary to help better understand the measurement process related to an assessment, and how it is impacted by sources of error.


Two well-known psychometric models that estimate a student's score or ability are classical test theory (CTT) and item response theory (IRT). Assessments based on these psychometric models provide performance information by assigning scores along a continuum—such as a scale from 0-100. Students are placed on a continuum by score or ability to provide information about where students fall compared to either other students or against a standard. In order for an assessment to accurately and reliably determine where a student's ability is along that continuum, and to place all students reliably with limited error, there needs to be enough test items on an assessment to separate out students and place the students in order. Unfortunately, assessments which require a large number of items to reliably place a student on a continuum tend to be lengthy and administering these tests takes time away from classroom instruction.

Given time constraints and the number of items required, reliable subscores are difficult to obtain when using these models. Because tests need to have a sufficiently large number of items, measuring multiple learning standards would require too many items to be practical. For example, a teacher testing students over the course of a math unit that included 5 learning standards may need 20 or more items per standard to measure each with some degree of reliability, totaling about 100 items. Because of these

constraints, instead of a 100-item test measuring 5 standards, a teacher will opt for a 20-item test measuring the overall unit. Both psychometric models, CTT and IRT, have limitations measuring multiple constructs on a single test resulting in a unit total score and not a specific standard score.

Although a lot of statistical energy goes into ensuring a score is reliable, it is often just a number and additional decisions still need to be made. This is a two-stage process. Standard setting is the methodology used to define levels of achievement or proficiency and the cutscores corresponding to those levels. A cutscore is simply the score that serves to classify the students whose score is below the cutscore into one level and the students whose score is at or above the cutscore into the next and higher level. In tests of educational progress, such as those required for federal accountability, students are typically classified into one of three or four achievement levels, such as below basic, basic, proficient, and advanced. As a result, with four achievement levels, three cutscores need to be determined. In a K-12 context, decisions based on cutscores affect not only individual students, but also the educational system. In the latter case, group test results are summarized at the school, district, or state level to determine the proportion of students in each proficiency category. As part of federal legislation, for example, a school's progress toward educational goals is expressed as the proportion of students classified as proficient. This method orders students and then classifies them based on a total score. Often the first question that is asked by the educator and test taker, is why the score or classification was obtained. Unfortunately, psychometric models often cannot provide this level of detail. In many situations, another, longer assessment is needed to provide diagnostic information.





So, how can we  
reduce the amount  
of time students are  
spending taking tests,  
while still obtaining  
reliable and actionable  
assessment data?

# An Innovative Approach to Psychometrics: the Diagnostic Classification Model →

The goal of the Mastery View Formative Assessments is to provide teachers with a high degree of confidence about whether students have mastered a learning standard—and to provide that level of confidence and reliability with the fewest number of test questions. While the assessments are formative (not interim), they provide a greater degree of specificity and accuracy than most teachers are able to achieve with their own classroom-created tests.

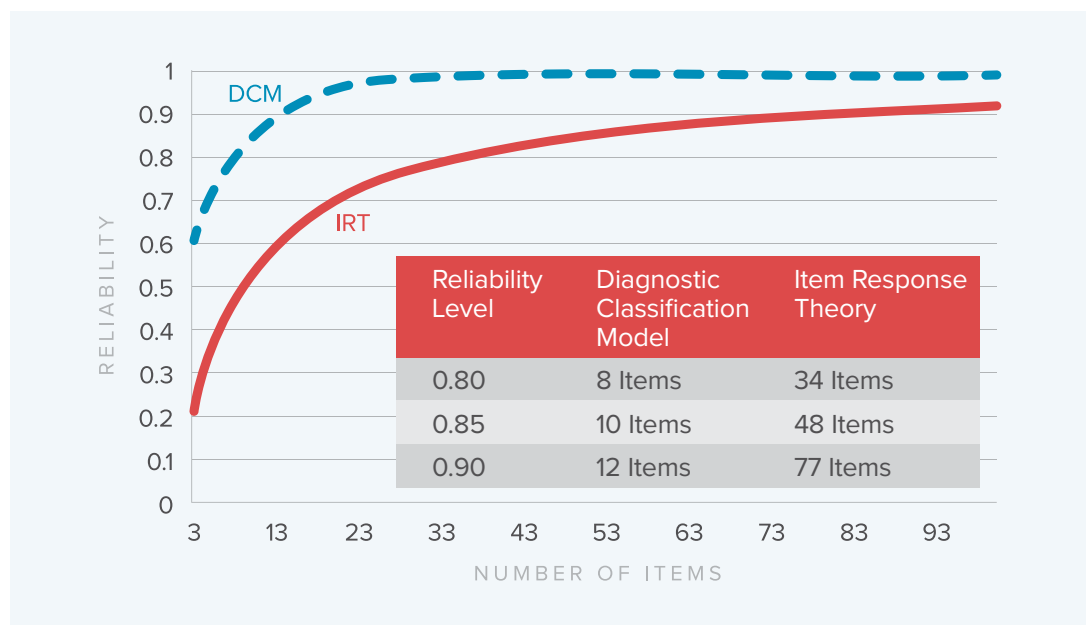
The Mastery View Formative Assessments employ a version of the statistically-reliable Diagnostic Classification Model (DCM), based on research by Jonathan Templin, Ph.D., E.F. Lindquist Chair and Professor of Educational Measurement and Statistics at the University of Iowa. In part, the DCM is designed to address some

of the limitations of traditional psychometric models. The goal of DCMs is to provide a reliable measure of student understanding based on mastery of a skill or set of skills. DCMs achieve this goal by providing profile estimates of probabilities of mastery, placing students in groups rather than on a continuum.

Often, most assessment decisions are about classification as described above, and DCMs are designed specifically for those classification decisions, performing the function of standard setting/cutscores and the judgment error involved in that process.

Since it is easier to place a respondent into one of two groups (masters and non-masters) or three groups (mastery, near-mastery and remediation, which is the model used in the Mastery View Formative Assessments) than to locate them along a continuum, far fewer items are required. The example on the previous page (Templin) illustrates the number of items needed to acquire a certain level of reliability.

With few items required for each assessment, more learning standards can be tested on a single test and single standards can be assessed with fewer questions. For example, a teacher testing students over the course of a math unit that includes 5 standards may need 8 items per standard to measure each with some degree of reliability, totaling about 40 items. Now a student will receive a profile specifying which of the 5 standards have been mastered and which have not. This provides students a direct link between what and why.



“

DCM-based assessments provide actionable information that can be used to help students learn while not burdening students with excessive testing. The combination of more focused assessments and an innovative psychometric model allows teachers to use shorter, standards-based assessments throughout the academic year as a reliable gauge of standards mastery.

JONATHAN TEMPLIN, PH.D.

”

# Putting the Diagnostic Classification Model to Work in the Classroom →

Administering classroom assessments that are short, reliable and fit seamlessly into the instructional cycle is important, because given the time constraints that teachers and students face, assessments compete against instruction and other valuable classroom activities. Teachers need the information provided by assessments to guide future learning, and this information needs to be immediately actionable. With shorter assessments, teachers can assess one or multiple standards in one sitting, providing

reliable and actionable feedback directly related to what has been taught: the learning standards!

With DCM-based assessments, teachers have actionable data immediately so they can adjust their instruction accordingly. The Mastery View Formative Assessments use the DCM algorithm to quickly and reliably set up groupings (specifically: mastery, near-mastery, remediation) based on item responses, allowing teachers to identify which students to group together for remediation or enrichment—with a strong rationale.

Administering DCM-based assessments and using the results is all about implementing a balanced assessment system and calibrating a formative practice. Teachers are formatively assessing students throughout the grading period, whether by asking questions during the lesson, having exit tickets, or practice assignments. All teachers can benefit from having an additional, external data point with a known reliability, to calibrate their formative practice and evaluate the rigor and appropriateness of the questions, activities, or assignments they use.

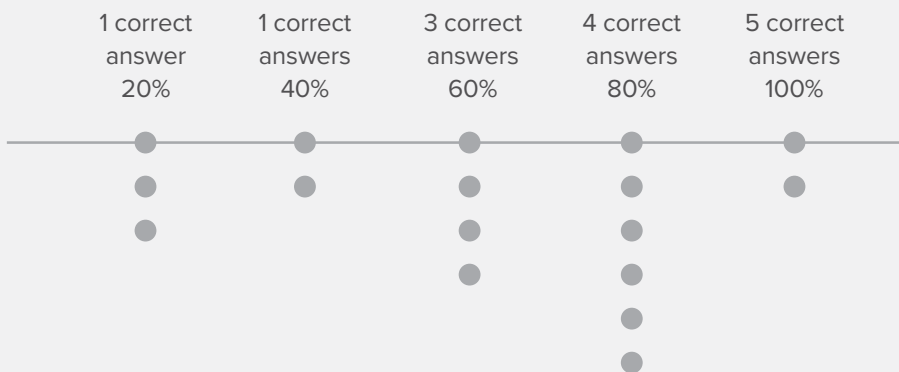
## Do you really know if your students have mastered the standard?

With the research-backed DCM algorithm, we can gauge with a greater degree of precision and clarity which students are masters, which students are near-masters, and which students require remediation.

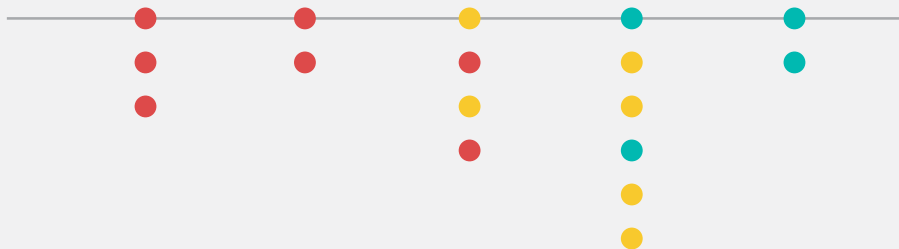
The DCM does not score each question equally, but reflects which questions a student gets right or wrong - not how many questions the student answered correctly or incorrectly.

- Mastery
- Near-Mastery
- Requires Remediation

### TYPICAL CLASSROOM ASSESSMENT: 17 STUDENTS



### MASTERYVIEW ASSESSMENT WITH DCM SCORING: 17 STUDENTS





A young child with dark hair, wearing white headphones, is focused on writing on a piece of paper with a purple pen. The child is wearing a grey hoodie over a yellow shirt. The background is slightly blurred, showing what appears to be a classroom setting with bookshelves.

“

With larger standardized tests, often you don't get the results immediately; they're too big, they're too broad, they're too generic. They don't give teachers that data that they need right away. DCM-based assessments give us the opportunity to understand where students are today, and what teachers can do to impact their learning moving forward.

**RICH WILSON**

Director of Curriculum and Assessment,  
Olathe Public Schools, KS

”

## Diving Deeper: Not All “3 Correct out of 5” are Created Equal! →

In DCM-based assessments, items are designed to separate students into categories as efficiently as possible. Each item is specifically designed to represent different difficulty levels of the standards. The specific item(s) answered incorrectly by a student will determine their probability of ending up in one category or the other. Some students who get 3 out of 5 correct can end up in different categories because of the different items they got correct. Here’s an example:

701 third grade students took a mathematics assessment. Using the accepted formative practice of 80/50 break points there were 133 students who scored 3/5, or 60%. With other methods, all 133 of these students would be categorized as near mastery. But, by using the DCM model, which looks at item difficulty and not just percent correct, it’s found that 64 of these 3/5 students are categorized as mastery for this standard. There were 51 students categorized as needing remediation, and only 18 students in the near mastery category. Here’s a sample set of the results:

%	PE/PF	Item 1	Item 2	Item 3	Item 4	Item 5	Mastery Level	Traditional
65.0%	13/20	0	0	1	1	1	Mastery	3/5
45.0%	9/20	0	0	1	1	1	Mastery	3/5
45.0%	9/20	0	1	1	1	0	Mastery	3/5
80.0%	16/20	1	0	0	1	1	Mastery	3/5
90.0%	18/20	0	0	1	1	1	Mastery	3/5
45.0%	9/20	1	0	1	1	0	Near-Mastery	3/5
55.0%	11/20	1	1	0	1	0	Near-Mastery	3/5
90.0%	18/20	1	0	1	1	0	Near-Mastery	3/5
70.0%	14/20	1	0	1	1	0	Near-Mastery	3/5
60.0%	12/20	1	1	0	1	0	Near-Mastery	3/5
35.0%	7/20	1	0	1	0	1	Remediation	3/5
35.0%	7/20	1	1	1	0	0	Remediation	3/5
35.0%	7/20	1	0	1	0	1	Remediation	3/5
40.0%	8/20	1	0	1	0	1	Remediation	3/5
40.0%	8/20	1	1	0	0	1	Remediation	3/5

## Expert Item Development →

The Mastery View Formative Assessments are developed using the Instructure methodology for creating rigorous, high-quality, standards-aligned test items by the team of curriculum and assessment experts which authors the Mastery View Predictive Assessments. A hallmark of the Instructure assessment development process is the attention to item rigor and learning standards alignment, certainly an attribute that contributes to the strong validity of the assessments.

The Instructure assessment development methodology include the following protocols:

- Items feature Webb's Depth of Knowledge (DOK) levels 1, 2 and 3
- Items include a range of item types including selected response, evidence-based selected response, and technology-enhanced items (TEIs)
- ELA items include sets which share the same reading passage and offer passages with appropriate Lexile® levels (using MetaMetrics' Professional Lexile Analyzer); and text complexity readability ratings
- ELA assessments utilize authentic reading passages, specifically literary and informational texts, and a variety of genres

During the item-development process, items are checked for content by at least two subject area specialists. Each item is then examined for conformity to learning standards for each grade level by content experts. Next, the items are reviewed for cultural and gender bias, sensitivity and cluing. This process may occur in multiple steps, often looping the same item through the process several times to ensure that

CURRICULUM SPECIALISTS DEVELOP ITEMS

CURRICULUM SPECIALISTS EVALUATE ITEMS

MULTI-PHASE REVIEW & EDITING PROCESS BEGINS

CONTENT AREA DIRECTORS APPROVE

FINAL GRAMMAR & FORMAT EDITS APPROVED

FINAL ITEMS APPROVED

reviewer bias does not affect modifications. Finally, each item is reviewed by a technical editor for spelling, grammar, and usage conventions. Each test and item sent through this review process is tracked using internal procedures documenting the reviewers and dates of review.

Items are rated for rigor using a rubric developed by each team of curriculum specialists (i.e., Math, ELA). The rubric will include content-specific vocabulary, the domain of the question, and the thinking skill (DOK) of the question, and difficulty level.

The Mastery View Formative Assessments are designed to require students to recall information, apply concepts and skills, make judgments, and justify thinking. All test items found on the assessments are calibrated to the DCM, which allows for highly reliable assessments with fewer questions and less testing time.



“

Assessments developed by the Instructure team have “provided our administrators and teachers with timely and precise data to guide instructional decisions.

**GREG PACZAK, PH.D.**  
Director of Research and Development,  
Madison County Schools, MS

”

# Immediate, Meaningful Data to Identify Learning Needs →

Mastery View Formative Assessments are available exclusively through **Mastery Connect**, an assessment management system designed to show student progress in an intuitive, visual way so educators can make informed instructional decisions. With Mastery Connect, teachers can administer MasteryView Assessments and immediately see student mastery levels of key learning standards. The Mastery Tracker tool provides an instant, visual view into three levels of student mastery – mastery, near mastery, and remediation. This stoplight approach lets teachers quickly identify learning needs and adjust teaching in real-time and in context. Mastery Connect also provides the ability to align Mastery View Formatives to Curriculum Maps, and use the assessments alongside a formative assessment item bank when teachers want to create additional quick checks.



## The Mastery View Formative Assessment Collection

Available in selected states beginning in the 2021-2022 school year.

### MATH

- Mastery View Formative Assessments for Math are available for Grades 3-8 and Algebra I.
- Math Assessments cover approximately 15 priority Math standards; the collection includes 2 assessments per standard, providing approximately 30 assessments per grade.
- Math Assessments offer a combination of multiple choice and alternate item types, including some technology-enhanced items (TEIs).
- Items will assess the standards for mathematical practice, including reasoning, modeling and attention to precision.
- Calculator-active and calculator-inactive items will be presented, depending on the standard being assessed.

### ENGLISH LANGUAGE ARTS

- Mastery View Formative Assessments for ELA address reading comprehension using informational and literary texts and a variety of genres.
- Each Assessment is passage-based and includes a selection of questions associated with the authentic text selections. The Lexile® level of reading passages is evaluated using the MetaMetrics' Professional Lexile Analyzer.
- The length of the ELA assessments will vary by grade (slightly longer for Grades 6 through English I & II); the collection will include approximately 8 assessments for Grades 3-5 and approximately 4 assessments per grade for Grades 6 through English I & II).





## Summary

As schools and districts begin to rethink how formative assessments fit into the learning process, reliable classroom assessments like the Mastery View Formative Assessments allow teachers and schools to:

- Administer formative assessments that evaluate student progress with the rigor, alignment and reliability of state and standardized tests
- Give high-quality tests using the fewest number of test questions to determine mastery
- Provide immediate, meaningful data to teachers so they can design interventions and remediation
- Implement a “through-course” model that supports districts’ instructional scope and sequence

Simply put, with Mastery View Formative Assessments, educators can accurately assess student mastery and use meaningful data to meet each student’s academic needs, something perhaps more important now than ever before.

“

With assessments based on the DCM—such as the Mastery View Formative Assessments—we can get to the heart of what students know and don’t know and give teachers the reliable and actionable data they need.

JONATHAN TEMPLIN, PH.D.

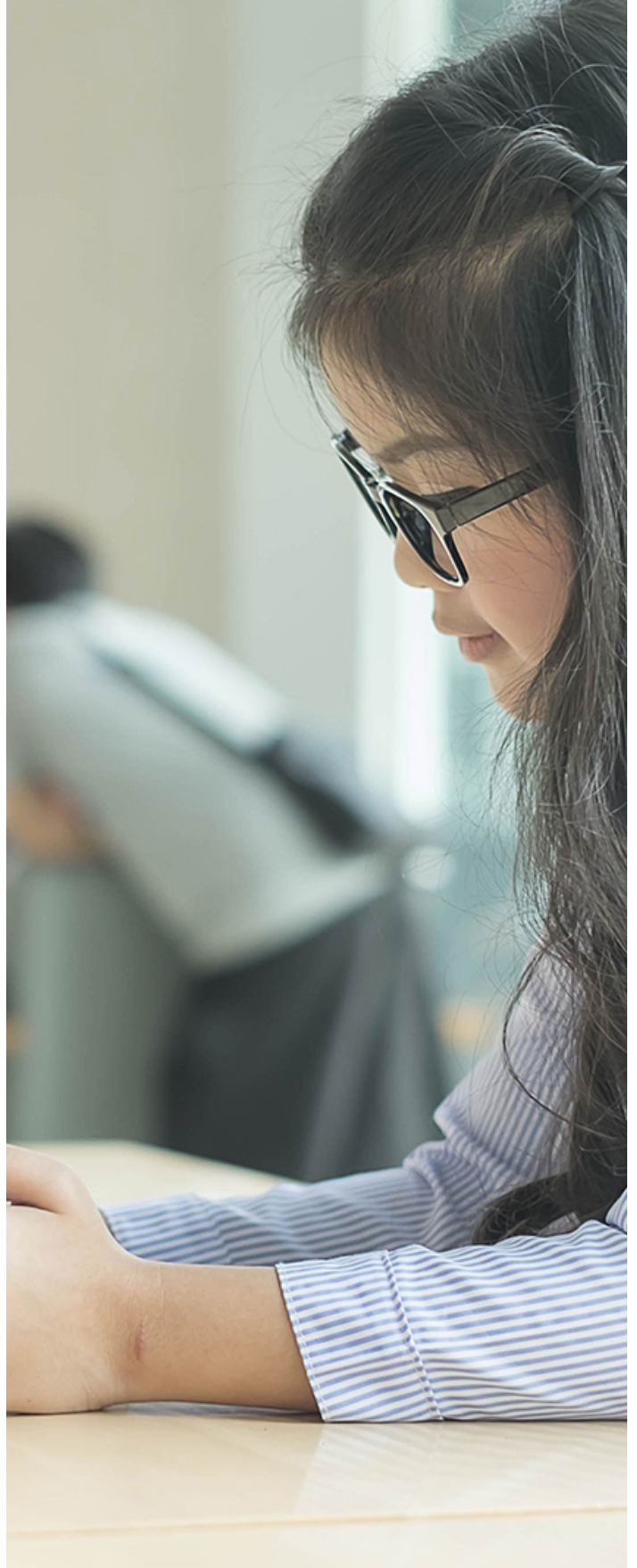
”





**Mastery by Instructure** brings together the assessment tools and quality content educators need to implement a successful assessment program.

Learn more about how our assessment offerings—including the **Mastery View Formative Assessments** and the **Mastery Connect** assessment management system—drive learning forward and positively impact student success.





# Powering the World's Smartest Classrooms.

Instructure is an education technology company dedicated to helping everyone learn together. We amplify the power of teaching and elevate the learning process, leading to improved student outcomes. Today, Instructure supports more than 30 million educators and learners at more than 6,000 organizations around the world.

The Instructure Learning Platform makes learning more personal and student success more equitable by delivering the solutions you need to support, enrich, and connect every aspect of teaching and learning. The Instructure Learning Platform includes:



**Canvas by Instructure:** providing a robust foundation for teaching and learning—both in and out of the classroom—for all students, while seamlessly integrating with other learning tools.



**Mastery by Instructure:** bringing together the assessment tools and quality content educators need to implement a successful assessment program that drives learning forward.



**Elevate by Instructure:** improving the use of data and technology to help solve district challenges and give educators the information they need to make impactful decisions.



**Impact by Instructure:** helping teachers and students adopt educational technology to promote deeper engagement with learning.

Learn more at <https://www.instructure.com/k-12>.