Why Data Matters

and how to find the right tools to improve student performance





CTB Smart Center eGuide Series Assessment Data Management

Why Data Matters

and how to find the right tools to improve student performance

Do you have a solid assessment plan?

Without one, you run the risk of wasting time and money purchasing and administering the wrong tests, receiving inadequate reports or misinterpreting results. With a well conceived assessment plan, you can be sure to collect data that's accurate, relevant, and actionable towards your efforts to focus instruction to meet the needs of all your students.

This e-guide will help you build your plan and undestand why data mastery is essential to success as teachers, administrators, and educational leaders. You'll learn the various uses of data in education and receive a wealth of resources for continuing down the path of mastery.

Table of Contents

Transforming Our Schools	page 3
Data's Growing Role	page 7
Blueprint for Data-Based Planning	page 1
Getting Immediate Feedback	page 2
Choosing Assessments You Can Trust	page 3
Engaging the Family	page 3
CTB Solutions	page 3
About CTB	page 4
Summary	page 4
Glossary/Resources	page 4





A better understanding of assessment data can transform our schools

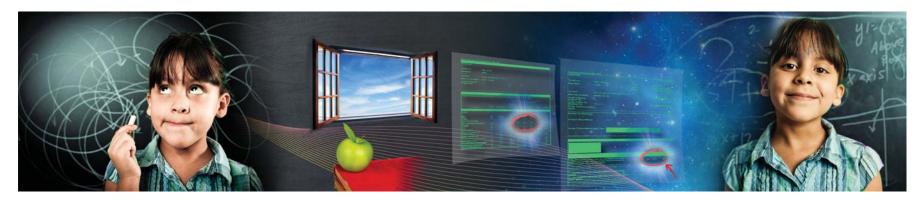
Quality educational assessments not only capture data on learning and achievement, but are fundamentally changing the way teachers teach, students learn, and parents engage in classrooms around the world. They provide essential information to guide genuine education reform and improvement and allow us to fairly hold ourselves accountable for providing students with the quality education they need to compete and thrive in their local community and in the world economy.

Administering all the assessments in the world does no good if they do not measure what was intended or yield meaningful reports, or if the data are misinterpreted and erroneously point our schools in the wrong direction. To make sure we have support for our learning, there is a wealth of instruction and professional development out there to support our data literacy, as well as assessment literacy. Like learning any new subject, understanding how to accurately collect data and properly interpret test results can be daunting at first, but is well worth the effort.

This e-guide will help you understand why data mastery is essential to success as teachers, administrators, and educational leaders; the various uses of data in education; and provides a wealth of resources for continuing down the path of mastery.







Why look deeper to measure achievement?

Too often we compare student scores to overall state averages and stop there. If performance is near average, we carry on as usual and just focus on our low achievers, trying to bring them up to standards.

But there is so much data gathered every year at every school in America that goes untapped! Imagine what we could achieve with a better understanding of its meaning.

Assessment results are most valuable when the testing information is used to support instruction in the classroom. They can also be used to help make critical administrative decisions that affect students and their families, such as eligibility for placement into special programs. Additionally, the data can be used to evaluate the effectiveness of the curriculum and the instructional program, and to keep the public informed of progress being made by schools.

In many schools, teachers are not trained in assessment data interpretation and haven't taken the time to develop these skills on their own.

A well-planned, comprehensive assessment program built upon a quality test provides valuable information about students that can support teaching in many areas, including:

- Evaluation of student progress
- Needs assessment
- Instructional program planning
- Curriculum analysis
- Program evaluation
- Teacher and principal evaluation
- Class grouping
- Administrative planning and direction

And higher achievement translates into a brighter future for our students. It can positively impact individual students as well as parents or guardians, groups or classes of students, schools, and entire communities:

- Fewer dropouts
- More college-bound graduates
- Greater success at jobs and in life
- Greater self esteem
- Less dependence on social services/more economic independence
- A tradition of achievement for generations to come





A paradigm shift: From effective schools to effective students

The more we learn about how students learn, the more aware we become that every student has different needs. There is no one-size-fits-all solution. It's no longer enough to teach the same lesson the same way to the whole class, letting those who don't master the material fend for themselves, and simply move on to the next lesson.

With today's sophisticated benchmark and interim assessments—such as CTB/McGraw-Hill's Acuity®—you can create short, targeted tests based on the material you just taught and learn instantly who "got it" and who didn't. Then you can remediate, offer practice, and retest to improve the performance of more students before moving on.

There are no more excuses for allowing students to fall between the cracks. When you begin turning borderline students into achievers, you will change their mindset from "I can't do this, and I hate school," to "I'm as good as anyone else. What else can I do?" Dropout rates will drop, scores on summative tests will rise, funding will increase, and teachers will have new tools to use to prevent further failures.

This turnaround can only occur, however, in the presence of valid and reliable data.

"Data Aversion" is Curable

We ask our students to learn something new every day. We teach them to ask questions and base their conclusions on facts. We teach them all sorts of evidence-based principles: mathematic equations, scientific methodology, and then we ask them to support their premises when they write essays. For teachers and administrators, learning how to use data to inform instruction is no more or less challenging. However, in terms of student motivation, achievement, graduation rates, and success in college and career, the rewards are incalculable.



Data's Journey- District Level





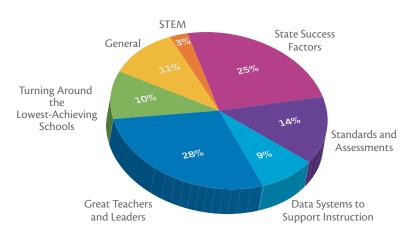


Today, data has a bigger role than ever before

These are unprecedented times in the history of education. Race to the Top (RTTT) requirements ask states to reform their education policies around four specific areas:

- Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy
- Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction
- Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most
- Turning around our lowest-achieving schools

Race to the Top funds are awarded on the following basis:



To support your state's eligibility to receive funds under RTTT, each of the areas above must be documented using data of one sort or another. Proof of improved performance must be presented, teacher effectiveness must be documented, student outcomes and achievement must improve, achievement gaps must close, graduation rates must increase, assessments that measure common standards must be adopted, and data systems must be implemented that support instruction. All this requires data.





Common Core State Standards require new ways to analyze data and report assessment results

Assessments that measure student achievement and progress on the Common Core State Standards will require new thinking, new technologies, new approaches to content, and new ways to analyze data and report assessment results. One big change is the shift from basic factual knowledge to an increased emphasis on performance and depth of knowledge. These next-generation assessments, as envisioned by the two Common Core assessment consortia, must address a new set of design considerations and criteria.

Innovative Item Types

Common Core assessments will include performance tasks, technologyenhanced items, and constructed-response items in addition to the traditional multiple-choice items to measure the depth, rigor, and complexity of comprehension required by the Common Core.

✓ **Performance tasks** help uncover deeper levels of student understanding by measuring a student's ability to think through a complex problem that may have more than one correct answer. They call for students to apply their knowledge within an authentic learning experience, which may take anywhere from minutes to days to complete. These items will be scored using rubrics based on the cognitive skills being assessed.

- ✓ **Technology-enhanced items** and assessments provide instant feedback to students and teachers, making it possible to quickly fill learning gaps and support students in moving steadily towards achievement. They reduce the turnaround time for student reports, save instruction time, and ultimately preserve resources.
- ✓ Extended constructed-response items allow educators to measure skills that are difficult to assess with traditional multiple-choice items. This could include writing an essay or answering an open-ended question in English language arts or mathematics.
- ✓ Writing items require students to write logical arguments based on substantive claims, sound reasoning, and relevant evidence. Emphasis on writing skills helps students improve their essay-writing skills and prepare for state and national writing assessments. In an RTI framework, the online essay-writing mentor and scoring tool supports a range of intervention approaches, from general screening to highly individualized intervention.





What other Common Core mandates require data literacy?

Informative Reports

These new item types require innovative new scoring techniques, as they're far more complex than multiple-choice test items. But they will also yield far more useful data and more robust reports. It will be up to the states, districts, school principals, and teachers to interpret the results to best take advantage of this new wealth of data.

Common Core assessment reports must provide teachers and administrators with useful, actionable interim/benchmark or summative information that links student performance to college- and career-readiness benchmarks.

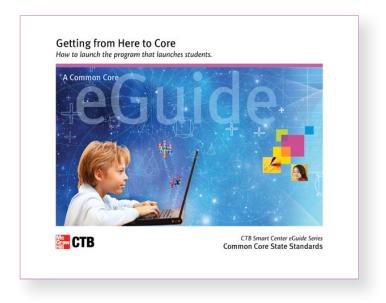
Challenging Cognitive Tasks

Common Core assessments must include items that provide measurable evidence of robust student thinking about important content.

Rigorous Validity and Reliability

These assessments must be engineered and developed to support appropriate and instructionally-meaningful uses of data. Further, the assessments must be research based and incorporate learning progressions and clearly-defined cognitive tasks. Online delivery in a computer adaptive environment is also a priority.

For more information about transitioning to the Common Core, download CTB's recent e-Guide or visit CTB.com/CommonCore







Bringing educators up to speed fast

Between the Race to the Top initiative, Common Core State Standards, and a general acknowledgement that every student deserves the best possible chance to succeed, it's essential to focus attention on the need for valid, reliable information for informing instruction. Traditional teacher training on assessment literacy or data literacy skills has been scant or absent. As a result, many teachers are unprepared for the massive changes to the educational landscape that will characterize the next 10 years. Some may be resistant to learning data literacy, believing that crunching numbers is for accountants and analysts, and not the job of a classroom teacher.

Even recent hires who have received some psychometric (the science that measures knowledge, abilities, and academic performance) training will need their skills refreshed as new requirements for reporting and test interpretation become the norm. New tests generating rich data and actionable insights need to have their results understood, communicated to stakeholders, then translated into classroom instruction and individual remediation.

All this will require professional development for classroom teachers, principals, and key administrative staff—a well-spent investment in time that will yield long-lasting benefits. Professional test developers like CTB can set up a program that can improve the assessment and data literacy of your key players and turn them into trainers for the rest of your staff.







A blueprint for data-based performance improvement

A thorough, effective assessment plan can be broken down into five essential steps, each requiring thoughtful planning and collaboration among stakeholders.

1. SET GOALS

2. CHOOSE MEASURES

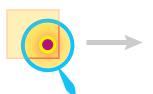
3. ANALYZE RESULTS

4. IDENTIFY OPPORTUNITIES

5. RE-MEASURE FOR PROGRESS















1. Set your goals for data-based performance improvement

What do you want to achieve?

A reduction in the drop-out rate?

More students headed to college?

Improved teacher effectiveness?

A 20% improvement in math scores?

System-wide Common Core readiness?

Make a list and prioritize your goals before embarking on any data collection program.

During this phase, you'll also want to make a realistic inventory of your current assessment readiness and create a plan to upgrade your system to accommodate some of the recent advances in administration, scoring, and reporting.

For example, are you still using paper and pencil tests or are you technologically prepared for online testing? Are there laptops for each student? Clickers? Tablets? Or are students rotating through a computer lab for standardized testing?

What do you need to do to improve your current state of assessment readiness? Make a list of what it would take to create a system that supports informed instruction and makes data advocates out all of your teachers and staff.

Laptops
Tablets
Clickers
Expanded computer lab
Teacher training
Hire Assessment Consultant
Committee to study data collection issues
Better assessments: do vendor review
Other







2. Choose the right assessments for the goals you've set

Different assessment types yield different types of data and reports. A comprehensive assessment program may include several different measures to track progress against your set goals:

✓ Interim/Benchmark assessments, given at the beginning of a year or new subject introduction, and again at the end, measure patterns in student performance and predict end-of-year performance on summative tests. These give school and district level administrators quick access to student data that can be aggregated or disaggregated at any level.

Example: CTB's Acuity

✓ **Summative assessments**, given at the end of a school year, measure student achievement and compare individual student growth to various subgroups and demographics. Norm-referenced tests measure basic concepts and skills commonly taught in schools throughout the country. Reports let teachers compare their students with a matched group of other students who took the test during the norming process, and then, compare each student's obtained scores to their anticipated achievement based on a national percentile scale.

Example: CTB's TerraNova, Third Edition

✓ Formative assessments are as much for the teacher as for the students. Given during instruction or immediately following the presentation of a key concept, they give educators critical information about student and classroom progress and uncover opportunities for further review, suggest adjustments to the teacher's approach, and make sure a subject has been mastered before moving on.

Example: CTB's Acuity and Yearly ProgressPro™

√ Through-course assessments provide continuous feedback to teachers throughout the year.





✓ **Performance assessments** are innovative new assessments based on tasks students are asked to perform to demonstrate their understanding and procedural skills relative to the content. More constructed-response than multiple-choice items, they are scored using rubrics relevant to the skills being assessed.

Example: CTB's Acuity and *TerraNova* Common Core

Additionally, there are two assessment types that support specialized learning situations:

- ✓ Writing assessments measure a variety of writing skills, including punctuation and grammar, persuasive argument, and higher-order thinking, and they give students unlimited practice of writing skills. Example: CTB's Writing Roadmap™
- ✓ Language proficiency assessments are given to English Language Learners (ELLs) students and English-native students who are learning another language.

Example: CTB's LAS Links®

Tests with multiple measures of achievement show a more complete picture of student knowledge.

The new generation of assessments requires new thinking, new technologies, new approaches to content, and new ways to analyze data and report assessment results. One big change is the shift from basic factual knowledge to an increased emphasis on performance.

In addition to multiple-choice (selected-response) items, you'll be seeing performance task items, technology-enhanced items, constructed-response items, and essay-writing items.

New ways to conduct scoring and reporting allow these new item types to be graded using rubrics that accurately score the cognitive skills being assessed, and even apply partial credits.

Make sure that the assessments you are using are taking full advantage of the latest assessment theory and scoring technologies to get the most accurate possible picture of your students' performance.

For information on Item Response Theory for both multiple-choice and constructed-response items, download the whitepaper,

Accuracy of Test Scores: Why IRT Models Matter.

(http://www.ctb.com/ctb.com/control/researchArticle/MainAction?art icleld=18746&p=ctbResearch)







3) Analyze your results to determine where each student stands in relation to your goal

Different reports crunch the data differently. The best reports are designed to meet the needs of the user. They are formatted to make usable insights readily available, be deliverable online, easy to customize, useful for identifying classroom, school, and district trends, and detailed enough to point out the learning strengths and weaknesses of individual students.

It's essential to choose assessments that yield accurate, actionable data, and it's just as important to become fluent in interpreting the results. Inadequate tests, poorly-constructed reports, or misinterpretation of results can result in a misdiagnosis of achievement at every level.

Districts need accurate reporting to help with data mining—measuring learning progressions year over year, identifying factors that contribute to poor or outstanding achievement, and reporting to parents or school boards. Teachers, students, and parents need accurate data to properly assess a student's level of achievement, to pinpoint learning gaps, and provide help before the students falls too far behind.

Acuity Algebra Readiness Exam - Class Assessment Report

GOODALL MR.			SAMPLE REPORT
Test Delivery:	Paper; Untimed	Teacher:	MR. GOODALL
Report Date:	04/25/07	School:	Branson Middle
Number Of Students Assigned Test	: 14	District:	Acuity Algebra Demo
Number Of Students Completed Test:	14 out of 14	Scoring:	Number Correct
Test Name:	Acuity Algebra Diagnostic Readiness Exam		
Test Window:	03/05/07 - 05/09/07		
Important Note:			

Summary Test Performance
Average Scaled Score : \$\\$\\$673(130)\$

▼673(130)

510 930
(Min Score) (Max Score)

		=Low M		Ioderate Master			ate Mastery Rang
strand	National Average	Local Average Class	Difference	Moderate Mastery Range	Performance by Strand		id
Assessment Totals:	46	52	6	4968			
Acuity Algebra					0 25	50	75 10
1 Number Sense, Computation, and Estimation	51	56	5	4567			
10 Statistics and Probability	51	44	-7	6180			
2 Variables and Expressions	52	48	-4	4567			
3 Inequalities and Linear Equations	49	58	9	5167			
4 Graphing Inequalities and Linear Equations	22	51	29	4160			
9 Geometry	43	55	12	5167			

Your students' average performance and the average national performance are displayed. In addition, the Low, Moderate and High Mastery range indicates the Material Restricts of the Moderate Mastery range indicates like Mastery for that particular strand. For example, in the strand I Number Sense, Computation, and Estimation, your students average performance is 50, compared with the average national performance of 51. This difference is 5. Strands on which students' average performance existed, the considered potential carse of strength. Conversely, strands on which students' average performance exceeds the energy antional performance can be considered potential carses of strength. Conversely, strands on which students' average performance is lower than the average national performance.

Note: See your Acuity Algebra reference materials for testing guidelines that provide the most reliable and accurate report data.

Page 1 of 1



A closer look at reports

Summative Assessment Reports

Here are just a few of the most commonly used reports for norm-referenced summative assessments, such as CTB's TerraNovaTM.

Home Reports identify a student's potential strengths and needs in major content areas.

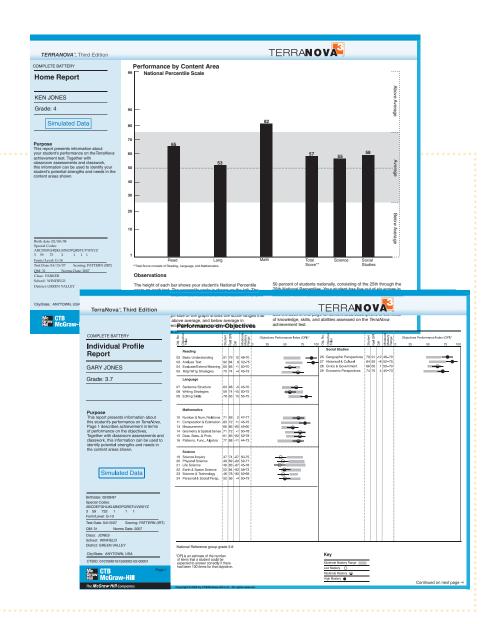
Individual Profile Reports describe achievement in terms of performance on objectives, present descriptions of the student's demonstrated skills and abilities, and outline opportunities for additional practice and/or intervention.

Group Reports summarize achievement data for a specified group.

Summary Reports help educators evaluate, plan, and establish educational priorities.

Longitudinal Summary Reports show a group's progress over multiple test administrations.

State Proficiency Reports give parents and teachers a comparison of student progress against state standards.



Formative and Interim/Benchmark Reports

Formative and Interim/Benchmark assessments measure proficiency, track progress, target instruction, and predict performance, and will ideally allow teachers to drill down to see individual student performance scores, item-by-item scores, and present instructional resources for practice and remediation. Here is a partial list of reports generated by Acuity, CTB's award-winning formative and interim/benchmark assessment:

Individual Student Reports show a student's strengths and areas for additional instruction. Drill downs allow cross-referencing and data mining.

Class Reports help teachers understand class strengths and areas of difficulty, and help suggest small-group and whole-class instruction. They also allow drill downs to give a complete picture of performance, right down to item analysis and comparisons of aggregate class data with individual student performance results.

School Reports help focus attention at the school level and suggest areas where additional support is needed for teachers, principals, coaches, intervention specialists, and professional development staff.

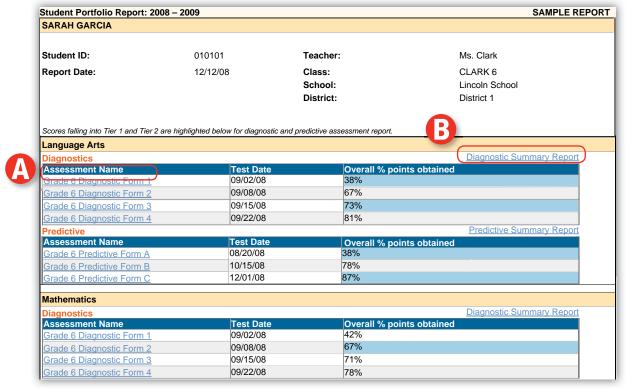
District Reports supply a comprehensive view of student achievement, yet still allow drill-downs to individual student performance or specific test questions. They help support meaningful decision making that translates to improved achievement across students, classrooms, and schools.





Student Portfolio Report

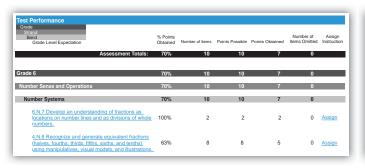
Delivers quick access to a library of reports maintained for each student





Specific Test Drill Down

Allows quick access to Diagnostic Form 1.





Summary Report Drill Down

Opens a Longitudinal Report showing performance results across multiple test administrations.

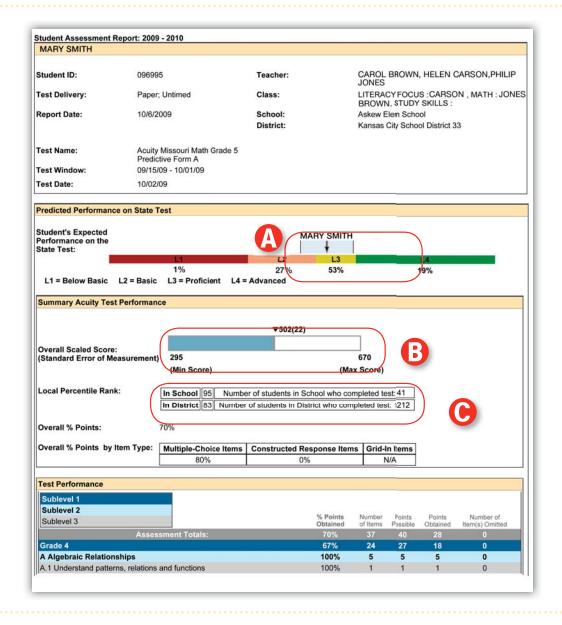
Tests:	Acuity NY Math Grade 6 Diagnostic Form 1		Acuity NY Math Grade 6 Diagnostic Form 2		NY Math Grade 6 Diag Form 3	
Strand	% Points Obtained	# of Items	% Points Obtained	# of Items	% Points Obtained	# of Items
Overall Assessment:	73%	37	54%	60	45%	40
Number Sense and Operations	72%	18	63%	25	61%	18
Statistics and Probability	80%	4	44%	8	20%	5
Geometry	80%	4	30%	8	0%	5
Measurement	57%	6	55%	10	57%	7
Algebra	80%	5	55%	9	40%	5



Student Predictive Assessment Report

Provides powerful interpretations of student performance to support progress and achievement

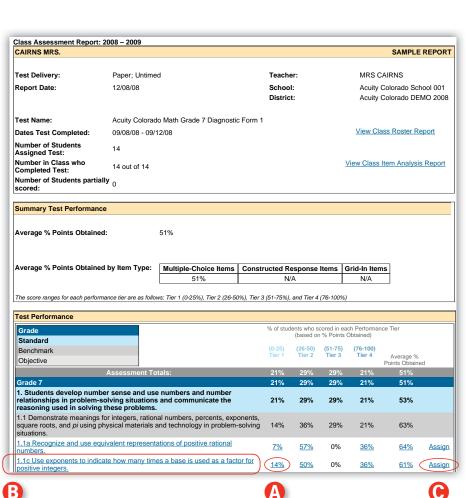
- Get clear forecasts with easy-to-read tables that indicate each student's expected performance level on subsequent state test
- Clearly understand student progress with scale scores that allow for longitudinal tracking of student progress within and across school years
- Review student achievement using Local Percentile Rank, the percent of scores in the school or district that were below a given student's score.





Class Assessment Report

Provides teachers with actionable data on how their classes are doing relative to your standards



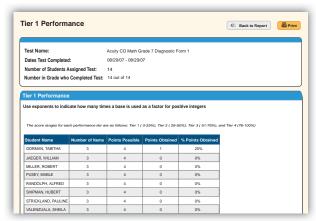


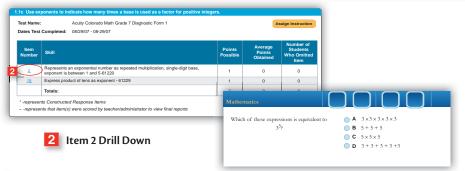
Tier 1 Drill Down

Notice that 14% of this class is in the Tier 1 performance level (0–25%) for items on the related standard. Educators can easily drill down into tiers to discover which students are in need of remediation and which are proficient.

Item-Level Drill Down

Educators can quickly link from reports to the skill level and the item for a better understanding of which items are most difficult for students.

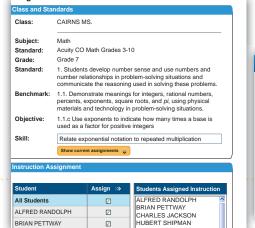




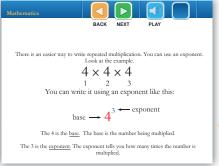


Assign Instruction to Class Drill Down

Engaging instructional resources are easily accessible from reports to assign to students in need of additional coaching and practice.



Student View of Instructional Resource

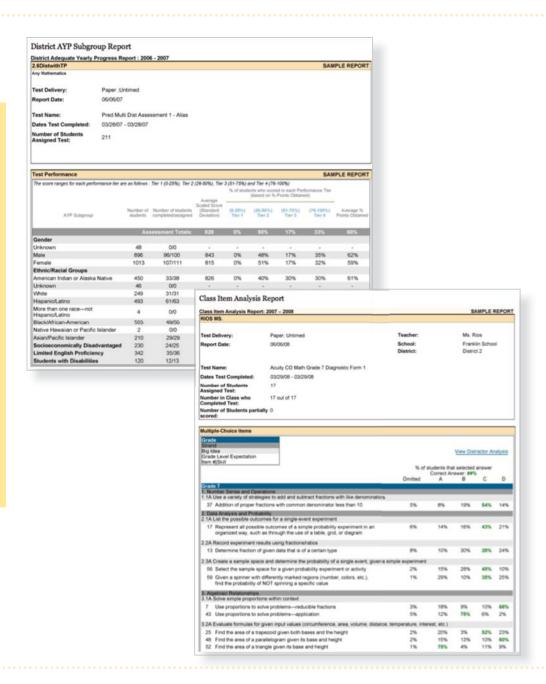




How one district used data to improve scores on their state test

The Park Hill School District in Kansas City, Missouri began using Acuity—an award-winning formative assessment—in 2006 to measure student progress against the GLE and predict student performance on the MAP, Missouri's state test.

Student-teacher collaboration was an integral part of the plan. Students were asked to keep a student data folder and participate in periodic consultations with teachers to compare their interim test scores with district and state data. In the process, they learned how to read and interpret their results and see where they needed improvement, then collaborate with teachers to set goals and work towards improvement.









4. Identify opportunities for improvement and modify instruction where needed

The best assessments don't simply report numbers. They offer detailed feedback on how students are doing relative to standards and pinpoint areas where individuals and groups of students will benefit from additional instruction or practice.

The items presented in a high-quality assessment can be directly linked to specific skill sets and standards. For example, the software behind an online assessment can not only identify specific strengths and deficits related to those standards, but it can also present students and teachers with instructional resources specifically tailored to help individual students work on the areas where they appear to need support.

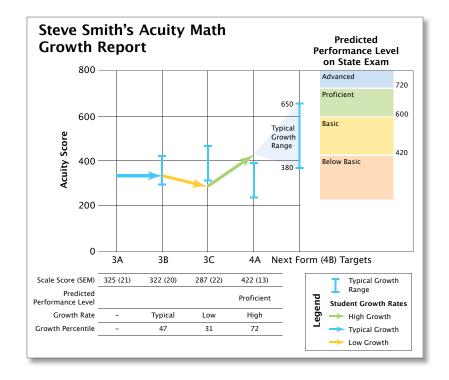
A student may have demonstrated a solid understanding in five of the six areas covered in that assessment. But if the student performed poorly in one area, an assessment instrument can link the student to worksheets or content focusing on the area that needs work. This technology can shorten the time from learning to mastery and improve future test results.

Studies have shown that students who fall too far behind will tune out and/or drop out. The feedback loop available with today's technology can prevent this from happening and can be used to reduce dropout rates, improve performance, and help build a stronger teaching staff.

With this feedback, teachers can react immediately to needed changes in instruction, design additional class work and homework, assign targeted work to groups, and retest as needed.







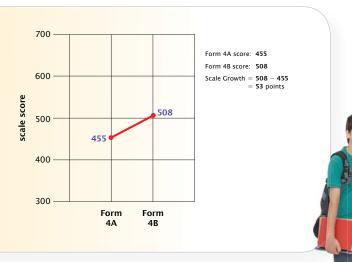


- Your student's scale score is clearly represented on the Y axis.
- Acuity clearly illustrates each student's progress relative to three important growth categories: Low, Typical, or High.
- Reports give additional context to student performance with clear percentiles for each growth category. Percentiles are presented in tabular format on each report for an easy, at-a-glance reference.

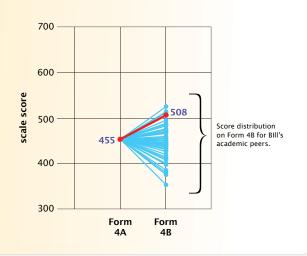


Growth Reports

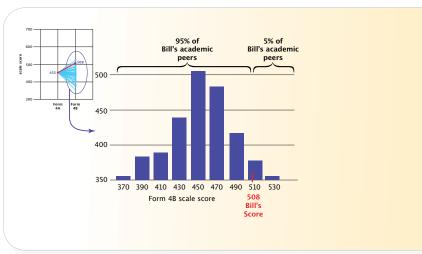
Bill's scaled score growth in Acuity Predictive Math



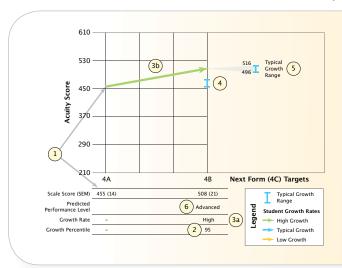
Bill's scaled score growth compared to his academic peers



Bill's Form 4B score in the distribution of his academic peers



Bill's Math Growth Report









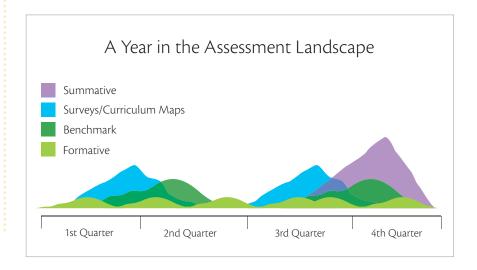
5. Re-measure to mark progress

To make sure student progress is happening according to your plan, retest periodically and compare results. If you're setting goals, choosing the right measures, using them to collect data and analyzing the data properly to target classroom instruction, you will begin to see results within a few months.

There will always be some sort of data collection going on in your class-rooms, whether formal or informal. Here is a plan that works for many educators:

- •Once/year administer summative district and state assessments. Reports will measure aggregated, disaggregated, strand, item, and student performance.
- 2–4 times/year collect data about people to reveal information on demographics, enrollment numbers, teaching practices, etc. These may take the form of surveys, interviews, observations, and/or curriculum maps.

- Quarterly or after each unit, administer benchmark assessments to confirm mastery of the subject matter
- 1–4 times/month conduct formative assessments to check on progress in math, writing, and other student work.
- Daily or weekly conduct formative assessments. This may take the form of student self-assessments, pop quizzes, student journal entries, and other assignments.





How I learned to love data

By Philip Adams

As a science teacher in the Midwest, I was one of the first people in my district to worry about how good or bad the items on standardized science tests were. I realized that I was going to have to learn about psychometrics and test construction myself if I wanted to translate test results into something that would be useful in the classroom. So I went back to school and got a doctorate in Educational Leadership and Policy Studies and, after graduation, became a senior examiner for the state of Missouri Quality Awards.

There's a huge variance in the use of data across the country. And it all starts with having good measures—because imperfect measures lead to imperfect conclusions, and that's just a waste of everyone's time. Today, as an Assessment Solutions Consultant at CTB/McGraw-Hill, I recommend that school districts that can afford it hire an assessment consultant or a full time staff member to help manage the program and train the teachers and principals.

For those who can't afford to hire a specialist, I recommend getting in touch with someone like CTB who offers professional development programs to teach data literacy and train people to go back into the schools and pass on this information to every teacher in every district.

The example I always use is: before you can lose those 15 lbs, you've got to step on the scale to see what you weigh today. Then climb back on to check your progress and make sure you're not slipping up. It's the same with data analysis. Measure at the beginning, during the process, and again at the end. And make sure those scales are accurate.

Test scores improved significantly, as did teacher effectiveness and parental involvement.





The importance of immediate feedback

The power of assessment to improve instruction and learning is greatest when the feedback is immediate. Fortunately, the old days of waiting months for tests to be handscored, analyzed, printed, and returned to the teachers are over.

New test administration, scoring, and reporting technology has enabled great strides to be made in the delivery of secure, accurate, and comprehensive data feedback. In many cases, tests administered online can be scored and reported with a click. The speed of results delivery—even for paper-and-pencil tests—has been greatly reduced, from what used to be months to just a few days. And secure online portals for warehousing test results and providing resources to districts, schools, teachers, and even parents, are in common usage.

Online dashboards and portals streamline the entire process. CTB's Online Reporting System, for example, lets you quickly and easily access data on demand, lets you slice and dice (aggregate and disaggregate) your data into customized reports, and export data and charts to PDF or Excel documents

To see the CTB Online Reporting System in action, visit:

CTBOnlineReporting.com

District login: demodistrict

Diocesan login: demodiocese

School login: demoschool

Teacher login: demoteacher

Password: ctbdemo





How fast is "fast"?

As you'd imagine, any test that's administered using paper and pencil needs to be mailed in, handscored, printed out and mailed back—a labor-intensive process that can take days or weeks. Computer adaptive testing, on the other hand, can yield instantaneous results. For example, Yearly ProgressPro, CTB's research-based assessment, instructional, and intervention tool, generates results on the spot. CTB's Acuity returns compiled results within two hours.

Funding is often contingent on accurate, accessible report generation, and fast online reporting is an invaluable tool to help you qualify for needed funds.

Parents can receive reports via email and quickly learn their student's strengths and opportunities for growth. This can facilitate involvement in a child's academic progress and help support his or her growth.

Teachers can generate reports for each class, then drill down to individual student performance data in great detail.

School Administrators can aggregate and disaggregate data by classroom or building as a whole for program evaluation, accountability, and professional development requirements.

K-12 Testing Directors can create easy-to-understand summary reports for all key stakeholders. They can export data and graphs to make presentations to school boards, teachers, and parents.

Superintendents and District Administrators can quickly generate report views for specific program needs and export them to make reports to school boards, staff, and the community.

Federal Program Administrators can access powerful information to meet accountability requirements for federal programs and grants.





Assessment analysis guidelines

While test reports and individual scores can tell an accurate and compelling story about student performance, there are some caveats to their use. Keeping these in mind will help ensure that test information is used in ways that best serve the interests of the student, the school, and the community.

When interpreting and using the results of any group-administered or standardized achievement test. remember to:

✓ Consider "standard error of measurement"

All tests—even those developed to the highest standards—have a certain measurement error range, because they are a sample of student performance at a single point in time. A student may be hungry, tired, distracted, or ill on test day and not perform at his or her best level. Before making critical decisions that affect the life of a student, make sure to consider all available information, including previous test scores, class grades, and teacher judgment.

✓ Don't hesitate to question results

If there is a good reason to question the accuracy of a test result, that test result should not be made part of a student's permanent record. Any result that is at great variance with other information about the student should be interpreted with caution.

✓ Make sure you're testing at the right level

If a test is too difficult, a student will usually resort to guessing, a practice that can yield unreliable results. Retesting with a lower level of the test would probably provide more useful information about the student's specific needs.

Also, it's best to interpret perfect scores with caution. When all items are answered correctly, retesting with a higher level of the test would provide better information about the student's true top performance level. Locator tests are available to help you determine the proper level to administer to students in question.

✓ Administer as directed

Norm-referenced scores are valid only when the test is administered following the same directions and time limits under which the test was standardized. For *TerraNova,Third Edition*, this includes administering the test with appropriate accommodations. When students are unable to take the test under standard conditions, administering the test with modified directions and time limits may yield valuable information for diagnostic purposes, but it may impact the norm-referenced scores. (See "Guidelines for Inclusive Test Administrations" on CTB's website at CTB.com)

Link to: http://www.ctb.com/ctb.com/control/researchArticle MainAction?topicId=390&articleId=475&p=ctbResearch

✓ Consider the size of the test group

Mean (average) and median scores based on small numbers of students—generally fewer than 25—should be interpreted with caution. One or two extreme scores can greatly affect the mean and median in such instances.

✓ Test using the latest norms

Note that the norms for a standardized achievement test describe the levels of achievement that prevailed at the time the test was standardized. Check with the publisher to make sure the test you are using is based on recent norms





Choose assessments from a publisher you trust

As we've described, the critical role of assessments and the data generated by them places a heavy responsibility on the test publisher, who must ensure that the results are accurate and that clear, reliable information about the interpretation of results is provided.

- Make sure you are using assessments from a publisher with an established reputation for excellence and innovation in educational assessment and measurement theory.
- Check to make sure their assessments are proven to be valid, fair, bias-free, and reproducible from classroom to classroom. Ask about their psychometric standards and whether they use the most effective methodologies for assuring reliable results.
- Preview their scoring practices and take a close look at their reports to be certain they are giving you clear, understandable data that will help you improve instruction at your schools. Do they provide longitudinal analyses, growth reports, classroom reports, and teaching evaluations on all subjects?
- Ask for a demo of their online reporting portal to assure your team that they will have 24/7 access to results, analysis, and custom test-generation tools.

- Do they provide a dashboard for creating instant reports from a variety of data sources? Can you easily aggregate and disaggregate the results generated so you can create custom reports for specialized applications, analyze the performance of demographically-distinct groups, and prepare special reports for community meetings and other special occasions?
- Find out if the publisher will customize assessments for you, which may include developing content and test items based on non-standard subjects such as Bible studies or advanced topics not normally covered in end-of-year summative tests.
- •Finally, ask about professional development. How willing are they to train your staff in using the assessments they provide? Do they offer a variety of solutions, including in-person trainings, interactive online guidance, CD-based instruction, etc.?





What about money, time, and technology considerations?

Yes, it costs money to upgrade your assessments to provide better data for better decisions. And it takes time and money to retrain teachers to do a better job of integrating data analysis into their daily routines. But, continuing to operate without these invaluable tools will yield the same results you're getting now. You can continue to suffer from lower scores, more dropouts, and lower funding, or begin the process that will turn around your schools, improve the morale and self-esteem of teachers and students, and become a high-achieving system.

Don't sacrifice substantial results for immediate cost savings

When it comes to assessments, you get what you pay for. Look closely at the assessment options you may be considering. Too often the low-cost ones are legacy products built according to outdated standards, and may lack the rigor, reliability, and/or validity required by today's academic initiatives.

Assessment qualities to check for:

- Correlation with recent norms and other independent measures of achievement and cognitive ability.
- Engaging content your students can relate to. (Studies show that engagement can lead to better test results.)

- Multiple measures that show depth of knowledge, not simply rote learning (constructed-response and performance items in addition to multiple-choice items).
- Strict controls in test development that result in valid measurements that support meaningful decisionmaking.
- Sophisticated scoring that facilitates rapid feedback to teachers, administrators, students, and parents.
- Reports that are useful in informing instruction and can be used to improve individual student performance.

Spend more time teaching smart, less time wasting time

A twenty-minute assessment can drive instruction for the next month. Once you know where your students are, you can skip unnecessary reviews and begin to focus instruction on filling gaps and moving forward. When you're teaching smart, more students stay engaged in the learning process. The slower learners receive the help they need to feel proud and successful, the faster learners remain engaged and challenged, and all students experience the excitement of learning at a pace that brings out their highest levels of productivity.



Ask your assessment publisher to help with professional development

A strong professional development program educates teachers on learning tools and methodologies, and enables teachers to:

- · Obtain an overview of what should be taught and how
- Analyze report data at the student, school, and district levels
- Transform data analyses into enhanced curriculum
- Target instruction at the individual and group levels
- Adjust instructional styles to meet student needs
- Acquire knowledge in areas of expertise
- Acquire new knowledge to expand skills base

Your assessment provider may be able to provide access to community resources that teachers and administrators can use to learn how other teachers and administrators have responded to situations—either in the classroom, when engaging with parents, or when trying to use the tools provided for working with the data. Publishers like CTB-McGraw/Hill can create a program as part of their service package that will help your entire staff become more data literate.

General support and product-specific support are available by phone, online, and through a variety of professional development programs that include in-person trainings, DVD/video trainings, and guided instruction.

In addition, CTB has an entire online community for teachers and administrators using Acuity. Teachers can communicate with other teachers, administrators with other administrators. They can swap ideas and experiences, even documents and materials they have created for use in their schools and districts that other schools and districts may find useful.

Training models at CTB include:

- · Train the trainer
- Train the individual
- Coaching leaders
- Job shadowing

- Product training
- Product positioning and launch strategy

Workshops:

- Interpretation and use of test results
- Targeted instruction
- Content development
- Item review

- Item banking
- Reporting
- Handscoring
- Test securityItem writing







The role of technology in test administration, data collection, and reporting

Technology will play an ever increasing role in data management as online assessments, online reporting, data warehouses, administrative portals, and parent networks gain popularity. The future of assessment and reporting lies in digital delivery, particularly as testing agencies begin to develop assessments that delve more deeply into student performance, such as problem solving, visual literacy, real world literacy, and measures of critical thinking. States are increasingly requesting online solutions to bring the benefits of technology into classrooms: to provide prompt feedback, reduce the turnaround time for student reports, save valuable instruction time, and ultimately save costs related to assessment and reporting programs. Most of the technology already exists to make a truly efficient, universal and yet highly customized digital-based system a reality.

How technology motivates students

Students take pride in probing the Internet for the latest research, discoveries, and up-to-the-minute news. While some love to read books, many prefer to get their information online. The use of technology at school replicates students' environments outside the classroom and serves as a familiar, energizing factor in the classroom.

Education technology, with its potential to engage and entertain, its instant feedback and ability to adapt to each user's needs in the moment inspires students to take control of their own learning and keeps them engaged in classroom activities.

Make the highest use of the technology you have

Your district has likely applied for grants and earmarked future budget spending towards technology, but until those funds arrive, there are steps you can take to make the most of what you've got. Even if your schools are lacking the technology assets of larger, betterfunded systems, as long as you have a computer lab in a school, you can make progress towards your data-driven performance improvement goals. Simply schedule computer time wisely—for student assessment and online practice, and for helping improve the data literacy of your staff. Also, keep in mind that the right use of technology can ultimately lower costs of assessment construction, administration, scoring, and reporting, while providing valuable resources for intervention, instruction, professional development for teachers, and parent engagement in student learning.





Engaging families in understanding test scores

A key component of student success is ongoing parental involvement in learning. An ideal assessment program not only encourages this involvement, but also provides the tools, resources, and access that enable parents and families to support student achievement at the times that are most convenient to them—especially important for working-parent households.

Key parent engagement tools include:

- Personal learning plans based on an innovative analysis of students' specific needs and strengths
- Activities that help prepare for state/district curriculum and state standard mastery
- Family-friendly home activities, planning tools, resources, and advice for each age group
- Multilingual formats to give both families and schools exceptional clarity around achievement
- Accessibility from any computer or Internet-access device
- Community information and resources for parent and student support to strengthen learning and the home/school connection

How administrators can give every student the best chance for success.

- Be the leader your students need to succeed.
- Become an advocate for the principles and practice of data-informed instruction.
- Help teachers transition from paper and pencil tests to online testing and scoring, and from multiple-choice items to more rigorous performance-based items.
- Allocate funds for professional development, technology, and high-quality assessments.



CTB-McGraw/Hill offers a full complement of affordable, top-quality assessment solutions in a variety of testing categories.















Acuity

Your interim and formative testing solution

The Acuity InFormative Assessment[™] solution helps you diagnose, predict, report, communicate, and provide individual instruction—all in one powerful and award-winning solution. Acuity now offers a complete and integrated system beginning in kindergarten and helping you advance student achievement all the way through high school.

You can use Acuity throughout the year to build achievement with individual students, classrooms, schools, and districts. Now Acuity delivers innovative resources for Grades K–12 that support your gradual and effective transition to teaching and learning relative to the Common Core State Standards (CCSS) in its latest product, Acuity for Common Core.

Learn More: CTB.com/Acuity









Writing Roadmap Practice for writing excellence

Written response items are now mandated by Common Core guidelines, but in many schools, writing instruction has been slashed from the curricula. Writing Roadmap helps teachers fill this knowledge gap efficiently with a set of essay-writing exercises designed to challenge and inspire students to learn to write better essays.

Writing Roadmap is a state-of-the-art online assessment solution that helps your students improve their essay-writing skills, while allowing teachers to spend less time grading and more time teaching. It provides students and teachers with instant feedback on writing assignments, including automatic scoring, instructional guidance, and robust reporting.

Learn More: CTB.com/Writing Roadmap







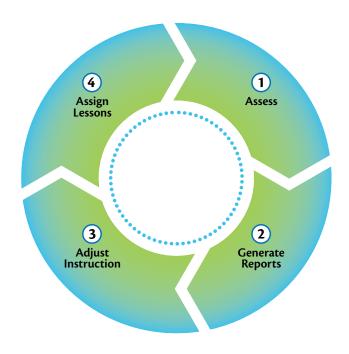


Yearly ProgressPro Monitor student learning in real time

Yearly ProgressPro is an award-winning online progress monitoring solution that is based on 25 years of research in Curriculum-Based Measurement (CBM), a methodology proven to improve student learning outcomes. CBM is particularly effective with Response to Intervention (RtI) programs, a method of academic intervention that is designed to provide early, effective support to struggling learners.

With Yearly ProgressPro's robust classroom assistance tools linked to an easy-to-understand data management system, you can receive detailed diagnostic curriculum-based measurements of a studen'ts progress towards a defined set of skills. Teachers can measure the effectiveness of instruction and track both mastery and retention of grade-level skills, with either CBM or their own customized assessments.

Learn More: CTB.com/YearlyProgressPro















TerraNova, Third Edition and TerraNova Common Core

Reading, English language arts, and mathematics for Grades 3–8

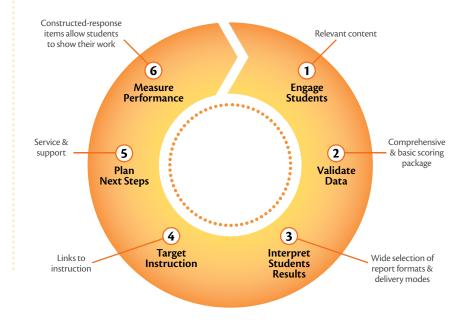
The assessment includes constructed-response, extended constructed-response, and performance task items in the same test, on the same scale. No other national achievement test developed to date offers this combination, all in a single test book. Results are available within seven days or less, and reports are designed to show administrators, students, and teachers where they stand on both national and the state standards today and over time.

Norm-referenced summative assessments aligned to the Common Core

The new *TerraNova* Common Core from CTB/McGraw-Hill is the right product at the right time. With content aligned to the Common Core, this new form of *TerraNova* features innovative items that deliver an authentic measure of the higher-order thinking skills and increased depth of knowledge highlighted by the Common Core.

TerraNova Common Core is the only field-tested, valid, and authentic measure of the Common Core currently available to districts today. Use it to familiarize your teachers and students with the new, more challenging test item types, to conduct early benchmarks, and to get an early start on your transition plan.

Learn More: CTB.com/TerraNova











I AS Links

Core-ready assessments for your English language learners

As classroom populations grow more culturally and ethnically diverse, so do your educational and assessment needs. The all-new LAS Links is the world's most advanced suite of English and Spanish English Language Proficiency tools available—an integrated system of assessment, instructional placement, instructional guidance, and support for all English language learners. LAS Links helps educators:

- Quickly and accurately place incoming students with the LAS Links Placement Test
- Measure Spanish language proficiency with LAS Links Español
- Monitor student progress with LAS Links Benchmark Assessments
- Build teacher skills using LAS Links Staff Development DVDs and Training Audio CDs
- Augment curriculum with LAS Links Instructional Guidance that is grade and proficiency-level appropriate

The latest edition of LAS Links aligns to Common Core State Standards, so while students are improving their language proficiency, they are also getting practice in answering the unique types of test questions required by the Common Core, including:

- Performance tasks that uncover deeper levels of student understanding
- Technology-enhanced items and assessments that provide instant feedback to support students in moving steadily towards achievement
- Extended constructed-response items to measure skills that are difficult to assess with traditional multiple-choice items
- Writing items that improve students' essay writing skills

Learn More: CTB.com/LASLinks









Assessment development rigor at CTB

Traditionally, testing was done by teachers, who simply wrote items similar to the textbook exercises and then counted the number of points correct. More recently in the last 30 years, educational testing has become a scientific discipline that is studied in doctoral programs in universities.

CTB frequently receives requests from educational authorities world-wide to provide solutions for a wide variety of testing needs. School, ministry, state, and provincial requests range from consultation on a single test form to the development of complex, high-stakes testing programs requiring rigorous security measures. We begin from the point of view that each client has unique testing needs and end with the implementation of an innovative testing program designed to address each of these needs.

We typically develop tests in three phases, with test delivery and support considerations embedded in each:

- 1. Needs analysis
- 2. Test program design and specification
- 3. Test development and validation

During needs analysis, the test purpose and use is defined and documented in detail. These definitions serve as the basis for the development of your testing program requirements. CTB works closely with clients to ensure that all testing requirements are identified and well documented at this stage.

During phase 2 (the design phase), CTB uses all information gathered during the needs analysis to design a testing program and develop detailed specifications that address all of your requirements. These design and specification documents can serve to support Requests for Proposals as well as provide a detailed map for test development and implementation. It will address each of the testing components in enough detail to build and validate a successful testing program that satisfies all stakeholder's needs.

Phase 3 is the test development and validation stage. In this phase we will execute the design produced in Phase 2 to build the test and testing infrastructure, as well as conduct all research analyses required to produce precise and meaningful examinee test scores. Supporting activities that CTB can provide during this phase are 1) technology services, 2) scoring and reporting, 3) program management, 4) capacity planning, 5) policy development support, and 6) professional development activities.

The process for creating a test item at CTB

To create each item on a CTB assessment, meticulous planning is required. The process starts with construction by assessment and content experts, followed by an examination for bias and fairness, and an evaluation to ensure that reliable inferences can be drawn from the item about the student's strengths and instructional needs. Numerous item reviews, specific to standards, ensure accurate skill evaluation and measurement of student growth.





About CTB

Established in 1926 to "help the teacher help the child," today CTB/McGraw-Hill ranks among education's leading assessment partners. With a reputation for delivering the highest quality psychometric research, accuracy, and client service, CTB/McGraw-Hill serves the global education community with a comprehensive range of assessment programs and services. Our clients include public, private, and religious educational institutions in all 50 United States and in 49 countries around the world. CTB provides powerful interim and formative, summative, norm-referenced, performance-based, and observational assessments delivered in both online and traditional formats.

Our scope of services and capabilities includes rigorous standards alignments for all of our assessments. Also, we have developed authentically aligned assessments to support the new Common Core State Standards, enabling states and districts to migrate to these new standards effectively. At the same time, we continue to ensure that all of our assessment solutions maintain alignments to existing accountability requirements under federal and state mandates.

All CTB programs and services are distinguished by innovative delivery modalities encompassing online and other technology-based approaches, thorough and timely reporting, the profession's most accurate and fastest scoring processes, and comprehensive data management services. We collaborate closely with educators in the development and testing of our programs and take pride in the depth and quality of our professional development programs that support our clients, from the classroom to the boardroom.







Summary

We are living in a data-based world, and the field of education is no exception. As the nation searches for ways to improve student achievement, educators and policy makers continue to evaluate and reform their education systems using data-based measures to create benchmarks and record progress. Data literacy is now seen as a necessity at every level, from top administrators to classroom teachers. With it comes a greater responsiveness to individual students' learning needs, improved performance, and readiness for a world of work and higher education.

In an increasingly fast-paced world, teachers are required to react quickly to data they collect and modify daily classroom instruction in response. Breakthroughs in technology are facilitating these changes, along with teacher training, better assessments, and programs such as Common Core State Standards and Race to the Top.

Educational assessments and related resources not only capture data on learning and achievement, but are fundamentally changing the way teachers teach, students learn, and parents engage in classrooms around the world. They provide essential information to guide genuine education reform and improvement, and they allow us to hold ourselves accountable for providing students with the quality education they need to compete and thrive in the world economy.

The data generated from carefully selected, quality assessments can be used to monitor educational systems for public accountability; help improve curricula; evaluate the effectiveness of teaching and instructional practices; measure student achievement; and determine a student's mastery of skills.

When test results are used to help make critical educational decisions that affect the lives of students, it's important to know that the tests are backed by the best research and consultant support possible. When selecting an assessment partner look for:

- Quality assessments that yield accurate, meaningful data
- Report data that makes it easy to apply these results to classroom instruction
- A wide range of assessment types from a single source
- A publisher with professional development capabilities
- The highest levels of rigor in developing test item development
- Shared resources and community for continued growth and improvement

High-quality tests selected for their appropriateness for meeting key educational goals, professional development for teachers, and targeted instruction based on individual student results are a proven combination for boosting achievement. Especially in these times of increasingly rigorous academic standards, data is key to the turnarounds so many of our schools are mandated to achieve.



Academic Aptitude Test

An aptitude test predicts achievement in academic pursuits. Ideally, in constructing this type of test, the developer tries to minimize the effect of exposure to specific materials or courses of study on the examinee's score.

Accommodation

An adjustment in the administration of an assessment to meet the needs of a student. Accommodations could be extended time or a test booklet with larger print.

Achievement Test

An assessment that measures a student's acquired knowledge and skills in one or more common content areas (for example, reading, mathematics, or language).

Adult Accountability Test

An assessment intended primarily for individuals 18 years old or older who are no longer attending elementary or secondary school.

Alternative Assessment

An assessment that differs from traditional achievement tests. For example, an alternative assessment may require a student to generate or produce responses or products rather than answer only selected-response items. This type of assessment may include constructed-response activities, essays, portfolios, interviews, teacher observations, work samples, and/or group projects.

Analytic Scoring

A scoring procedure in which a student's work is evaluated for selected traits or dimensions, with each dimension receiving a separate score.

Aptitude Test

A test consisting of items selected and standardized so that the test predicts a person's future performance on tasks not obviously similar to those in the test. Aptitude tests may or may not differ in content from achievement tests, but they do differ in purpose. Aptitude tests consist of items that predict future learning or performance; achievement tests consist of items that sample the adequacy of past learning.

Authentic Assessment

An assessment that measures a student's performance on tasks and situations that occur in real life. This type of assessment is closely aligned with, and models, what students do in the classroom.

Battery

A test battery is a set of several tests designed to be administered as a unit. Individual subject-area tests measure different areas of content and may be scored separately; scores from the subtests may also be combined into a single score.

Bias

A situation that occurs in testing when items systematically measure differently for different ethnic, gender, or age groups. Test developers reduce bias by analyzing item data separately for each group, then identifying and discarding items that appear to be biased.

Composite Score

A single score used to express the combination, by averaging or summation, of the scores on several different tests.

Comprehensive Equal-Interval Scale

A scale marked off in units of equal size that is applied to all groups taking a given test, regardless of group characteristics or time of year. Each test yields its own scale. On TABE, for example, scale scores are expressed in numbers ranging from 0 to 999. The continuity of the scale among levels comes from administering special test forms containing items from adjacent test levels to random groups of students. This allows the TABE scales to be calibrated so that a given adult learner is expected to obtain the same scale score regardless of the form or level of the test he or she takes. However, the standard error of measurement associated with that student's score will vary systematically from level to level.



Computer Adaptive Tests

Computer adaptive tests (CATs) are computer-administered tests that tailor the selection of test items during the administration of the assessment based on the responses of examinees. By adapting the difficulty of the test to the ability level of the examinee, CATs generally provided greater precision and/or enable shorter testing time, when compared to non-adaptive tests.

Construct

The concept or characteristic that a test is designed to measure.

Constructed-Response Item

An assessment unit with directions, a question, or a problem that elicits a written, pictorial, or graphic response from a student. Sometimes called an "open-ended" item.

Criterion

A standard or judgment used as a basis for quantitative and qualitative comparison; that variable to which a test is compared to constitute a measure of the test's validity. For example, grade-point average and attainment of curricular objectives are often used as criteria for judging the validity of an academic aptitude test.

Criterion-Referenced Test

A test in which every item is directly identified with an explicitly stated educational behavioral objective. The test is designed to determine which of these objectives have been mastered by the examinee.

Culture-Fair Test

A test devised to exclude specific cultural stimuli so that persons from a particular culture will not be penalized or rewarded on the basis of differential familiarity with the stimuli.

Curriculum-Referenced Test

An assessment that measures what a student knows or can do in relation to specific, commonly taught curriculum objectives.

Derived Score

A test score pertaining to a norm group (such as a percentile, stanine, or grade equivalent) that is an outgrowth of the scale scores. Derived scores are useful descriptors; however, they are not calibrated on an equal-interval scale, so they cannot be added, subtracted, or averaged across test levels the way scale scores can.

Diagnostic Test

A test intended to locate learning difficulties or patterns of error. Such tests yield measures of specific knowledge, skills, or abilities underlying achievement within a broad subject. Thus, they provide a basis for remedial instruction.

Discrimination Parameter

The property that indicates how accurately an item distinguishes between examinees of high ability and those of low ability on the trait being measured. An item that can be answered equally well by examinees of low and high ability does not discriminate well and does not give any information about relative levels of performance.

Distractor

An incorrect answer choice in a selected-response or matching test item.

Early Childhood Test

An assessment intended for students in kindergarten and grades 1 through 3.

Educational (Instructional) Objective

A statement that defines an intended outcome of instruction. It describes what a successful learner is able to do at the end of the lesson or course, defines the conditions under which the behavior is to occur, and often specifies the criterion or standard of acceptable performance.



Equal-Interval Scale

A scale marked off in units of equal size that is applied to all groups taking a given test, regardless of group characteristics or time of year. Each test yields its own scale. On TABE, for example, scale scores are expressed in numbers ranging from 0 to 999. The continuity of the scale among levels comes from administering special test forms containing items from adjacent test levels to random groups of students. This allows the TABE scales to be calibrated so that a given adult learner is expected to obtain the same scale score regardless of the form or level of the test he or she takes. However, the standard error of measurement associated with that student's score will vary systematically from level to level.

Equated Score

A score from one test that is equivalent to a score from another test. Equated scores are usually obtained by administering the two tests of interest to a representative sample of students. Scores from one test are then aligned with scores on the other test using equating analysis.

Face Validity

An evaluation of a test based on inspection only.

Floor

The opposite of ceiling, it is the lowest limit of performance that can be measured effectively by a test. Individuals are said to have reached the floor of a test when they perform at the bottom of the range in which the test can make reliable discriminations. If an individual or group scores at the floor of a test, the next lower level of the test, if available, should be administered.

Formative Assessment

Assessment questions, tools, and processes that are embedded in instruction and are used by teachers and students to provide timely feedback for purposes of adjusting instruction to improve learning.

Frequency Distribution

An ordered tabulation of individual scores (or groups of scores) showing the number of persons who obtained each score or placed within each range of scores.

Functional Range

The functional range of a test is the range of grades for which the test can be administered in order to obtain accurate norm-referenced data. For most tests, this range is two grades above or below the grade for which the test was intended.

Grade Equivalent

A score on a scale developed to indicate the school grade (usually measured in months) that corresponds to an average chronological age, mental age, test score, or other characteristic of students. A grade equivalent of 6.4 is interpreted as a score that is average for a group in the fourth month of Grade 6. Grade equivalents do not compose a scale of equal intervals and cannot be added, subtracted, or averaged across test levels the way scale scores can.

Grade Norm

The average test score obtained by students classified at a given grade placement.

Guessing Parameter

The probability that a student with very low ability on the trait being measured will answer the item correctly. There is always some chance of guessing the answer to a multiple-choice item, and this probability can vary among items. The guessing parameter enables a model to account for these factors.

Holistic Scoring

A scoring procedure yielding a single score based on overall student performance rather than on an accumulation of points. Holistic scoring uses rubrics to evaluate student performance.

Intelligence Test

A test that measures the higher intellectual capacities of a person, such as the ability to perceive and understand relationships and the ability to recall associated meaning--in other words, measures the ability to learn.



Interim Assessment

An assessment that occurs multiple times throughout the academic year rather than just at the end. Through an interim assessment, teachers can see weaknesses and strengths of students that otherwise may have gone unnoticed.

Interpretation

The act of explaining test scores to students so they understand exactly what each type of score means. For example, a percentile rank refers to the percentage of students in the norm group who fall below a particular point, not the percentage of items answered correctly.

Item

A question or problem on a test.

Item Bias

An item is biased when it systematically measures differently for different ethnic, cultural, regional, or gender groups.

Item Response Theory

The basis of various statistical models for analyzing item and test data. In TABE, the three-parameter model was used in the selection and scaling of items. This model takes into account discrimination, difficulty, and chance level of success (guessing) to describe each item's statistical characteristics.

K-1 Assessment

An assessment intended primarily for students in elementary and secondary schools. CTB assessments may assess students in the entire K-12 range or just in selected grades, e.g., Grades 2-12.

Local Norms

Norms that have been obtained from data collected in a limited locale, such as a school system, county, or state. They may be used instead of, or along with, national norms to evaluate student performance.

Location Parameter

A statistic from item response theory that pinpoints the ability level at which an item discriminates, or measures, best.

Mean

The quotient obtained by dividing the sum of a set of scores by the number of scores; also called "average." Mathematicians call it "arithmetic mean."

Median

The middle score in a set of ranked scores. Equal numbers of ranked scores lie above and below the median. It corresponds to the 50th percentile and the 5th decile.

Mixed-Format Tests

An assessment that includes different forms of questions. The questions could include a mix of multiple choice, essays, or performance tasks.

Mode

The score or value that occurs most frequently in a distribution.

Multiple Measures

Assessments that measure student performance in a variety of ways. Multiple measures may include standardized tests, teacher observations, classroom performance assessments, and portfolios.

Multiple-Choice Item

A question, problem, or statement (called a "stem") which appears on a test, followed by two or more answer choices, called alternatives or response choices. The incorrect choices, called distractors, usually reflect common errors. The examinee's task is to choose from, among the alternatives provided, the best answer to the question posed in the stem. These are also called "selected-response items."

Norm-Referenced Test

A standardized assessment, in which all students perform under the same conditions. This type of test compares a student or group of students with a specified reference group, usually others of the same grade and age for K–12 students, or for adults, those with similar characteristics, such as those in an adult basic education class



Normal Distribution Curve

A bell-shaped curve representing a theoretical distribution of measurements that is often approximated by a wide variety of actual data. It is often used as a basis for scaling and statistical hypothesis testing and estimation in psychology and education because it approximates the frequency distributions of sets of measurements of human characteristics.

Norms

The average or typical scores on a test for members of a specified group. They are usually presented in tabular form for a series of different homogeneous groups.

Number Correct or "Raw" Score

The Number of Correct Responses (NCR) is the number of items answered correctly by a student on any given test section.

Objective

A desired educational outcome such as "constructing meaning" or "adding whole numbers." Usually several different objectives are measured in one subtest.

Objective Test

A test for which a list of correct answers, one for each test item, can be provided so that subjective opinion or judgment is eliminated from the scoring procedure. Multiple-choice, true/false, and matching-item tests are purely objective, while short answer and completion-item tests are less so.

Percentile

One of the 99 point scores that divide a ranked distribution into groups, each of which contains 1/100 of the scores. The 73rd percentile denotes the score or point below which 73 percent of the scores fall in a particular distribution of scores. (See also the table under "stanine.")

Performance Assessment

An assessment activity that requires students to construct a response, create a product, or perform a demonstration. Usually there are multiple ways that an examinee can approach a performance assessment and more than one correct answer.

Performance Standard

A level of performance on a test, established by education experts, as a goal of student attainment.

Power Test

A test that samples the range of an examinee's capacity in particular skills or abilities and that places minimal emphasis on time limits. A "pure" power test is sometimes defined as one in which every examinee has sufficient time to complete the test.

Predictive Validity

The ability of a score on one test to forecast a student's probable performance on another test of similar skills. Predictive validity is determined by mathematically relating scores on the two different tests.

Prompt

An assessment topic, situation, or statement to which students are expected to respond.

Raw Score

The first score obtained in scoring a test, which is often the number of correct answers. Sometimes it is the number right minus a fraction of the number wrong, the time required to complete the test, the number of errors, or some other number obtained directly from the test's administration.

Readiness Test

A test of ability to engage in a new type of specific learning. Level of maturity, previous experience, and emotional and mental set are important determinants of readiness.



Reliability

The consistency of test scores obtained by the same individuals on different occasions or with different sets of equivalent items; accuracy of scores.

Rubric

A scoring tool, or set of criteria, used to evaluate a student's test performance.

Scale

An organized set of measurements, all of which measure one property or characteristic. Different types of test-score scales use different units, for example, number correct, percentiles, or IRT scale scores.

Scale Scores

Scores on a single scale with intervals of equal size. The scale can be applied to all groups taking a given test, regardless of group characteristics or time of year, making it possible to compare scores from different groups of examinees. Scale scores are appropriate for various statistical purposes; for example, they can be added, subtracted, and averaged across test levels. Such computations permit educators to make direct comparisons among examinees, compare individual scores to groups, or compare an individual's pre-test and post-test scores in a way that is statistically valid. This cannot be done with percentiles or grade level equivalents.

Selected-Response Item

A question or incomplete statement that is followed by answer choices, one of which is the correct or best answer. Also referred to as a "multiple-choice" item.

Special Admissions Test

A test of a student's ability to participate in special programs or advanced learning situations. For example, an honors-level class or a magnet school may require the attainment of high scores on an assessment for admission.

Speed Test

A test in which one aspect of performance is measured by the number of tasks performed in a given time. A "pure" speed test is one in which examinees make no errors and that cannot be completed by any examinee in the allotted time.

Standard Deviation

A statistic used to express the extent of the divergence of a set of scores from the average of all the scores in the group. In a normal distribution, approximately two-thirds (68.3%) of the scores lie within the limits of one standard deviation above and one standard deviation below the mean. One-sixth of the scores lie more than one standard deviation above the mean, and one-sixth lie more than one standard deviation below the mean.

Standard Error of Measurement

A measure of the amount of error to be expected in a score from a particular test. The smaller the standard error of measurement, the greater the accuracy of the test score. The standard error of measurement is the standard deviation of a theoretical distribution of a set of variations, each of which is the difference between the obtained score and true score. Thus, if a standard error of measurement is 5, the chances are two to one that an obtained score lies within five units of the true score.

Standard Score

A derived score scaled to produce an arbitrarily assigned mean and standard deviation. For example, deviation IQs are standard scores with a mean of 100 and, usually, a standard deviation of 16.

Standardization

The process of administering a test to a nationally representative sample of examinees using carefully defined directions, time limits, materials, and scoring procedures. The results produce norms to which the performance of other examinees can be compared, provided they took the test under the same conditions.

Standardization Sample

That part of the population that is used in the norming of a test, i.e., the reference population. The sample should represent the population in essential characteristics, some of which may be geographical location, age, or grade for K-12 students, or, for adults, participation in a specific type of program (for example, adult basic education).



Standardized Test

A test constructed of items that are appropriate in level of difficulty and discriminating power for the intended examinees, and that fit the pre-planned table of content specifications. The test is administered in accordance with explicit directions for uniform administration and is interpreted using a manual that contains reliable norms for the defined reference groups.

Stanine

A unit of a standard score scale that divides the norm population into nine groups with the mean at stanine 5. The word stanine draws its name from the fact that it is a STAndard score on a scale of NINE units.

Stem

The part of an item that asks a question, provides directions, or presents a statement to be completed.

Stimulus

A passage or graphic display about which questions are asked.

Test Battery

A test battery is a set of several tests designed to be administered as a unit. Individual subject-area tests measure different areas of content and may be scored separately; scores from the subtests may also be combined into a single score.

Test Developer

One who prepares and develops tests.

Test Item

A question or problem on a test.

Test Objective

A desired educational outcome such as "constructing meaning" or "adding whole numbers." Usually several different objectives are measured in one subtest.

Test User

One who uses test results for some decision-making purpose

Test-Taker

One who takes a test whether by choice, direction, or necessity.

Validity

The capability of a test to measure what its authors or users intend it to measure.

Web-Based Assessment

An assessment that is delivered over the World Wide Web and is accessed via a Web browser.



Resources

It's never too soon to brush up on your understanding of assessment data and its role in the classroom.

Here are some ways to get started.

Webinars:

http://www.youtube.com/user/ctbassessments?feature=results_main

Documents:

Why Data Matters e-Guide

http://www.whydatamatters.com

Getting From Here to Core e-Guide

http://www.assessmentsforcommoncore.com/eguide/index.html

Beyond the Numbers

http://www.whydatamatters.com/pdf/beyondthenumbers.pdf

Building the Best Student Assessment Solution: An Integrated Approach to student Achievement Data in an Era of Accountability

http://www.whydatamatters.com/pdf/buildingthebestdatasystem.pdf

Data requirements for the Common Core:

http://www.cde.ca.gov/ta/tg/sa/smarterbalanced.asp

Data requirements for Race to the Top:

http://www2.ed.gov/programs/racetothetop/index.html



Copyright © 2012 by CTB/McGraw-Hill LLC. All rights reserved. Acuity and LAS Links are registered trademarks and TerraNova, Writing Roadmap, Yearly ProgressPro, and InFormative Assessment are trademarks of The McGraw-Hill Companies, Inc.

Links:

CTB/McGraw-Hill Assessments and the Common Core http://www.assessmentsforcommoncore.com

CTB/McGraw-Hill Professional Development http://ctb.com/ctb.com/control/servicesDetailAction?serviceId=402&p=services

CTB/McGraw-Hill Research [http://ctb.com/ctb.com/control/ctbResear chMainAction?p=ctbResearch]

Corporate Headquarters 20 Ryan Ranch Road Monterey, California 93940 U.S.A.

Customer Care 800.538.9547

CTB.com