Assessment for Learning:

12 recent studies on formative assessment and aligning assessments with learning goals

Published by Educational Research Newsletter
Foreword

A growing theme in the field of assessment is the notion that assessment is a tool for advancing learning, not only a tool for measuring what students have already learned.

Teachers have made informal use of formative assessments for years, but now many districts want to take a more formal approach to monitoring students’ progress during the year so that teachers can modify instruction for individual needs and give students feedback to help meet learning goals.

Before adopting a formative assessment system, districts must make many preliminary decisions. One of the key questions is to ask who will be using information from formative assessments—will it be used by administrators, teachers or both? What infrastructure and competencies (e.g. ability to analyze student work and design effective instructional improvement strategies) do the schools have? What investments will be needed in infrastructure and in building staffing competences before a formative assessment process can be implemented?

Teachers often complain that preparing students for statewide assessments takes away from instruction time. Classroom-based performance assessments can bridge the gap between learning and the reporting and accountability functions of statewide tests. New York State has developed a classroom-based assessment process for K-3 students called the Early Literacy Profile intended to be an instructional assessment.
According to a recent study in *Educational Assessment*, a formative assessment system for New York K-3 students was intended to set itself apart from traditional testing:

- by examining students’ literacy skills in the context of everyday classroom activities;
- by reporting student performance over time rather than in a one-time snapshot;
- by referencing itself to standards and stages of development; and
- by providing information useful for reporting without the high stakes attached to most tests.

A new special report from *Educational Research Newsletter*, *Assessment for Learning: 12 recent studies on formative assessment and aligning assessments with learning goals*, includes more information about New York’s formative assessment process as well as strategies used by schools in Wisconsin, Texas, the UK and Canada to advance learning goals with assessments.

The 12 studies featured in this special report provide a range of perspectives from the latest educational research on how learning and assessment work together. This set of studies is intended to give educators a sense of current thinking on the interplay between learning and assessing.
# Table of Contents

## At the district level...

**Key decisions a district must make before choosing a formative assessment system**  
*American Journal of Education*  
Page 6

**Frequent assessments of student progress reduce pressure on teachers to teach to the test**  
*Education and Urban Society*  
Page 9

**Aligning assessments with major reforms in math curriculum: Lessons from Ontario**  
*Assessment in Education: Principles, Policy & Practice*  
Page 12

**Getting formative information for the classroom from large international and national math assessments**  
*Assessment in Education*  
Page 15

**Kentucky’s demanding portfolio assessment for writing provides enriching—and cost-effective—professional development for teachers**  
*Educational Forum*  
Page 17

**Scale measures teacher confidence in developing standards-aligned classroom assessments**  
*Educational and Psychological Measurement*  
Page 20

## In the classroom...

**Classroom-based assessment makes preparing students for statewide test more instructionally useful**  
*Educational Assessment*  
Page 22

**How classroom-based assessments affected one teacher’s practices**  
*The Reading Teacher*  
Page 25
Questions to and from students in the classroom can work as ‘assessment conversations’
Journal of Research in Science Teaching 28

Teachers describe challenges of developing formative assessment tasks
Assessment in Education 31

Class size has little effect on teachers’ assessment practices
The Alberta Journal of Educational Research 33

Assessment may influence students’ learning approaches but not always in the expected ways
Educational Studies 35
Key decisions a district must make before choosing a formative assessment system


Teachers have made informal use of formative assessments for years, but more and more districts want to take a more formal approach to formative assessments to ensure that their students are on track to meet statewide assessments. Before adopting a formative assessment system, districts must make many preliminary decisions, reports this study in the American Journal of Education.

Conclusion: Before choosing a formative assessment system, one of the most important decisions a district must ask is: Who will be using the formative assessment results—administrators or teachers? A district also must assess its own infrastructure and set of competences. Schools where educators have already learned to work collaboratively to analyze student work and design effective instructional improvement strategies have different needs for implementing such a system than do schools that have little capacity to do this work.

Main research question: What are the key decisions to make before developing or choosing a formative assessment system?

Participants: 42 individuals from one large urban district that was implementing reform based on the increased use of data. Participants included the superintendent, district-level administrators, teachers, school-based administrators and the vendor that provided the computer-based assessment program to the district.

Method: Interviews with participants that focused on how teachers and administrators use student assessment results. Researchers also reviewed documents related to district’s recent history and current challenges.
LESSONS LEARNED:

- If a computer-based system is chosen, additional technology investment such as data storage and software may be needed by the district, depending on its existing infrastructure.

- Teachers and administrators may need professional development to help them analyze student assessment results and translate the knowledge effectively.

- It’s important to consider the existing competences of educators for working collaboratively on designing instructional improvement.

Findings: What is the purpose of the assessment is a key question that should guide the choice of a formative assessment system. The superintendent in the school district at the center of this study saw four roles. Formative assessments could assist:

- teachers in planning instructional interventions;
- principals in developing improvement plans for students with lagging skills;
- the superintendent in identifying schools that are not making adequate progress; and
- updating the school board on the likely performance of students on the next round of the state’s math tests.

Among the key decisions districts must make in selecting a system are:

Compulsory or voluntary? If the assessments are to be used by the central office to identify struggling schools, then assessments will have to be scheduled at the same time with the same test administration and security methods, the researchers note. One disadvantage of mandatory assessments is that faculty may see the process as one more externally imposed mandate rather than as a source of information for them. If the primary goal of formative assessments is to provide educators with information, the design choices would be different than if the goal is to hold principals accountable.
**Make or buy?** Buying a system is faster than developing a homegrown system, researchers report. Vendors have done the work of aligning the assessments with state standards and assessments. However, the cost of purchasing a commercial system is usually higher (cost can range from $5 to $75 per student per year). Teacher buy-in is usually higher with a system that is developed internally.

**Computer-based or paper-based?** One major advantage of computer-based assessments is quick and easy scoring as well as graphic summaries of results. Another advantage is that students like taking tests on computers. An important disadvantage of computer-based testing is that all items are multiple-choice questions so that educators do not get information about students’ abilities to answer open-ended questions. Computer-based testing can also place a strain on a district’s computer infrastructure.

**Computer-adaptive assessments?** Computer-adaptive testing allows for testing of students with widely divergent skills. A short test can test the abilities of students with low skills while a longer test is needed to test students with higher skills. One disadvantage of computer-adaptive assessments is the considerable amount of time that is needed to test the higher-ability students—an hour in one of the pilots tested in this school district. Another disadvantage, as with any computer-based testing, is the demands it places on the infrastructure. Because of bandwidth issues at the district in this study, there was also a long lag between the student answering a question and the next question appearing on the screen, causing the student to lose interest in the process.
Frequent assessments of student progress reduce pressure on teachers to teach to the test

Study: “Can Rapid Assessment Moderate the Consequences of High-Stakes Testing?” by Stuart S. Yeh, Education and Urban Society, Volume 39, Number 1, November 2006, pp. 91-112.

A major criticism of No Child Left Behind is that it puts pressure on teachers to drill students so that they will perform adequately in statewide assessments with the potential consequence of a narrowed curriculum. Texas is considered a forerunner in changes that are expected under NCLB because of its early adoption of a high-stakes testing program. To help reduce the pressure of testing on teachers and students, some school districts have implemented assessment programs that provide rapid diagnostic information about student progress to teachers.

Conclusion and implications: Frequent assessments of student progress can give teachers a sense of greater control over student achievement and reduce the pressure teachers feel to increase test scores through whole-class drill and practice, concludes this study from Education and Urban Society. The study found that most teachers in a district that used rapid assessments had positive attitudes towards state-mandated testing and used test results to guide their teaching. Results suggest that use of a rapid feedback system can help teachers prepare students for state-mandated tests while maintaining a balanced curriculum.

Main research questions: How did teachers in a district that had used a rapid assessment system for several years view the impact of state-mandated testing on the curriculum? Did they view it as having a positive or negative effect? Did the teachers feel that district- and state-mandated test results improved instruction in a balanced way? How did teachers and administrators in the district feel about state-mandated testing?
Participants: Teachers, administrators and other staff in the McKinney, Texas school district. The participants included 37 teachers for grades 2-12, 11 principals and assistant principals and librarians in 8 schools. The district had used a curriculum-based rapid assessment system from StandardsMaster since 1999.

Method: Participants were interviewed individually for 50 minutes using a semi-structured interview protocol. Researchers conducted observations in 10 randomly selected classrooms and interviewed three to five randomly selected students from each classroom to assess quality of instruction.

LESSONS LEARNED:
- Frequent assessment of student progress can reduce pressure on teachers and help them avoid teaching a narrowed curriculum. Previous research has shown that districts respond well in conditions where stakes are high, but pressure is low.

- Providing teachers with periodic assessments of student progress supports more effective and efficient individualized instruction that keep students at grade level.

- High-stakes tests may be viewed by teachers as having positive effects if a state-mandated test is aligned with the curriculum.

Findings: A majority of teachers (77.6%) in this district believe that state-mandated testing had a positive impact on the curriculum. They agreed that the state-mandated test encouraged teachers to focus instruction in a positive way on key skills and knowledge.

A majority of teachers (89.8%) reported using results from the periodic curriculum-based tests and state-mandated tests to systematically improve instruction. Teachers said they used results to identify areas of difficulty and to decide how much time to spend on particular topics.
Some 65.3% reported that they used the results to identify and teach missing skills, including higher order thinking skills and basic skills. Teachers noted that the curriculum is determined by the state and that it was not the test results that caused narrowing of the curriculum.

Although one-half of the teachers (49%) reported that they have reduced the amount of time they spend on projects, almost all (91.7%) reported that they eliminated unimportant projects and refocused the remaining projects in a way that benefited students.
Aligning assessments with major reforms in math curriculum: Lessons from Ontario

**Study:** “The challenges and possibilities of aligning large-scale testing with mathematical reform: The case of Ontario.” Assessment in Education: Principles, Policy & Practice”, by Alex Lawson and Christine Suurtamm, Volume 13, Number 3, November 2006, pp. 305-325.

In 1997, the Ontario government undertook an ambitious reform of its elementary school (K-8) mathematics program. The reform emphasized math instruction using challenging problems, student construction of multiple solution methods, and mathematical communication and defense of ideas. In a recent issue of Assessment in Education, two researchers examine Ontario’s efforts to align a large-scale assessment with a reformed math curriculum.

**Conclusion:** Math reforms emphasize skills such as problem-solving, student construction of multiple solution methods, and mathematical communication and defense of ideas. When aligning math assessments with a reformed curriculum, educators face the challenge of how to assess these distinct skills that students rarely use in isolation.

Testing should attempt a more global assessment students interrelated uses of these skills. When revising assessments, educators should incorporate the latest research in math learning and education. A more in-depth and up-to-date understanding of the mathematical development of children would also benefit teachers in the classroom as they attempt to make shifts in instruction and gauge students’ understanding and progress.

**Main research question:** How can schools that have reformed the way they teach mathematics align assessments with the new math curriculum and with the latest research on math education?
**Participants:** Educators involved in a major reform of the mathematics program for K-8 launched in Ontario in 1997.

**Method:** The authors report on the experience of Ontario educators as they worked to refine assessments to measure student math development in reform-inspired classrooms.

**LESSONS LEARNED:**
- A reformed math curriculum requires a major re-examination of how students should be assessed.
- Use of a simple matrix of skills and content areas, while a popular tool for designing an assessment, can lead to tests of these skills in isolation. Since students make interrelated use of these skills, more global assessments are needed.
- New assessments should reflect the latest research in how children learn mathematics.

**Background:** The Education Quality and Accountability Office (EQAO), which was developed at the time of reform, developed a matrix of five content areas and four cognitive categories to revise assessment of students:

<table>
<thead>
<tr>
<th>Content areas</th>
<th>Cognitive areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeration and number sense</td>
<td>Communication</td>
</tr>
<tr>
<td>Data management and probability</td>
<td>Problem-solving</td>
</tr>
<tr>
<td>Algebra and patterning</td>
<td>Understanding concepts</td>
</tr>
<tr>
<td>Geometry and spatial sense</td>
<td>Procedures</td>
</tr>
<tr>
<td>Measurement</td>
<td></td>
</tr>
</tbody>
</table>

Teachers and consultants helped create and test field items for each cell in the matrix, resulting in 20 performance items. The rest of the test was made up of multiple-choice items to provide continuity from year to year.
**Findings:** While the matrix seemed to offer a simple solution, the problem was test items developed by educators had isolated these processes into discrete categories. For example, while communication was one of the cognitive areas the province wanted to measure in the assessment, it is difficult to measure communication apart from mathematical ideas.

A monograph from the Third International Mathematics and Science Study (TIMSS, 1996), warns that use of a process-category-by-content-strand matrix, while it has a long history, generates such serious difficulties. Use of such a matrix, states the TIMSS monograph, “failed to take into account the interrelatedness of content or of cognitive behaviors, and that this forces the description of information into unrealistically isolated segments.”

Many problems require all four processes at once: problem-solving, calculating, understanding and communication. So rather than assessing one cognitive process at a time, each item should evaluate all processes, the researchers note. This would support the vision of mathematics as “problem-solving in action” as opposed to the view that it is that is “mechanistic and atomized”.

The researchers stress the importance of incorporating the latest research on children’s mathematical learning in revising assessments, i.e., what do mathematical understanding and proficiency look like in Grade 3? Research could provide a more accurate description of what a child is capable of doing and help educators better assess progress.
Getting formative information for the classroom from large international and national math assessments


Large-scale international and national assessments monitor the health of an education system, but they can also provide valuable formative information, argues this study from Assessment in Education.

Conclusion: Large-scale international and national mathematics assessments can provide formative information for teachers if the data is reworked and re-presented in formats more appropriate for educators, according to a recent article in Assessment in Education. Information about student achievement can be used to make inferences about teaching practices that can be used at the classroom level, writes researcher Brian Doig, who presents examples of some re-analyses of national and international data.

Main research question: How can data from large-scale international and national mathematics assessment reports be used to inform teachers about the effects of their classroom practices?

Method: The author provides examples of how researchers and research organizations, using Item Response Theory, have re-analyzed mathematics data from the Third International Mathematics and Science Study (TIMSS) and other large-scale assessments to provide data that could be used by educators for formative purpose.
LESSONS LEARNED:

- Large-scale math assessments contain valuable formative data but it is hidden unless results are re-analyzed and presented in ways more suitable for educators.

- To be useful, this data must include implications for the classroom and the basis for inferences about educators’ teaching practices.

- Data from national and international assessments can help increase understanding of student needs and achievement and contribute to reforms of educational systems.

Findings: Below are some examples of how researchers have reanalyzed data to provide formative information to educators:

- Scale anchoring of TIMSS data by one research organization helped provide a greater understanding of what 9-year-old and 14-year-old students have learned in mathematics. Researchers identified items that students at different levels had a high probability of answering correctly. Those items were provided to mathematics education experts, who described what understandings and skills students needed to answer the items correctly.

- In New South Wales, Australia, researchers reanalyzed results for a half-million students from its Basic Skills Testing Program for Year 3 and Year 6. Students’ possible responses to questions were examined in terms of the percentage of students that selected each possible response. Selection of responses was also broken down by the percentage of students that selected each response by gender and subgroup. To emphasize the classroom relevance of the expert analysis, each section concluded with a teacher’s observations on classroom implications.

- Exploring errors by students can also provide useful information, according to researchers from the University of Manchester in the UK who did an error analysis of national mathematics tests taken by English and Welsh students. Error analysis is based upon the notion that it can provide important diagnostic information for teachers. Ability estimates were used to compute the mean ability of students who made each error. A student’s total score and the errors likely to have been made provide information about student achievement. This type of test review can help support and educate teachers and lay the foundation for better practice, according to the UK researchers.
KY’s demanding portfolio writing assessment doubles as professional development for teachers


As more is spent on testing, fewer resources are available for meaningful professional development, says an article in Educational Forum. One benefit of the state of Kentucky’s demanding portfolio assessment for writing is that it doubled as a cost-effective professional development program, the authors write. Kentucky was among the innovators in taking portfolios out of the classroom into the arena of large-scale testing and assessment. This sophisticated and demanding approach to evaluating writing provided teachers with a rich professional development experience.

Conclusion: Kentucky’s statewide use of portfolio assessment for writing is labor-intensive and poses significant accountability and reliability issues. However, a major benefit is that it has stimulated better professional development in teaching writing. The portfolio review process stimulates reflection and exchange of ideas among the team of teachers who perform the writing assessments for their schools. Student portfolios are scored by a cross-curriculum team of teachers within the school. To support the state’s teachers, the state department of education also has created sophisticated professional development courses and materials for teaching writing,

Main research question: Is high-stakes writing assessment compatible with high-quality writing and high-quality professional development?

Method: Since 1992, the authors, a writing specialist and English educator and a literacy assessment developer for a national reform coalition, have analyzed and criticized Kentucky’s portfolio systems and studied its impact on schools, English departments, teachers and students.
**LESSONS LEARNED:**

- Portfolio assessment is a more sensitive measure of quality writing that demands thoughtful and sophisticated reading and scoring.

- Portfolio assessment can stimulate expanded and cost-effective professional development for teachers and can focus resources on the complex process of nurturing good student writing.

- High-quality professional development encourages learning through joint productive and professional relevant activities among teachers. It contextualizes teaching and challenges participants to find more complex solutions in addressing problems.

**Findings:** Teachers outside of the English department were expected to create writing assessments, teach students how to complete them and evaluate the results, presenting additional challenges for this already ambitious assessment system.

One teacher and researcher described the first cross-curricular scoring session in her high school as “certainly the single, most important professional development activity the faculty had that year, or perhaps any year. The faculty moved toward the goal of all teachers using writing as a learning tool in their classrooms.” Teachers have spent many hours discussing the characteristics of good writing, the types of assessments that generate it, and about potential changes in the curriculum.

The state department of education has developed workshops and materials, which now include videos on such topics as teaching poetry and using writing to learn. The materials have become more sophisticated and increasingly directed toward improving classroom instruction, the researchers note. The “Marker Papers”, for example, are collections of student writing from every grade level that illustrate how writing develops over time.
Kentucky writing teachers also participate in professional development at eight active Writing Project sites. The Writing Project, which has received significant funding from the state, brings together elementary, secondary and college teachers from all disciplines to create professional communities outside of school and district boundaries. From 1997-2003, the Writing Project sites served more than 80,000 teachers and students.

“Everyone knows that Kentucky’s portfolios are a labor-intensive accountability test, but few are aware that they also are a stunningly cost-effective form of professional development,” the researchers conclude. “Dollars that would have gone to testing companies to provide tests and scoring are being used instead to improve the knowledge and skills of Kentucky teachers.”
Scale measures teacher confidence in developing standards-aligned classroom assessments


Most states have standards that serve as guidelines for what students should be learning at each grade level. But, unless teachers believe that standards are reasonable and usable and unless they feel comfortable and confident in their understanding of those standards, it is unlikely that classroom assessments will be successfully aligned with statewide learning standards, say the authors of a recent study in Educational and Psychological Measurement. This article reports on a scale that could guide efforts to train teachers in aligning classroom assessments with standards.

Conclusion: A program evaluation developed as part of a professional development program (Standards-Aligned Classroom Initiative in Illinois) could be used as a scale to identify teachers’ familiarity with and attitudes towards state standards, the authors report. The scale could be used to identify issues that need to be addressed in training teachers to use standards for developing classroom assessments and to gauge the impact of any training activities.

Main research questions: Does the Teacher Assessment Efficacy Scale (TAES) provide a valid measure of teacher efficacy in developing standards-aligned classroom assessments?

Participants: 642 Illinois teachers who participated in a program conducted by the Standards-Aligned Classroom Initiative

Method: Teachers completed a 42-item questionnaire that was designed to assess teachers’ attitudes towards learning standards. The scale measured six domains:
   a) Confidence—teachers’ confidence in using the state standards as a basis for creating assessments;
   b) Impact—teachers’ beliefs about whether using the standards benefits instruction;
   c) Use—whether teachers plan to use the standards as a basis for assessments
d) Utility—whether teachers believe using the standards for assessments is productive and useful;

e) Experiences—teachers’ familiarity with the standards; and

f) Students—teachers’ beliefs about the benefits of including students in process of developing standards-aligned assessments

**LESSONS LEARNED:**

- An important hurdle in helping teachers in their efforts to create standards-aligned classroom assessments is building their confidence and comfort level with standards.

- Professional development activities can give teachers a better understanding of standards and train them in aligning classroom assessments with those standards.

- A scale of teacher efficacy can help gauge teachers’ level of comfort with this task before and after professional development activities.

**Findings:** The study found that it is useful to differentiate teacher attitudes toward educational standards with the six domains listed above.

Other observations from the study:

- Professional development activities on developing standards-based assessments resulted in large increases in TAES scores.

- Because this scale was developed to be consistent with the Illinois state standards, validity and generalization would have to be conducted to support wider use of the instrument.

- The scale could be used to guide efforts to develop teacher competence in creating standards-aligned assessments and to refine professional development programs by indicating which areas of the assessment alignment process are most problematic.
Classroom-based assessment makes preparation for statewide test more instructionally useful


Teachers often complain that preparing students for statewide assessments takes away from instruction time. Classroom-based performance assessments can bridge the gap between learning and the accountability functions of statewide tests, says this study from *Educational Assessment.* The study reports on New York’s Early Literacy Profile for K-3 students. Created in the late 1990s as an “instructional assessment” useful for both teaching and reporting, the Early Literacy Profile monitors the progress of K-3 students toward meeting state standards.

Conclusion: In field-testing, the Early Literacy Profile classroom-based assessment was found to be instructionally useful by educators. 98% of survey respondents felt that Profile data provided information about students’ literacy progress that was useful to teaching. Teachers noted that looking at student work in relation to standards increased their knowledge of individual students and their abilities and guided them in what they needed to do next for those students. Many teachers reported a shift toward relying more on evidence rather than subjective feelings as the basis for instructional decisions.

Main research questions: What are the components of a research-based classroom assessment of literacy? What is the reliability and validity of such an assessment? How useful is such an assessment for both teaching and reporting purposes?

Participants: 19 schools from 19 New York State school districts, 63 teachers and 1,215 students Grades 1-3 involved in field-testing the Early Literacy Profile.
Method: Technical studies on the Profile were conducted to see if the Profile could (a) measure important literacy traits; (b) be scored reliably; (c) predict student performance on state tests (d) be consistent with state standards (e) provide information useful to teachers, parents and the community. Participants were surveyed to assess their views on curriculum validity and on how the Profile impacted their teaching.

Background: With the Profile, teachers take four early literacy measures of each student:

- reading in use,
- early reading skills,
- student’s writing, and
- oral language development.

Teachers use standardized multiple tasks and gather evidence over time throughout the year. A 10-20 minute “reading interview”, during which the teacher observes the student reading, is a key part of the Profile. Students and teachers also record a list of works that students read independently and students write a written response to a text as a measure of reading comprehension.

Data is collected and evaluated in the fall and spring of each year in relation to scales that describe key qualities of literacy learning along a developmental continuum. As teachers examine each student’s work in relation to the scales, they select the level of proficiency based on descriptions of skills—emergent, beginning, independent and experienced. Proficiencies do not fall in one column or scale point, but are spread across the columns.

LESSONS LEARNED:

- Examining a student’s work in relation to the Profile scales deepens teachers’ knowledge of literacy development.

- Embedding assessment into classroom life and involving teachers in the scoring process can help make statewide testing more instructionally relevant for teachers.

- Classroom-based assessments give teachers immediate feedback on student progress.

- Double-scoring of classroom-based assessments is a valuable professional development exercise.
**Findings:** Profile assessments were found to be predictive of performance on the state’s English Language Arts (ELA) exam. Researchers looked at students’ fourth-grade scores on ELA and found “moderately strong” correlations between performance on the ELA and the Profile. All 63 teachers were surveyed about the curriculum validity of the Profile—98% of respondents said that the ways of collecting evidence for the Profile resembled the kinds of activities they provided for students in their classrooms.

Some 98% felt that Profile data provided information about students’ literacy progress that was useful to teaching. “I learn much more about a child from the evidence I collect about his learning in the context of classroom activities than by the score received on the standardized test,” wrote one teacher. All of the teachers and administrators felt that the Profile was useful in helping parents understand their child’s literacy progress. Among the challenges, according to 20% of field-test teachers, is to do the Profile evaluation of students in a reasonable amount of time.
How classroom-based assessments affected one teacher’s practices


Standardized tests provide a snapshot of a student’s comprehension on one day, according to a recent article in The Reading Teacher. But, classroom-based assessments measure a student’s often-uneven progress over time. A classroom-based assessment process can be an important complement to standardized measures by helping teachers to modify instruction to fit student needs and to give their students more valuable feedback on their progress, say the authors of this study of one teacher in a Wisconsin school district.

Conclusion: In a school district that developed a comprehensive, classroom-based assessment process for reading comprehension, the impact on one fourth-grade teacher’s instruction was significant. Classroom-based assessments helped her to gain insight into the students’ thinking while they were reading, get a more accurate impression of the capabilities of students who were not performing at grade level and develop a more accurate assessment of the skills of one student with a learning disability. They also helped her to develop reading comprehension goals for students.

Main research question: How does a comprehensive approach to classroom-based assessment influence teacher practice?

Method: Researcher Judy Fiene spent a year investigating how a Wisconsin school district’s comprehensive, classroom-based assessment process influenced the teaching of one experienced fourth-grade teacher. Co-author Susan McMahon served as a consultant to the district, working to develop, evaluate and modify the assessment process.
Background: The district encouraged teachers to examine students’ specific, ongoing comprehension needs and to design instruction accordingly. Teachers applied a four-point rubric at least twice a year and district reading specialists worked with teachers to support ongoing assessment. The district encouraged teachers to provide differentiated instruction to students.

LESSONS LEARNED:

• Essential to identifying students’ needs is monitoring their comprehension processes while reading.

• Teachers need to find ways of encouraging students to make their thinking and the strategies they are using visible.

• Asking students to take notes while they are reading provides a rich source of material for teachers on what students are thinking while they are reading.

• Strategic questioning is an important aspect of comprehension. Students should be taught that active readers question the author, the text and themselves during reading.

Findings: Among the researcher’s observations:

• One of the challenges teachers face is figuring out how to monitor students’ comprehension processes while reading. The teacher gave students sticky notes to paste their questions and comments as they were reading. To get a more in-depth look at student comprehension, the teacher asked students to use the notes as a basis for extended writing in a journal.

• There were three dominant patterns to the teacher’s approach: She modified instruction based on the comprehension needs of her students, often modeled asking questions as one facet of comprehension, and stressed text structures through the use of graphic organizers. Graphic organizers, such as using separate columns to separate factual information, connections, opinions and the author’s craft, deepen a student’s understanding. They also provide the teacher with a
level of understanding about student comprehension not possible with standardized testing.

- With the classroom-based assessments, the teacher could see that one student with a learning disability was using new comprehension strategies sometimes, showing that she had the potential to make inferences. Another student who was performing at grade-level was raising questions in her reading notes that could not be answered in the text, potentially creating frustration for her. The teacher helped the student understand that some questions were her own questions and that the answers were not to be found in the text. An instructional focus for this student was to help her express her thinking more clearly and help her identify inferences in a text so that she could find answers to her own questions.
Questions to and from students in the classroom can work as ‘assessment conversations’


Questions to and from students in the classroom can be more than a tool to encourage participation or to check that students are paying attention, according to a recent article in Journal of Research in Science Teaching. Questions can serve as opportunities for “assessment conversations” that help teachers recognize students’ conceptions, mental models, strategies or language and to use that information to guide instruction.

Conclusion: Students performed better in science when their teachers make effective use of informal formative assessments in the classroom. An ESRU model was developed to guide teachers in conducting these informal assessments. The ESRU model is as follows: The teacher elicits a question, the student responds, the teacher recognizes the student’s response and then uses the information to support student learning. The ESRU model differs from the more common IRE/F model (initiation, response, evaluation/feedback) in that there is greater emphasis on using responses from students to reach learning goals, e.g. asking another question that challenges or redirects the student’s thinking.

Main research question: Can a model of informal formative assessment be used to judge the quality of informal practices across different teachers? Can the quality of teachers’ informal formative assessment practices be linked to student performance?

Participants: Three middle school teachers who were teaching the first physical science unit of the Foundational Approaches to Science Teaching (FAST 1), “Properties of Matter”, and their students. The FAST curriculum is a constructivist, inquiry-based, middle-school science education program developed by the
Curriculum Research and Development Group (CRDG) and aligned with National Science Education Standards.

**Method:** The teachers were provided with digital video cameras, microphones and videotapes and asked to videotape every session of the unit. Videotapes (30 lessons by the three teachers) were transcribed and 26 assessment conversations that involved teacher-whole-class interactions were identified in the transcripts. The conversations were coded as complete or incomplete ESRU cycles and they were also coded for whether they focused on the “epistemic” (relating to knowledge) or the “conceptual.”

To assess student learning, researchers used two sources of information: a multiple-choice achievement test as a pretest and three assessment activities embedded in three sessions—graphing, prediction question and predict-observe-explain.

**LESSONS LEARNED:**
- Informal formative assessment can take place in any student-teacher interaction (whole-class, small group or one-on-one).
- A student’s incorrect response or question can give the teacher assessment information about a student’s misunderstanding.
- Teachers can “use” information from classroom conversations by encouraging the contrast of students’ ideas, making connections between new ideas and familiar ones or increasing the difficulty of the task at hand.
- Evaluations by themselves (“good” or “excellent”) are not part of the ESRU pattern unless they are embedded in a rationale for the evaluation provided.
- Frequent and ongoing assessment conversations and assessment activities can help develop habits of scientific inquiry in students; instructional responsiveness is crucial in teaching scientific inquiry.
Findings: The ESRU model provided important information about the teachers’ informal assessment practices in the classroom, the study found. Based on this small sample, the researchers conclude that the teacher whose whole-class conversations were more consistent with the ESRU cycle, had students who performed better on the assessment activities embedded in the sessions.

Most of the conversations in the transcripts were coded as conceptual conversations; there were few epistemic conversations in the classrooms. Examples of epistemic conversations include comparing and contrasting observation, data or procedures, making predictions, formulating explanations and evaluating the quality of evidence.

The highest percentage of the conceptual questions focused on asking students to define concepts (44%) and checking students for understanding (32%). Few conceptual conversations focused on comparing or contrasting concepts. The most common epistemic conversations focused on applying known procedures and making observations.
Teachers describe challenges of developing formative assessment tasks


Many teachers lack a clear understanding of the differences between formative and summative assessment, according to a recent study in Assessment in Education. This shaky understanding is one of the challenges teachers face when they try to develop formative assessment tasks for their classrooms, write the authors of this study of 12 UK teachers developing assessments tasks for geography.

Conclusion: With the introduction of the National Curriculum in England, teachers had to change their assessment practices to assess pupils against national standards. In order to plan formative assessment tasks, teachers needed to have a working map of progression in geographical learning throughout the key stages. Among the challenges in developing formative assessments for the classroom were: communicating the differences in levels of achievement to students and developing differentiated tasks for students of varying abilities.

Main research questions: How do teachers devise assessment tasks that evaluate differentiation and progression? How are assessment tasks, feedback, feedforward (target setting) and pupil self- and peer-assessment related? How do teachers respond to level descriptions in interim assessment activities?

Participants: Heads of geography and humanities departments

Method: Observations and pupil and teacher interviews with a semi-structured framework. Teachers were expected to refer to levels of achievement in geography to assess students based on four distinct progression strands in geography: inquiry and skills, place, patterns and
processes and environmental change and sustainable development.

**LESSONS LEARNED:**

- One of the key concepts of formative assessment is teacher feedback and feedforward, which works in two ways—to help the pupil improve learning and enable teachers to make adjustments to their teaching. Feedback is designed to close the gap between pupils’ current and desired level of learning and performance.

- If information about a gap in learning is just recorded, then the action is not formative. Active involvement of the pupils in assessment of their learning is needed with a particular emphasis on how they are progressing.

- Teachers reported using checklists to provide feedback and feedforward. The checklists helped pupils understand what they were expected to achieve and how they were going to be judged.

- When they reviewed their written comments to students, teachers found that providing positive reinforcement, especially to low-achieving students, detracted from the quality of the formative assessment.

**Findings:** Among the challenges teachers reported were:

- Providing differentiated tasks, especially for higher-achieving students. Differentiation is defined as planned intervention to enable pupils with different learning needs to achieve their highest levels. It was easier for teachers to provide differentiated tasks for students at lower ability levels.

- Eight of 12 teachers found it difficult to articulate progression, or what might be meant by increasing depth in pupils’ understanding.

- Target setting was used by teachers for communicating to their pupils the next steps in their learning. Nine of 12 teachers used assessment checklists for target setting and to guide self- and peer assessments.

- Teachers found it difficult to translate the assessment criteria into language their pupils could understand and also to communicate the levels of achievement.
Class size has little effect on assessment practices


With the major shift toward using classroom assessments to promote learning rather just to measure it, it’s important to have a better understanding of teachers’ classroom assessment practices and what factors influence these practices. A recent study in The Alberta Journal of Educational Research looks at how class size, school size and subject area affect teachers’ grading and assessment practices.

Conclusion: Class size has surprisingly little effect on teachers’ assessment practices, concludes this study from The Alberta Journal of Educational Research. School size was not a significant factor either, according to this research based on a survey of Canadian high school teachers. A more important factor in grading and assessment practices is the teacher’s subject area. Mathematics teachers were less likely to take into account behaviors such as motivation, effort, paying attention, etc. in their grading. They were less likely to use more subjective assessments (constructed-response assessments such as structured observations and essay-type questions).

Main research questions: How do factors such as class size, school size and subject area affect teachers grading and assessment practices?

Participants: 513 secondary school teachers in one Western Canadian province

Method: Teacher responses to a 34-item questionnaire were analyzed. The questionnaire included a section for demographic information about grade level taught, subject area, number of students in the class and total number of students in the school. Teacher responses were collected with a six-point Likert-scale (not at all=1; completely=6).
LESSONS LEARNED:

- The study did not support the position that teachers actually vary their assessment practices based on class size.

- High school teachers used non-cognitive practices such as effort, improvement and behavior, indicating that many teachers perceived a need to “pull for borderline cases.”

- Understanding why teachers make the assessment decisions that they do will help modify assessment principles, ultimately improving instruction and student learning.

Findings:

- Class size and school size were not significant factors in teachers’ grading and assessment practices.

- Mathematics teachers placed less emphasis on academic enabling behaviors (i.e., ability level, student effort, paying attention, improved performance, work habits and level of disruptive behavior) than teachers in English, social studies and practical arts (industrial arts, home economics) and other subjects.

- Mathematics teachers preferred using grouped quizzes with objective assessments (performance quizzes, multiple-choice tests, assessments that measure student recall and major exams) rather than more subjective assessments.

- Science teachers also emphasized grouped quizzes with objective assessments more than teachers in English, practical arts and other fields.
Assessments may influence students’ learning approaches but not always in expected ways

Study: “Students’ assessment preferences and approaches to learning: Can formative assessment make a difference?” by David Gijbels and Filip Duchy, Educational Studies, December 2006 Volume 32, Number 4, pp. 399-409.

A n important challenge for educators is fostering students’ use of deep learning strategies such as making inferences, thinking critically and analyzing and synthesizing ideas and information. Research suggests that assessments that demand use of higher-order thinking skills can help stimulate deep approaches to learning and help discourage a “surface” approach to learning. This study explores what happens to a group of students who are exposed to formative assessments that demand use of deep-learning skills.

Conclusion: Students who have direct experiences with formative assessments encouraging deeper approaches to learning, actually showed higher preferences for more surface-learning and assessment approaches, concludes recent research in Educational Studies. These students also changed their learning approaches in favor of more surface approaches. These findings are in line with other research findings, the authors report, suggesting that inducing a deep approach to learning with assessments is problematic. Other factors in the learning environment could account for these results.

Main research question: What is the relationship between students’ assessment preferences and their approach to learning? Do students’ assessment preferences change when they have experienced a formative mode of assessment? Do students adopt a deeper approach to learning when they have experienced a formative mode of assessment?

Participants: 108 first-year university students in a criminology course (a total of 298 students participated in pre-test and 221 students in post-test)
Method: To characterize their approach to learning (deep vs. surface approach), students took the Revised two-factor study process questionnaire. Surface learning approaches refer to student intentions to learn by memorizing and reproducing the factual content of study materials. Deep learning approaches refer to student intentions to understand and construct the meaning of the content to be learned.

Students completed group assignments that were used as a formative mode of assessment. The four group assignments were as follows:
1) Define a scientific, researchable topic;
2) develop a step-by-plan for searching for resources;
3) search for relevant resources according to the plan; and
4) formulate a research question based on the resources.

LESSONS LEARNED:
- Unless difficulties in the learning environment are resolved, formative assessments promoting higher order thinking skills are unlikely to influence students’ learning approaches.
- Assessment of students’ achievement is seen as an important factor in encouraging students to adopt these kind of learning strategies.
- Students generally shift between surface and deep approaches to learning to suit the assessment demands of their courses.
- It may be difficult to “induce” a deep approach to learning if it is not already there.

Findings: In general, students with learning preferences for deeper approaches to learning preferred assessments that demanded use of higher order thinking skills. Students with a surface approach to learning were less likely to prefer assessments that featured higher-order thinking tasks.

This study found that formative assessments encouraging higher-order thinking skills had the opposite effect on many students’ learning and assessment preferences after they had an experience with such assessments.
Other factors in the learning environment could help account for these results, the researchers say. Previous research by Nijhuis et al. (2005) indicates that three elements of the learning environment influence students’ learning strategies: clarity of goals, appropriateness of the workload and usefulness of the literature. Other research has shown that the quality of the learning environment and the educational setting influence students’ learning approaches. Heavy course workload and other features of the learning environment such as structure or feedback can deter students from pursuing deeper approaches to learning.