The Road

Opting out of standardized state testing, Nebraska has developed its own system of local assessments.

Chris W. Gallagher and Suzanne Ratzlaff

Suzanne and 4th grader Micah sit side by side in the computer lab, comparing notes. Suzanne explains how she has filled out the visual rubric assessing Micah's research report on Buffalo Bill Cody. Then Micah explains his answers to the questions on the self-assessment form: "I learned that Buffalo Bill was a rider for the Pony Express," and he discusses the resources he used. As she listens, Suzanne scans the room. Several students peck away at their computer keyboards. John is meeting with the resource teacher; Sam is meeting with the speech pathologist. Sarah and Brittany are practicing their presentations on J. Sterling Morton, the founder of Arbor Day, for Amy and Beth. Jordan and Liz are working on their poster boards on Fort Hartsuff. But Joe looks lost—again. He's sitting at his desk, staring into space. Suzanne makes a mental note to go over and see how he's doing.

Micah finishes explaining how he would do his research differently next time. He says that he will write short sentences on an index card to help him remember what to say during his presentation. He adds that he hopes his next project is even longer than this several-page report. Suzanne smiles at this student whose first report, at the beginning of the year, consisted of a single sentence.

She also thinks of Katie, who, for a full quarter, answered every math problem the same way: 7. When asked why, Katie responded, "Because last year it was 3. This year it's 7." This continued until the day Suzanne showed Katie her dismal scores on a visual rubric—*1*, *1*, *1*, *1*—and Katie broke down, sobbing. Seeing her performance somehow made it real. Soon, she began connecting. She didn't like the look of those *1*s, and she set out to change them. Now, several years later, she's in high school, doing just fine.

Proud, beaming Micah will do just fine, too.

We would wager that at least some features of this scene from Suzanne's 4th grade classroom are familiar to most teachers. The frenetic activity. The joy of watching students gain independence and find their own direction and rhythm. Teachers' team efforts to ensure that these diverse learners can access the resources they need. The heart-warming displays of competence and confidence. The struggling, as-yet-unreachable student.

But for all its familiarity, something unusual is happening here: Although most of the students don't know it, they're completing assessments for state reporting.

Suzanne, one of the coauthors of this article, teaches in Heartland Community Schools, a small district in eastern Nebraska. Many of her classroom assessments are part of her district's assessment process, which, in turn, is part of Nebraska's unique statewide system of local assessments: the School-based, Teacher-led Assessment and Reporting System (STARS).

Instead of opting for standardized state testing, as other states have done, Nebraska chose the assessment road less traveled to preserve the kind of teaching and learning we see in Suzanne's classroom. These 4th graders need not stop their wonderful, furious, stimulating activity to take a sterile pencil-and-paper test. Suzanne need not administer a standardized test that she had no hand in designing and that is not tied to her curriculum.

And that, as the poet writes, has made all the difference.

Preserving the Good Stuff

For several years, Suzanne and her colleagues have been using student-involved visual rubrics. Students help design these assessment guides, which include both pictures and words to help

Less Traveled

guide and assess research projects. It all began when Suzanne and a fellow teacher conducted an action research study (Ratzlaff & Diercks, 1995) in which they found that students who participated in developing and using visual rubrics for their own research projects achieved higher scores than those who were given traditional, teacher-designed rubrics or no rubrics at all:

> Students reported that working together to create the rubrics with pictures helped them better understand the research process. . . . Students expressed confidence in being able to conduct research on any topic. Teachers observed students asking for more conferences to assess their progress during the projects. Students showed more enthusiasm for the projects and often did extra research on weekends. (p. 16)

These rubrics have been a fixture in Suzanne's school ever since.

Using the rubrics in classrooms improves student performance and increases student motivation, but

Nebraska has become a kind of laboratory for classroom assessment.

continually *revising* the rubrics—on the basis of conversations with both colleagues and students—has proved equally important. As collective articulations of shared goals, the rubrics invite teachers into searching discussions about what they value in student work. Both within the school and among teachers in neighboring districts, the conversations spawned by these rubrics have become an important vehicle for professional development.

Students also participate in revising the rubrics—not only their content, but also their form. For example, students proposed adding a column to indicate that a student has gone "above and beyond" the rubric's expectations. They insisted that this column be left blank—so the teacher could fill it in with specifics—because there are unlimited ways to go above and beyond. Similarly, the earliest rubrics that Suzanne and her colleagues designed had the lowest score point on the left and the highest on the right (as most rubrics do). But her students asked, "Why do you have us read the worst first? Since we read from left to right, why don't you have the highest score on the left so we can read what is best right away?" Since then, the student-involved rubrics have had the highest score point on the left.

Through their own experiences as well as research on student self-assessment, visual literacy, and brain research, Suzanne and her colleagues have become convinced of the educational value of student-involved visual rubrics. But we want to emphasize that the rubrics aren't the "good stuff" that teachers smuggle into their classrooms when they're not furiously prepping for or administering standardized state tests. On the contrary—in Nebraska, classroom assessments "count." Their primary purpose is to inform teaching and learning. However, the data they provide may also be used for accountability purposes.

In fact, Heartland teachers had chosen their own road less traveled before the advent of STARS. When Nebraska's new system was implemented in 2000, Suzanne and her colleagues were required to document the assessment quality of the rubrics they used for state reporting. (This involved double scoring by qualified adults, for instance.) But they were *not*

built a system that challenges each school to improve every year.

But Does It Work?

Nebraska's assessment system has led to significant improvements in student performance at all grade levels. For example, at the 4th grade level, the percentage of Nebraska school districts earning the top rating of exemplary in student performance in reading rose from 31.8 percent in 2000–01 to 66.1 percent in 2004–05; at the 8th grade level, it rose from 34.1 percent to 59.2 percent; and at the 11th grade level, it rose from 18.3 percent to 41.3 percent.[2]

Data also show that Nebraska's district assessment quality—on which classifications of student proficiency are based—is steadily improving, according to the trained assessment experts who rate district portfolios. For example, the percentage of districts rated exemplary on their assessment quality in math rose from 30.2 percent in 2001–02 to 68.8 percent in 2004–05 at the 4th grade level, from 32.3 percent to 80.5 percent at the 8th grade level during that same period, and from 33.3 percent to 84.1 percent at the 11th grade level.

Nebraska students' scores on the statewide writing assessment—which is the only state test—have also risen sharply. In addition, students' traditionally high performance on nationally normed tests has remained stable, even though the state's changing demographics over the past decade might be expected to cause slippage in this performance, as has been the case in other states experiencing similar changes. Nebraska continues to struggle to close achievement gaps, but most demographic groups are showing improved performance, graduation rates are up, and dropout rates are down in the state as a whole, and especially in its urban districts.

The Comprehensive Evaluation Project (CEP), which evaluated STARS across 100-plus schools from 2001 to 2004, has identified a number of important shifts in the culture of Nebraska schools since the inception of the program (see Gallagher, 2007).

A Focus on Student Learning

Nebraska teachers describe moving away from a teaching-for-coverage model, in which they focused on getting through material in time for tests, toward a learning-for-understanding model, in which they focus on individual student mastery. One teacher described the role of classroom assessment this way:

> Instead of focusing on just what your objectives are, [classroom assessment] focuses on what your objectives are for each student. In essence, then, each child almost has an individual education plan.

Several of the schools in the evaluation study, for example, use student portfolios that students take with them across their classes and grades. One public school on an American Indian reservation has seen significant increases in student achievement, as well as in attendance and graduation rates, since it implemented its portfolio system, which culminates in a public presentation to teachers, parents, and other community members, including tribal elders. These portfolios are often color coded so teachers and students can see at a glance which standards students have and have not met. This information helps teachers adjust their instruction and frees them from moving through content or skills in lockstep. Many schools have moved away from textbooks and off-the-shelf curriculum and assessments, relying instead on teacher-designed materials, which teachers continually adjust and update to support individual learners.

Assessment Literacy

To design reliable and valid assessments, teachers need to be assessment literate. For a start, they need to know various assessment methods and their purposes, the rudiments of technical assessment quality, how to embed assessments in instruction and curriculum, how to interpret assessment data, and how to use data to adjust curriculum and instruction. Nebraska is developing this expertise within its teaching corps through courses, workshops sponsored by the Nebraska Department of Education and Educational Service Units, and district-level data retreats.

Although most of the students don't know it, they're completing assessments for state reporting.

However, the most prevalent and meaningful forums for teachers to design, review, and revise assessments are school based. These include teacher learning teams, action research projects, inquiry groups, and the like. Many districts are partnering with other districts, or a consortium of districts, to build or borrow expertise; but as one superintendent of a small, rural district put it, "We've learned that the consortium can help us with the technical aspects of [assessment] but that we need to really write assessments that fit with our curriculum."

Teachers are finding that the process of creating valid and reliable assessments is becoming easier. One noted,

required to put these assessments aside to make room for state tests. Indeed, STARS encouraged the teachers to keep experimenting with and improving the homegrown assessments that had become so important to teaching and learning in this district.

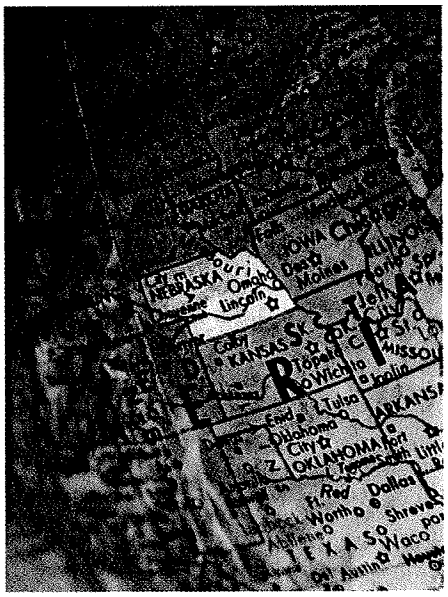
How STARS Works

The Heartland Community School district is not alone in building its local assessment system on high-quality classroom assessments; indeed, Nebraska has become a kind of laboratory for classroom assessment. Under the School-based, Teacher-led Assessment and Reporting System, assessment is in the hands of educators, not policymakers or test makers.

This decision was made in 1999, when Nebraska became the 49th state to adopt an assessment and accountability system. At that time, most states had implemented high-stakes standardized tests, and many of the pitfalls of such regimes were already in evidence, including narrowing of curriculums, deprofessionalization of teachers, emphasis on rote memorization rather than higher-order skills, misuse and misreporting of data, cheating scandals, and so on (see Gallagher, 2007; Nichols & Berliner, 2007). So Nebraska set out on the road less traveled, designing a statewide system of local assessments.

STARS is complex in detail but simple in its philosophy: Change can be forced on schools, but meaningful improvement must come from within—through commitment and capacity building, not compliance and control (see Barth, 2001; Darling-Hammond, 1997; Elmore, 2004). The most important decisions about teaching and learning happen in classrooms, not in conference rooms or boardrooms. Those decisions should be informed by the trained professional judgment of educators. Also, the primary purpose of assessment is to inform teaching and learning. So assessment must be embedded in instruction and curriculum, not imposed from outside. If assessment is valid and reliable, then the data it generates may also be used for accountability purposes.

Here's how STARS works. Each district in Nebraska is responsible for developing its own assessment



process—although many do so in collaboration with other districts or with regional education service units. The Nebraska Department of Education provides assistance for developing K–12 and cross-curricular district systems. However, annual *reporting* to the state on reading, math, science, and social studies is selective; it represents only one piece of a comprehensive, locally meaningful process. Most reporting occurs at the "guidepost" grades of 4, 8, and 11, although here, too, districts have some discretion.[1]

Nebraska asks all its districts to meet the same learning standards, but *how* districts do so varies widely. Some develop districtwide criterion-referenced assessments. Others, like Heartland, rely heavily on classroom assessments. Most devise a combination of the two.

To ensure that local assessments are of high quality, districts refer to a set of guidelines for assessment review. Local panels of educators and content experts review assessments used for reporting on standards on the basis of six quality criteria:

- Assessments align to state or local standards.
- Students have had the opportunity to learn the content on which they are being assessed.
- Assessments are free from bias or offensive language.
- The level is developmentally appropriate for students.
- Scoring is consistent.
- The mastery levels are appropriate to subject and grade levels.

Districts document the process they use to ensure that assessments meet these criteria, and then local and national assessment experts review the process, including assessment quality documentation, during an annual on-site review by the Nebraska Department of Education. Districts receive both written and oral feedback and two public ratings: one for student performance and one for assessment quality.

Data on student performance and assessment quality are combined with demographic and other key information about each school in the state and compiled in an annual *State of the Schools Report*. This report card provides a rich portrait of schools, districts, and the state, encouraging readers to understand student performance in context. Because STARS comprises local processes, it does not generate or report data that enable rank-ordering of schools. Instead of building a system that puts schools in competition with one another—the school-reform-as-beauty-contest model—Nebraska has

> We've gotten better at looking at what needs to be changed. I've learned about good questioning techniques, looking for biases, and just knowing what a good assessment looks like.

Nebraska is witnessing the birth of "a new breed of assessment literate educators" (Lukin, Bandalos, Eckhout, & Mickelson, 2004, p. 31).

Improved Professional Development

The team-based approach to assessment has become the predominant form of professional development in Nebraska schools. Instead of sending teachers off to a conference or bringing in an expert for inservice training, most schools have adopted a "teachers teaching teachers" approach. In one suburban district with 700 teachers, all instructional staff members sit on two learning teams: one organized by grade level and the other by content area. In both, they share assessments and samples of student work. Upward of 90 percent of teachers in the district responded favorably to these meetings on a survey. Indeed, the evaluation study reveals that teachers generally find this kind of professional development more relevant and immediately useful to them. One teacher noted,

> I feel as if my voice matters, that what is best for students is being brought up in [learning team] meetings. I am able to discuss my opinions with other teachers, and we can bounce ideas off one another.

Data-Informed Improvement

One important task of these teams is to interpret assessment data and decide how to respond. When STARS was first implemented, many schools went from having little meaningful data to having an overwhelming amount. Our researchers heard many educators say that they'd only just begun to understand how to read data, that data scared them because they didn't really know how to make the most of the information, and that probably everyone—from the physical education teacher right through to the high school principal and the superintendent—needed better training in how to use data to make sound education decisions.

Over time, with the kind of professional development we have described, teachers have become smarter about collecting, interpreting, and using data. These data then feed school improvement, including ongoing revision of curriculum, the professional development program, public engagement, and so on. But the primary purpose of data is to inform instruction, as one assessment coordinator suggested:

> If we're good assessors, if we really know where students are and what their knowledge and skills are, then we can also identify the gaps that they have and, if they are not meeting the assessment criteria, provide additional instruction, reassess, and hopefully, we won't have . . . students falling through the cracks. That, to me, is the real goal of assessment, the real goal of the whole STARS process.

However, we would be remiss if we failed to note that Nebraska educators generally resist the term *data driven*. They insist that data might provide a map, but that teachers must decide where to drive.

Teacher Leadership

As the driver metaphor suggests, *teachers* lead the school improvement charge in Nebraska schools. The relationship between administrators and teachers has undergone a transformation as teachers have taken the school-improvement reins. Many principals now see their role as that of a supporter. One principal referred to this new model of shared leadership in this way: "I make sure teachers have the things they need to be successful and to achieve the goals they're setting for the school and district." Nebraska teacher leaders may be school or district assessment coordinators, but most exert their leadership in less formal ways, such as by convincing colleagues to try student-led parent conferences, serving on a school improvement task force, taking a turn facilitating a learning team, or just letting their voice be heard on important teaching and learning matters in the school.

The promise and hope of Nebraska's unique system is that all schools will find their own roads less traveled.

The Better Road

There are bumps, of course, along Nebraska's road less traveled. Although the trends are overwhelmingly positive, researchers have uncovered causes for concern, including lack of time, a need for professional recognition for teachers, and more flexible approaches to determining the quality of classroom assessments. And a handful of districts continue to struggle along well-worn paths to nowhere in particular—or have lost their way entirely.

But for many districts, STARS provides an opportunity and encouragement to move forward along their chosen paths. Still other districts are using STARS to chart whole new courses. The promise and hope of Nebraska's unique system is that all schools will find their own roads less

traveled. And that will make all the difference. EL

[1]For information about reporting, correlations among different measures of student performance, and the annual reports of the Comprehensive Evaluation Project and the *State of the Schools Report*, go to www.nde.state.ne.us.

[2]For complete information about student performance and assessment quality ratings, see http://reportcard.nde.state.ne.us/Page/AccountabilityStateSummary.aspx?Level=st.

References

Barth, J. (2001). *Learning by heart.* San Francisco: Jossey-Bass.

Darling-Hammond, L. (1997). *The right to learn.* San Francisco: Jossey-Bass.

Elmore, R. F. (2004). *School reform from the inside out: Policy, practice, and performance.* Cambridge, MA: Harvard University Press.

Gallagher, C. W. (2007). *Reclaiming assessment: A better alternative to the accountability agenda.* Portsmouth, NH: Heinemann.

Lukin, L. E., Bandalos, D. L., Eckhout, T. J., & Mickelson, K. (2004, June). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice, 23,* 26–32.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools.* Cambridge, MA: Harvard University Press.

Ratzlaff, S., & Diercks, R. (1995). Student-teacher created visual rubrics: Models to guide and assess research projects. In *Assessment in action: Collaborative action research focused on mathematics and science assessments* (pp. 16–18). Denver, CO: McRel.

Chris W. Gallagher is Associate Professor of English at the University of Nebraska–Lincoln and author of *Reclaiming Assessment: A Better Alternative to the Accountability Agenda* (Heinemann, 2007); cgallagher2@unlnotes.unl.edu. **Suzanne Ratzlaff** is a 4th grade teacher in Heartland Community Schools in Henderson, Nebraska, and 2003 Nebraska Teacher of the Year; sratzlaff@esu6.org.

SCHOLASTIC SRI READING INVENTORY™

The Best Choice to Manage Reading Progress

This research-based, computer-adaptive assessment quickly and accurately measures reading comprehension on the Lexile Framework® for Reading.

SRI provides the critical components to help raise reading achievement by:

- monitoring student reading progress
- differentiating instruction
- generating meaningful reports
- matching students to appropriate texts

LEXILE®

1.800.387.1437
Scholastic.com/SRI

PreschoolFirst

Why do our users prefer PreschoolFirst?

- Combines intentional teaching with ongoing assessment
- Strengthens home-school connections
- Links to state-specific Pre-K, Head Start standards
- Assists developmentally delayed children
- Formative assessment creates progress reports

FREE Web Tour and Trial! Call 1-866-584-2900 or visit www.preschoolfirst.com

From NiTV a non-profit learning and technologies corporation.