You’re a principal or a school district administrator, and you want more detailed achievement information for individual students, in particular subjects, and in particular classrooms. You want evaluation methods that mesh with the No Child Left Behind Act (NCLB) indicators and performance targets. Detailed information is becoming more available through the efforts of a unique partnership. For the past 7 years, WCER staff has worked with Milwaukee Public Schools (MPS) district staff to develop MPS capacity to analyze and use data on students and schools. That partnership has developed a sophisticated system for measuring and tracking the productivity of MPS schools, producing data that forms the core of the district’s school report card and accountability system. (NCLB requires school districts that receive federal funds to provide a report card on how its schools and the school district are doing. The report includes the combined test scores of the students at all the district’s schools.)

During the first 3 years of the project (1998–2001), MPS lacked the necessary data to launch a comprehensive value-added system. Meyer and colleagues helped build capacity within the district to understand options for using data to measure school performance and using value-added performance data as part of a district accountability system.

Too often, administrators must rely on weak and limited student achievement data. Limited data results in inaccurate estimates of school and program productivity. But value-added measures paint a more detailed and more accurate picture (see sidebar).

Beyond NCLB

Some educators express concern that NCLB accountability provisions are unfair to schools. The NCLB law judges schools primarily on the percentage of children who perform at the "proficient" level on state tests. Schools don’t get credit for students who show substantial achievement in a given year but still fail to reach the proficiency bar, or for advanced students who continue to progress. Value-added measures enable states and districts to set and monitor explicit performance objectives for schools with low-scoring students, students of color, transient students, and other policy-significant groups. That’s important because school characteristics such as racial mix and poverty level can influence the rate at which children learn. Unless the analyses account for such differences, they will not accurately isolate the contributions of teachers or schools to student learning.

(continued on next page...)
Evaluation, alignment, hypertext, and realistic mathematics

For the past 7 years WCER staff have worked with Milwaukee Public Schools district staff to develop the capacity to analyze and use data on students and schools. That partnership has developed a sophisticated system for measuring and tracking the productivity of MPS schools, producing data that forms the core of the district’s school report card and accountability system. Value-added measures enable states and districts to set and monitor explicit performance objectives for schools with low-scoring students, students of color, transient students, and other policy-significant groups.

Special education leaders and policymakers now have a tool for gathering evidence of the validity of their states’ assessments. A nationally recognized alignment procedure developed at WCER has been applied to Wisconsin’s alternate assessment. The procedure is described in this issue.

Hypertext learning environments allow students to view the same material from multiple perspectives. As part of her research into optimal learning environments, Sadhana Puntambekar has developed a learning environment called the Concept Mapped Project-based Activity Scaffolding System. The CoMPASS hypertext system and curriculum modules help students learn science in the context of design-and-build challenges.

A Realistic Mathematics Education (RME) conference held at UW-Madison in November drew 120 educators from Italy, England, Ontario, and the U.S. The first RME conference held in the U.S. showcased RME theory and its applications. Starting with context-linked activities, students in Realistic Math classrooms develop mathematical tools and models and receive support in classroom interaction, which lead to higher levels of mathematical thinking and understanding.

For example, Milwaukee has a highly mobile student population, in part because of the district’s school choice program. When students move into a district or state late in the year, after regular standardized assessments are administered, a well-designed value-added system assesses these students at the point of entry and incorporates them into an appropriately generalized value-added model. Historically, these students have not been included in standardized assessments.

Obtaining better evidence of students’ academic growth is needed not only out of fairness, but also because this evidence provides a more accurate picture of school effectiveness when combined with measures of absolute achievement.

Equity issues

Value-added analysis makes it possible to separate out individual student characteristics, including gender, previous test results, number of absences, eligibility for free school meals, length of time in the school, and special education status. It also accounts for school-level characteristics such as the percentage of students receiving subsidized meals and the number of children enrolled in the tested grade.

Conventional assessments don’t measure the degree to which teachers and schools reduce achievement gaps among different student groups (high and low achievers, poor and non-poor, etc.). Equity-oriented value-added indicators, however, make it possible to determine whether the productivity of schools and teachers differs for students with different characteristics. For example, a given school might be very effective with talented and gifted students but less so with students with low prior achievement.

Data-driven improvement

Meyer’s research represents a comprehensive approach to harnessing information on educational resources and student achievement, or, to put it another way, system inputs and outputs.

“We want to help create a learning environment characterized by continuous, data-driven improvement,” Meyer says. “Value-added indicators aim to find out ‘what works.’ In particular, we want to pinpoint the determinants of high teacher productivity, effective professional development, and alternative instructional strategies.”

The MPS value-added system has until now measured the performance of schools at the grade level, but not at the teacher or classroom level. Improved value-added methods promise to make available a much-needed quantitative component for teacher evaluation, to be combined with other sources of information, such as classroom observations.

Meyer credits Milwaukee Superintendent William Andrekopoulos and Director of Assessment and Accountability Deborah Lindsey for their leadership in this work. MPS has committed its own resources to support this research, rather than relying on external funding agencies.

Specifically, researchers and district staff are enhancing the current MPS value-added system to enable it to:

- Measure teacher and classroom performance, as well as school productivity;
- Measure how well teachers and schools reduce achievement gaps among different population groups;
• Incorporate diagnostic tools that help schools, teachers, and district administrators respond to accountability results with concrete steps to improve performance;

• Yield data on long- as well as short-term teacher and school productivity, removing any incentive for schools to teach narrowly to test outcomes and to emphasize short-term gains; and

• Develop and apply value-added statistical tools that explicitly control for and eliminate student selection bias.

The problem of selection bias is likely to be especially acute in districts like MPS that provide substantial opportunities for school choice. For example, students with persistently high or low achievement may systematically choose (or be assigned to) different schools and teachers. Test-taking rates may also differ across students due to student choices or to school policies. Statistical tools that adjust for this bias yield more accurate and useful information.

A new center

The Value-Added Research Center (VARC) was launched this winter at WCER as an umbrella organization for Meyer’s many theoretical and applied research projects. VARC research is comprehensive in that it encompasses policy-relevant expansion of the current MPS value-added system, dissemination of value-added information in a clear and accessible fashion, and policy implementation and professional development.

A hallmark of Meyer’s research is what he calls full-spectrum inquiry. His team conducts theoretical research, statistical research, applied qualitative and quantitative research, program evaluation, program design, and program implementation—with the goal of eventually helping other schools, districts, and states enhance their evaluation programs and professional development.

VARC extends Meyer’s prior work by helping MPS use student transcript data to document various paths Milwaukee students take to proficiency across the district and by evaluating the performance and effectiveness of schools, teachers, programs, and policies over the long term. The project has grown into a network that now includes the states Michigan, Minnesota, Wisconsin, MPS, and the Minneapolis Public Schools. More states are expected to join in as time goes on.

A generalized approach

Meyer’s generalized value-added model is useful in several ways. First, it can produce valid estimates of school performance free of student selection bias. Second, it can be used as a standard to evaluate simpler models. (Most, if not all, existing value-added models are special cases of this generalized model.) The generalized value-added model also can be implemented in situations in which at least some students change schools from one grade to the next.

In this work, qualitative and quantitative research inform each other. “Our qualitative research generates hypotheses about what works. Then quantitative research (surveys and other resources) takes the qualitative work to scale,” Meyer says. “Next, we use a validity study to ‘marry’ the results of qualitative work with survey results. Finally, we analyze large-scale student outcomes data.”

Along the way, the researchers and MPS staff develop hypotheses about student achievement in the context of their schools.

“Through research we find out which hypotheses matter, which innovative ideas are successful. The final task will be building diagnostic tools to help schools, teachers, and district administrators respond to accountability results with concrete steps to improve performance.”

Funding

VARC is funded through the Longitudinal Data Systems grants recently awarded by the U.S. Department of Education’s Institute of Education Sciences. VARC principal investigators led a Tri-State Partnership (Minnesota, Michigan, and Wisconsin) that won 3 out of 14 grants. VARC will provide overall project coordination, data warehouse design assistance, and support for researchers who are embedded in the districts.

**What is value added?**

The term value added can mean many things, but fundamentally it refers to the contributions teachers and schools make to student achievement.

**Five features** make Meyer’s project unique:

1. It explicitly addresses student-level selection bias and provides solutions for eliminating it.
2. It allows for possible decay over time in the effects of interventions.
3. It tracks student achievement across time.
4. It allows for the likelihood that school effects are not random, but may be correlated with student-level variables such as race, ethnicity, and income status. Such a correlation arises, for example, if poor students are disproportionately served by low-performing schools, an issue of policy interest.
5. It explicitly addresses measurement error in student achievement.

“We want to know why super-successful schools work, and why underperforming schools don’t,” he continues. “What is the ‘unknown’ in those schools?”
Aligning Alternate Assessments

Some students have disabilities that make their participation in state- and district-wide tests impractical. In many cases, students with disabilities who participate in alternate assessments receive curriculum and instruction that differ significantly from those received by other students. The mainstream tests don’t accurately measure their academic achievement. It’s important to know how well alternate assessments align with the curriculum and instruction provided to students with disabilities.

Special education leaders and policymakers have a tool for gathering evidence of the validity of their states’ assessments. A nationally recognized alignment procedure developed by WCER Senior Scientist Norman Webb, this tool has been applied to Wisconsin’s alternate assessment.

Norman Webb and colleagues Andrew Roach and Stephen Elliott analyzed the degree of alignment between Wisconsin’s Model Academic Standards and the Wisconsin Alternate Assessment (WAA), which assesses the academic performance of students with significant disabilities and offers an alternative to the traditional achievement tests. Specifically, the study asked: Does the WAA adequately measure the concepts and skill areas represented in Wisconsin’s Model Academic Standards? The study found content-related evidence for the WAA’s validity.

Measuring alignment

Webb says effective schooling depends on coordinating three components of the educational environment: curriculum, instruction, and assessment. The degree to which these elements work together toward student learning is termed alignment, and that provides the foundation of standards-based education reform.

Roach convened a panel to code and rate the alignment between the WAA items and Wisconsin’s Model Academic Standards. The primary role of the panel members was to:

1. reach consensus on a depth-of-knowledge level rating (see below) for each objective in the Model Academic Standards;
2. rate the depth-of-knowledge level for each item on the WAA rating scale; and
3. identify the one or two objectives from the Model Academic Standards corresponding to each WAA item.

Four levels were used to rate depth of knowledge:

- **Level 1 (recall)** requires simple recall of such information as a fact, definition, term, or simple procedure.
- **Level 2 (skill/concept)** involves some mental skills, concepts, or processing beyond a habitual response; students must make some decisions about how to approach a problem or activity. Keywords distinguishing a Level 2 item include *classify, organize, estimate, collect data, and compare data*.
- **Level 3 (strategic thinking)** requires reasoning, planning, using evidence, and thinking at a higher level.
- **Level 4 (extended thinking)** requires complex reasoning, planning, developing, and thinking, most likely over an extended time. Cognitive demands are high, and students are required to make connections both within and among subject domains.

Why alignment is important

Strengthening the correspondence between students’ individualized education programs (IEPs) and the state’s alternate assessment and academic standards helps ensure that students with significant disabilities are included in a meaningful way in instructional improvement efforts and standards-based reform.
The Promises of Realistic Math Education

Students in mathematics classrooms should not be considered passive recipients of ready-made math. Instead, students should be guided toward using opportunities to reinvent mathematics by doing it themselves.

That’s one of the principles underlying Hans Freudenthal’s concept of mathematics as a human activity. Students start with a context-linked activity. For example, they could be asked to make drawings to show different ways that a given number of candies, packed in rolls of ten, might be arranged in a storeroom. For example, 143 candies might be packed up into one ten-roll box and four rolls, with three single pieces, or they might be stored as twelve rolls and twenty-three pieces.

Then students gradually develop mathematical tools and understanding at a more formal level. Models emerge from the students’ activities, are supported by classroom interaction, and lead to higher levels of mathematical thinking.

The Freudenthal Institute, founded by Hans Freudenthal in 1971 at Utrecht University in the Netherlands, and WCER have collaborated on research and development over the past 15 years. This relationship has resulted in the establishment of an international research institute for mathematics education: Freudenthal Institute–USA. Its mission is to establish a comprehensive network of collaborating researchers and developers. They call their approach to teaching mathematics Realistic Mathematics Education (RME).

Freudenthal Director Jan de Lange says:

The desired competencies, not the mathematical content, are the main criteria for mathematics literacy. And these are different at different ages and for different populations. . . . We all need to understand how important, how essential, mathematics literacy is for every student, and mathematicians in particular need to understand that mathematics literacy will contribute to a better perception about what constitutes mathematics and how important that field is to our lives.

A Realistic Mathematics Education conference held at UW-Madison in November—the first such conference in the US—drew 120 people, including educators from Italy, England, and Ontario. Sponsored by the Freudenthal Institute–USA (http://www.fi.uu.nl/en/fius/welcome.html), the Mathematics in Context Satellite of the Show-Me Center Project (http://showmecenter.missouri.edu/), and WCER, the conference showcased RME theory and its applications in classrooms.

Keynote speakers included Gail Burrill, Michigan State University; Eric Gutstein, University of Illinois at Chicago; Cathy Fosnot, City College of New York; and the Freudenthal Institute’s Jo Nelissen, Jan de Lange, and Marja van den Heuvel-Panhuizen. They discussed topics such as the transformation of teaching and learning, the dialectical relationship between teaching mathematics for social justice and teaching Mathematics in Context, ways of increasing students’ access to mathematics, and the role of context in assessment problems in mathematics.

“In a knowledge society everyone shares interest in improving education,” said de Lange. “We need to connect research in different disciplines—cognitive science, neuroscience, mathematics, and math education. Researchers don’t talk with or work with other disciplines enough.”

De Lange also encouraged educators to learn to maximize informal learning opportunities and after school learning using the Web.

One person representing the WCER team at the RME conference was Tom Romberg, whose research over the past 30 years has analyzed young children’s learning of initial mathematical concepts, methods of evaluating students and programs, and integrating research on teaching, curriculum, and student thinking.

“For research to be productive and useful in any discipline,” Romberg said, “it must be conducted within a research community. Mathematics education is a relatively young academic field, and research on the teaching and learning of mathematics is even younger. A research community in mathematics education has only gradually begun to emerge in the past half century, and the process is still ongoing.”

Romberg said much more is known today than was known in the 1960s, however. For example, it is now understood that learning mathematics with understanding involves more than being able to produce correct answers to routine problems. Mathematics should be viewed as a human activity that reflects the work of mathematicians—finding out why given techniques work, inventing new techniques, and justifying assertions. Learning with understanding occurs when students are given time to discover relationships and learn to use their knowledge, and when they reflect about their thinking and express their ideas. Doing mathematics cannot be viewed as a mechanical performance or an activity that solely involves following predetermined rules.

Romberg, who announced his retirement at the RME conference, was a leader in the development of the National Council of Teachers of Mathematics (NCTM) Curriculum and Evaluation Standards for School Mathematics (1989). He is former director of WCER’s National Center for Improving Student Learning and Achievement in Mathematics and Science and has become Emeritus Sears Roebuck Foundation–Bascom Professor of Education at the University of Wisconsin–Madison. Despite his retirement, Romberg remains active at WCER and is completing a revision of his Mathematics in Context curriculum.
Using Hypertext Navigation to Deepen Learning

Hypertext learning environments hold great promise, but they also present challenges for learners and designers. As part of her research into designing optimal learning environments, UW-Madison education professor Sadhana Puntambekar developed the project she calls CoMPASS (Concept Mapped Project-based Activity Scaffolding System).

CoMPASS consists of a hypertext system and curriculum modules based on the pedagogical framework of Learning By Design, which has students learn science in the context of design-and-build challenges. The CoMPASS system includes two tightly integrated representations of content, one textual and the other visual (in the form of concept maps). Each CoMPASS screen (see illustration) represents a concept such as mass or gravity and provides both a concept map (left half of the screen) and a textual description (right half of the screen).

Concept maps display a fish-eye view—that is, the focus concept appears at the center, with the most closely related concepts displayed in the first ring and the less closely related concepts in the outer ring. The maps show connections between science phenomena and give students alternative paths to pursue so they can see how different phenomena relate to each other.

In contrast to a hierarchical index, which orients the learner but does not show all the relations between concepts, concept maps present ideas in the form of nodes linked by a word. In its simplest form, a concept map is just two words connected by a linking word to form a semantic unit (for example, friction and force connected by the link type of). Concept maps help students visualize the interconnections between concepts.

Acquiring advanced knowledge requires students to revisit the same material at different times, in different contexts, for different purposes, and from different perspectives. The alternative views in CoMPASS make this possible, helping students study science concepts and phenomena in depth by encountering them in multiple contexts. For example, a student might want to learn about the phenomenon of force in the context of objects falling in air. In CoMPASS, the student could change views to study the same concept (force) in other contexts, such as linear motion or simple machines.

To take another example, in one physics unit parachutes serve as an example of falling objects. If a student selects the concept motion of falling objects, CoMPASS provides an explanation of that phenomenon, along with a map of related concepts and principles. The student may then choose any concept from this map, and CoMPASS creates the corresponding fish-eye view. The maps reflect the semantic relatedness of the concepts within a topic and show students what related topics they can go to.

Evaluating CoMPASS’ effectiveness

Puntambekar recently evaluated the effectiveness of CoMPASS by studying two classes of eighth graders who used the system to learn about force and motion. One class, the maps class, used CoMPASS; the second, the index class, used an index version of the system that presented list of concepts instead of navigation maps. Both classes used the same text, teacher, and curriculum.

In the study, students were asked to design roller coasters. As the students raised questions about how roller coasters work, they encountered the concepts of force and motion. After finishing their initial designs, students built models, exploring concepts about the physics involved. During each class, students used CoMPASS to answer their questions.

Puntambekar found that, compared to students in the index class, students in the maps class:

- visited more goal-related concepts and spent more time on them;
- improved more on an essay question; and
- performed better when tested on their depth of knowledge.

The two classes did not differ significantly in their knowledge of facts as demonstrated on the pretest and posttest.
Roach’s panel found that the WAA generally met the criteria for acceptable alignment between assessments and curricular expectations as articulated in academic content standards. By establishing the alignment and curricular relevance of the WAA, Webb’s study provided evidence of the validity of the WAA results as a measure of students’ mastery of the academic concepts and skills outlined in the standards. The investigation also demonstrated the use of a formal procedure—the convened panel—to establish the alignment of an alternate assessment.

To improve the validity and utility of WAA results, the panel did recommend adding items to the WAA Science scale to improve its alignment with the state’s academic standards and to the IEPs and classroom curricula of students with significant disabilities. Three new items were added to the WAA Science scale before the initial WAA implementation year.

This study speaks to the role of sequential development and expert review in promoting the alignment between the policy elements of curriculum, instruction, and assessment systems. Sequential development involves creating and accepting one policy element and then using it as a blueprint for additional policy elements.

But students in the maps class improved more on the essay question than did the students in the index class. In fact, students in the index class performed worse on the essay question in the posttest than in the pretest. Puntambekar speculates that the index presentation, which was linear, may have been detrimental to learning.

The study suggests that the type of navigational aid may have affected students’ navigational decisions and thereby their learning. The index class largely followed the structural aids provided in the index, no matter what the goal. Although the goals were increasingly open-ended over the course of the 3 days, the structure of the index consistently drove the navigation. In contrast, the navigation used by the maps class clearly reflected the particular goal on each of the 3 days.

For more information visit the CoMPASS web site: http://www.compassproject.net/info/ Puntambekar’s research is supported by an NSF CAREER grant.


**ALIGNING ALTERNATE ASSESSMENTS**

(...continued from page 4)

For more information about alignment, see http://facstaff.wcer.wisc.edu/normw/MIAMI%20FLORIDA%20FINAL%20slides%2011-15-05.pdf.

To try the Web-based alignment tool, see http://www.wcer.wisc.edu/WAT/index.aspx.

For more information about the technical aspects of the evaluation study, see “Issues Related to Judging the Alignment of Curriculum Standards and Assessments,” http://facstaff.wcer.wisc.edu/normw/AERA%202005/2005papers/Alignment%20paper%202004-25-05%20edited%20nlw.doc.
It's time to refresh our mailing list. If you'd like to continue to receive WCER Research Highlights, simply drop this card in the mail with your name or updated information. Alternatively, send an email to the editor, pbaker@wisc.edu. If we don't hear from you, we'll assume you no longer wish to subscribe.

☐ My information is correct, please keep me on the mailing list.

☐ Please keep me on your mailing list but update my contact information.

NAME: _______________________________________________________________________

ADDRESS: __________________________________________________________________

CITY: __________ STATE: ______ ZIP: __________ EMAIL: ____________________________