Marshall Memo 342

A Weekly Round-up of Important Ideas and Research in K-12 Education
June 28, 2010

In This Issue:

Quotes of the Week

"The real harm is that low-performing students, fed a steady diet of test-prep worksheets, can sometimes follow familiar steps well enough to be declared proficient on the test but cannot answer parallel questions when asked in a slightly different format and cannot use their knowledge in applied contexts."

Lorrie Shepard (see item #2)

"Their knowledge isn't, 'Oh, there is a puddle on the sidewalk and when the sun comes out it evaporates and goes magically into a cloud and then it comes down again when it rains.' They understand how those molecules start to move and bounce around like mad as they evaporate. They really understand this in a way that they didn't before."

Boston third-grade teacher Candace Chick (see item #4)

"Many studies have shown that if a student enters class with a misconception, a good, clear explanation of the correct answer often fails to displace the student's incorrect idea. Research has found that we must allow our students a chance to apply their own ideas because only if these fail are students likely to replace their previous concepts with ones we have taught them."

Douglas Duncan (see item #3)

"I expected that you would teach me. I didn't expect that I would have to learn!"

A student commenting on the use of "clickers" in his classroom

"Instead of talking about what students can't do or aren't doing, we need to figure out how to convey belief about intelligence. We also need to figure out how to convey our belief in their capacity – even when they have given up on themselves. Then, and only then, can we engage them in learning how to 'get smart.'"

Laura Cooper (see item #5)

1. Thinking Through the Details of Merit Pay for Teachers

In this helpful *District Management Journal* article, Nicholas Morgan and Daniel Schiff provide a systematic analysis of performance pay for teachers. They begin with three theories of action on why performance pay might improve instruction and student achievement:

- Incentive to work harder – Clear recognition and rewards will motivate teachers to put in extra effort, improve classroom instruction, and thereby improve student achievement.
- Incentive to work smarter – Increased transparency of goals and outcomes will motivate teachers to get the best training, upgrade their professional skills, and use more effective classroom strategies, which will improve student achievement.
- A way to improve the faculty – Clear rewards for top performers will attract and retain more-effective teachers, create a culture in which less-effective teachers self-select out, thereby improving instruction and student achievement.

There have been some promising pilot programs, mostly abroad, but so far there's no definitive research on which of these theories is most valid. Nevertheless, merit pay is a popular idea these days, and most superintendents are enthusiastic about it, while worrying about teacher union resistance and whether their test results are accurate enough for high-stakes pay decisions.

Morgan and Schiff go on to pose the key questions that need to be answered to design a performance-pay system:

• *What's measured* – Should teachers be rewarded for classroom performance (judged by observations), for their students' achievement, or for a combination? The big advantage of including classroom observation is that teachers in non-tested grades (who constitute well over half the teaching force) can be included in the program.

• *What's rewarded* – If student achievement is the criterion, should teachers be rewarded for absolute test scores or value-added? There are simple-fairness advantages to the value-added approach – but serious technical issues in getting valid data in a timely fashion. Using test scores alone raises concerns about narrowing the curriculum and teaching to the test.

• *Who gets rewarded* – Should merit pay go to individual teachers, teacher teams, or the school staff? If the latter, should all campus personnel be eligible, only certified staff, or only core-subject teachers? Rewarding individuals can undermine teacher teamwork, but group

rewards can create a "freeloader" problem (team members who don't contribute but still get rewarded), which can demotivate top performers.

• *The scope of rewards* – Should merit pay go to all who meet the criteria, or only to a fixed number? The first option seems fairer, but it's financially risky since if many teachers do well, awards might exceed the budget. A fixed number of rewardees, on the other hand, could create unhealthy competition among teachers, undercutting collaboration and a common sense of mission.

• *The size of rewards* – A number of districts believe that paying teachers in the vicinity of $1,000 extra a year isn't that enticing. Houston came to the conclusion that teachers needed the chance to earn up to 20 percent above their base salaries for the program to make a difference, but most plans are nowhere near that generous. The problem is that substantial rewards might be unaffordable and/or unsustainable over time. To launch a performance pay program, districts need to tap into district funds, state grants and appropriations, federal grants, and philanthropic and corporate support, none of which are rock-solid over time.

Districts around the country have created programs that answer the above questions in a variety of ways:

- Performance-based (student achievement and/or evaluations): Houston, Columbus, Dallas, Lamesa, Florida;
- Knowledge- and skill-based (based on licensure, certifications, and professional development): Coventry, Manitowoc, LaCrescent, Cincinnati;
- Needs-based (rewards for serving in underperforming and hard-to-staff schools): Fairfax, Philadelphia, Massachusetts, California, and Utah;
- Performance and knowledge/skills-based: Minneapolis, Chicago, Douglas County;
- Performance and needs-based: Charlotte-Mecklenburg, Chattanooga, New York City, New York, Virginia;
- Four components (Student growth, professional evaluation, knowledge and skills, and market incentives): Denver
- Rewards to individual teachers: Minnesota, Hillsborough County;
- Rewards for schoolwide achievement: Texas, North Carolina, Toledo.

Morgan and Schiff say the jury is still out on which of these approaches is most effective at improving teaching and learning. Here are their recommendations for districts and schools thinking about setting up a performance-pay system:

- Think through each of the design questions above and come up with a good overall rationale and plan.
- Align the program design with the strategic objectives and local context.
- Simplify the program as much as possible to promote broad understanding.
- Use performance indicators that are objective and measurable.
- Create enough flexibility to respond to new challenges.
- Offer rewards large enough to motivate employees to excel.
- Ensure that the program has sufficient funding and is sustainable over time.

"Pay-for-Performance Programs: Strategies, Structures, and Funding" by Nicholas Morgan and Daniel Schiff in *The District Management Journal*, Spring 2010 (Vol. 4, #1, p. 29-36, 41-44), no e-link available; the authors are at nmorgan@dmcouncil.org and dschiff@dmcouncil.org

Back to page one

2. Concerns About Interim Assessments

In this *Peabody Journal of Education* article, University of Colorado/Boulder professor Lorrie Shepard summarizes eight articles on interim assessments in this issue of and says the two most glaring weaknesses are the benchmark tests themselves and the quality of teacher analysis and follow-up with students. Here are the details:

• "Of course, the idea of using data on an ongoing basis to find weak spots and make necessary improvements has intuitive appeal," says Shepard, "and we know it works for automobiles in Japan." But checking for student understanding, whether it's done on the spot or every 6-8 weeks, isn't enough to improve achievement, she says.

• So far, there has been very little research on interim assessments (the type given every 6-8 weeks, designed to mirror end-of-the-year state assessments), whereas there is a robust literature on formative (on-the-spot) assessments.

• Shepard casts a skeptical eye on interim assessments. She is particularly concerned about "teaching the test" – narrowing the curriculum to reading and math and focusing on items like those on state tests. This may improve test scores short-term, she says, but doesn't translate into long-term learning. "This is more than a problem of schools 'faking good' on accountability measures," says Shepard. "The real harm is that low-performing students, fed a steady diet of test-prep worksheets, can sometimes follow familiar steps well enough to be declared proficient on the test but cannot answer parallel questions when asked in a slightly different format and cannot use their knowledge in applied contexts."

• Shepard says that interim assessments, as they are being implemented in many schools, are likely to foster a "performance orientation" in students – working for grades to *appear* competent – versus a "learning" or "mastery" orientation, where the focus is on *being* competent, taking responsibility for one's own learning, and moving on and mastering more difficult material.

• Shepard believes that the typical commercial interim assessment feeds this syndrome – it's likely to "exacerbate the problems of teaching the test and test score inflation", has "negative effects on motivation", "limits learning", and is "inappropriate and fraudulent." "Overall," concludes Shepard, "the findings dispute the claims by vendors that data will immediately improve instruction and student achievement." She says we need a *Consumer Reports*-type of analysis of the interim assessments that test companies are producing.

• Three of the articles in this *Peabody Journal of Education* issue analyze the implementation of interim assessments in Philadelphia's public schools starting in 2004. The interim tests are criticized for containing only multiple-choice items and focusing mainly on lower-order thinking skills. The assessments aren't designed "to provide conceptual or diagnostic insights about what students are thinking," says Shepard.

• Another concern with Philadelphia's implementation of interim assessments (the schedule called for five weeks of instruction, tests, and a week of remediation) is how well teachers use the data to fix learning problems. Shepard notes that teachers' proficiency in this area varies dramatically from school to school, "with teachers in the lowest performing schools typically being less able to invent new strategies to compensate for what had not gone well in the preceding 5 weeks" – whereas schools that already have effective leadership, a commitment to accountability, and a problem-solving ethos use the data well. Budget cuts in Philadelphia eroded the professional development intended to level the playing field.

• There is a tendency, especially in the lower-performing schools, to focus on the "bubble kids" – those just below the proficiency cut-score – and use interim assessment results to predict how students will do on end-of-year state tests. This helps some schools eke out AYP gains but neglects the needs of many other students.

• One Philadelphia elementary school in the study made spectacular, multi-year gains in student achievement by focusing grade-level team meetings on thoughtful analysis of interim assessment data and systematic follow-up with students. But in most schools, teacher teams drew relatively superficial conclusions from the assessments and were less successful in boosting their students' achievement. Sadly, the successful school's achievement lapsed back to mediocrity when the principal left.

• The difference between high-gain and low-gain schools was that in the former, teachers analyzed not only *what* students got wrong on the interim assessments but *why* – gaining conceptual insights about where students' understanding broke down, about their skill and knowledge deficits and misconceptions. In less successful schools, teachers plodded through the assessment data item by item, focused on getting students to practice more, and often blamed students' errors on procedural goofs, poorly-worded questions, or the fact that the skill hadn't been covered yet. The researchers say that most teachers learned "almost nothing" about what math their students knew or understood.

• In the reteaching week, most Philadelphia teachers "attempted to reteach in the same way that they had taught before," says Shepard. If teachers did try a different approach, it "was not based on anything they had learned from the assessment."

• Very few teachers (only 2 of the 25 in one study) made the connection between interim assessment math results and the *Everyday Mathematics* curriculum they were using in their classrooms. In almost all cases, the assessment data weren't used to fix learning problems and tweak lessons in daily classroom instruction.

• Shepard concludes by looking at the three purported benefits of interim assessments and suggesting that there is a better source of information for each one:

- *Predicting student performance on end-of-year state tests* – Shepard says the previous year's state test scores are a better and cheaper predictor.
- *Evaluating instructional programs* – Similarly, end-of-year test results from the previous year do just fine.

- *Improving instruction for individual students or the entire class* – Shepard says that in-class, on-the-spot assessments do a better job of providing substantive insights into student thinking.

"What the Marketplace Has Brought Us: Item-by-Item Teaching with Little Instructional Insight" by Lorrie Shepard in *Peabody Journal of Education*, May 2010 (Vol. 85, #2, p. 246-257), no e-link available; Shepard can be reached at Lorrie.Shepard@Colorado.edu.
See Marshall Memo 324 #1 and Memo 293 #1 for more information on Philadelphia's implementation of interim assessments.

Back to page one

3. The Effective Use of "Clickers" in Classrooms

In this helpful booklet, University of Colorado professor Douglas Duncan presents a rationale and specific suggestions for using clickers (personal response devices) in K-12 and college classrooms. The basic problem that clickers address is passive learners. Duncan says clickers can greatly increase student engagement and boost learning – but only if they are used correctly. He agrees with Harvard physics professor Eric Mazur (see Marshall Memo 241 for a summary of Mazur's book, *Peer Instruction*) that it's essential to follow up well-framed clicker questions by asking students to "convince your neighbor." The peer interaction that ensues (before repolling the question and revealing the correct answer) instantly creates numerous teachers in the classroom and dramatically boosts students' understanding – both those who got the initial clicker question wrong and those who understood it up front. This real-time accountability is the key to clickers' success – and it's not always popular with students. One commented, "I expected that you would teach me. I didn't expect that I would have to learn!"

"It is still not entirely clear why peer instruction works so well," says Duncan. "Much research remains to be done in this area… After observing peer interactions for about 6 years, we suggest that one reason for their success is the strong peer pressure that college students feel. Certainly, most students behave as if the opinion of others in their group is more important than the instructor's. After all, we just give them their grades; but if their friends think they're dumb, that's much worse!" One student commented, "Professor X is an excellent teacher, but it is obviously better to learn from another student who thinks like me."

Duncan is particularly interested in the stubbornness of students' misconceptions. "Many studies have shown that if a student enters class with a misconception, a good, clear explanation of the correct answer often fails to displace the student's incorrect idea," he says. "Research has found that we must allow our students a chance to apply their own ideas because only if these fail are students likely to replace their previous concepts with ones we have taught them." Well-formulated clicker questions can surface misconceptions and get students to think them through, see their weaknesses, and change their minds.

Duncan closes with a number of specific suggestions on the use of clickers, which he thinks should be used on a daily basis:

• Use clickers to enhance interactive teaching, not as a tool to take attendance or give on-the-spot quizzes that count toward the final grade. "Simply having students vote on a problem without peer-to-peer discussion is not effective," says Duncan.

• Keep the stakes for clicker questions low – don't count them for more than 5 percent of the overall grade. "A higher percentage leads to anxiety among students, and they will be more focused on getting the answer right rather than on thinking critically," says Duncan.

• Avoid asking questions that require a calculation. Rather, ask questions that require critical thinking, conceptual understanding, and active learning.

• Keep the level of difficulty of clicker questions at an intermediate level – not so easy that the answer seems trivial, but not so hard that students will be lost and guess. Ideally, about half of students should get the correct answer, setting the scene for lively and productive "convince your neighbor" dialogues.

• Make sure that the final exam for the course tests conceptual understanding. "If the classroom focus is on conceptual understanding and the students are examined only on their problem-solving skills, then the student would see no benefit to classroom discussion or the importance of thinking through a conceptual problem," says Duncan.

Clickers in the Classroom: How to Enhance Science Teaching Using Classroom Response Systems by Douglas Duncan (Pearson Education, 2005)

Back to page one

4. In Science Teaching, Facts Alone Are Not Enough

In this *Harvard Education Letter* article, Patti Hartigan reports on The Inquiry Project, which aims to increase clarity and conceptual depth of science topics like weight, volume, density, and the water cycle in grades 3-5. Students in schools piloting the program wrestle with questions like, *Can two objects of the same size weigh different amounts?* and *Does a tiny piece of clay have weight?* Candace Chick, a teacher at the Mason School in Boston, who considered herself borderline science-phobic before the program, waxed enthusiastic about her students' progress: "Their knowledge isn't, 'Oh, there is a puddle on the sidewalk and when the sun comes out it evaporates and goes magically into a cloud and then it comes down again when it rains.' They understand how those molecules start to move and bounce around like mad as they evaporate. They really understand this in a way that they didn't before."

The Inquiry Project aims to change what its creators believe is the problem with the current science curriculum – disjointed facts, definitions, and topics taught in short units with little continuity from year to year, a mile wide and an inch deep. For example, the concept of matter is often introduced in first grade and then not mentioned again until middle school. The Inquiry Project wants to knit ideas together into progressions from kindergarten through high school and beyond. "If you are going to prepare students, you have to develop a network of concepts, not just topics, and devise instructional units that move ideas forward to prepare kids to take the next step," says University of Massachusetts/Boston psychologist Carol Smith.

This work is tied in with the American Academy for the Advancement of Science's *Atlas of Science Literacy* and a new draft of national science standards that will be released in July 2010. The progressions will build on how students' understanding evolves over time and how good teaching can help develop knowledge and skills. There are big gaps to be filled – for example, we know that primary-grade children can describe how the sun rises in the east and sets in the west, and upper-elementary students can make the connection to the rotation of the earth, but researchers struggle when it comes to helping middle-school students conceptualize the difference between the earth's rotation and orbit.

"A lot of assumptions being made about the nature of learning are wrong," says Gregg Solomon of the National Science Foundation. Little children are not blank slates when they enter school, he continues. They are already theorists with knowledge – and lots of misconceptions. "You have to take the students' perspective into account," says Amelia Gotwals of Michigan State University. "The previous standards have been based on what scientists say, but they don't identify the common ideas and misperceptions that students bring with them. How can we design curriculum and standards that take those things into account?"

One challenge is shifting students from the "force-dynamic" reasoning typical in elementary schools – looking at the world in terms of "actors" with purposes. For example, students typically see a tree as an actor whose purpose is to grow, and in order for it to grow, it gets water, air, sunlight, and soil. The more correct scientific explanation is that the tree uses the sun as an energy source to convert carbon dioxide and water into glucose, which is the building block of the tree. A tree's growth is really the story of the transformation of matter and energy. Getting students to look at the natural world this way helps them have a more realistic grasp of concepts like the carbon cycle and climate change.

In Inquiry Project fourth-grade classrooms, students watch time-lapse videos of plants growing, then grow plants under different conditions (with and without water and light), then measure the results over time and see that plants need water, air, and light to grow. Then they look through microscopes to examine the cellular makeup, which leads to an understanding of how plants make their own food and store it. Taking care of plants helps students see that plants aren't actors – they're part of an ecosystem.

"Learning Progressions in Science" by Patti Hartigan in *Harvard Education Letter*, July/August 2010 (Vol. 26, #4, p. 1-3) http://www.hepg.org/hel/article/472.

Back to page one

5. An High-School Administrator's Valedictory Thoughts

In this *Harvard Education Letter* article, veteran Evanston Township (IL) assistant superintendent Laura Cooper offers these thoughts as she rides off into the sunset:

• *We know a lot about teaching* – "Twenty-five years ago we routinely engaged in debates about whether teaching was an art or a science, arguing whether good teachers were 'born' or 'made,'" says Cooper. "Today, that debate is overshadowed by a powerful knowledge base on teaching and learning. We know that good instruction matters and that we don't get different student results without changing instruction." She touts Jon Saphier's classic

book, *The Skillful Teacher,* as a guide to continuously improving teaching and tuning it to students' learning.

• *Kids can get smart* – Cooper says she and her colleagues have been inspired by the work of Carol Dweck and Jeff Howard on the malleability of intelligence and the importance of getting students to attribute their successes to effective effort as a way to build their confidence, motivation, persistence, and sense of control over learning. "Instead of talking about what students can't do or aren't doing," she says, "we need to figure out how to convey belief about intelligence. We also need to figure out how to convey our belief in their capacity – even when they have given up on themselves. Then, and only then, can we engage them in learning how to 'get smart.'"

• *Talk honestly about race* – Cooper notes a tendency for educators to talk *around* the issue of race – the outstanding achievement of some black and Latino students, poverty and other root causes of underachievement, etc. But race *is* an issue. "People like Beverly Tatum and Glenn Singleton have provided us with the rationale, tools, and resources to learn how to talk honestly about race and ethnicity," she says. "Developing our skills and tapping our courage to have these conversations will give us a shared equity lens through which to examine our decisions about what to teach, how to teach, how to work together, and how to restructure our policies and practices as a school."

• *Collaboration is key* – Cooper credits Roland Barth with helping her to see how important the relationships among "grown-ups" are to student achievement. "Sometimes teachers blame administrators, administrators blame teachers, or we blame last year's teachers for not doing their part," she says. "We need to move beyond blame and learn to work together in many small groups – course-level Professional Learning Communities, department teams, and administrative-teacher teams. In PLCs we need to do the nitty-gritty work of matching instruction to our students' needs by setting course-specific goals, examining data including student work, and developing interventions to meet those goals. In departments we need to clarify what 'teacher autonomy' is and what it isn't and not let it get in the way of creating common, high-quality curricula and assessments." And professional development needs to be tied closely to this process.

• *Rigor for all students* – Cooper describes how Evanston was inspired by "a few brave schools" like Rockville Center, NY that experimented with detracking through tenth grade. She and her colleagues restructured some core courses so a broad range of students could have access to the most rigorous curriculum, and students rose to the challenge when they got the right support. Evanston also expanded access to AP courses and today more than a third of black and Latino seniors are enrolled (as are three-quarters of white seniors). "In order to increase achievement, we need to stop debating whether or not students have the knowledge or skills to qualify for an honors or AP class," she says. "Instead, we should work together to debate how to explicitly teach students how to be successful in these classes and how to align our courses to begin this preparation in ninth grade." The message to students should be:

- This work is really important.
- You have the capacity (or can develop it).

- I won't give up on you.

"Building on What We Know" by Laura Cooper in *Harvard Education Letter*, July/August 2010 (Vol. 26, #4, p. 8, 6-7) http://www.hepg.org/hel/article/471

Back to page one

6. Boy-Girl Differences in Reading

In this *Harvard Education Letter* article, Bard College professor Michael Sadowski explores the so-called "boy crisis" and concludes that there is indeed a worldwide boy-girl achievement gap in literacy. In the U.S. it holds true across race, ethnicity, and family income. But while girls outscore boys in reading in every geographic region, there are big differences *between* regions. For example, the percentage of eighth-grade boys scoring at the proficient level on the NAEP reading assessment in Connecticut is 16 points higher than that of girls in Mississippi. Boy-girl differences are also dwarfed by 20-, 30-, and even 40-point differences between white/Asian and African-American/Hispanic students.

What's going on here? "Most researchers agree that on average, boys develop the skills associated with reading and writing 12 to 24 months later than girls," says Sadowski. The key is attending to this achievement gap early on to prevent a "Matthew Effect" in which proficient readers surge ahead and struggling readers fall further and further behind. "There is a cumulative disadvantage associated with not reading very well," says Harvard literacy expert Catherine Snow. What is to be done? Sadowski offers the following:

• Bring reading material into classrooms that appeals to boys, including action-packed stories (yes, even some violence) and graphic novels that hook boys' interest and give them the skills they need to be successful in more academic texts.

• Provide support for struggling readers across the grades, using local data to identify boys (and girls) in need.

• Focus on the biggest achievement gaps which, within a state or district, may very well be racial-economic gaps. Researcher Naomi Chudowsky says that addressing gaps between subgroups is ultimately about improving the achievement of all students. "The goal for most educators would be for all groups to make gains," she says, "but for low-performing groups to make larger gains so that they catch up."

"Putting the 'Boy Crisis' in Context" by Michael Sadowski in *Harvard Education Letter*, July/August 2010 (Vol. 26, #4, p. 4-6) http://www.hepg.org/hel/article/473

Back to page one

Do you have feedback? Is anything missing?

If you have comments or suggestions, if you saw an article or web item in the last week that you think should have been summarized, or if you would like to suggest additional publications that should be covered by the Marshall Memo, please e-mail: kim.marshall8@verizon.net

About the Marshall Memo

Mission and focus:

This weekly memo is designed to keep principals, teachers, superintendents, and others very well-informed on current research and effective practices in K-12 education. Kim Marshall, drawing on 37 years' experience as a teacher, principal, central office administrator, and writer, lightens the load of busy educators by serving as their "designated reader."

To produce the Marshall Memo, Kim subscribes to 44 carefully-chosen publications (see list to the right), sifts through more than a hundred articles each week, and selects 5-10 that have the greatest potential to improve teaching, leadership, and learning. He then writes a brief summary of each article, pulls out several striking quotes, provides e-links to full articles when available, and e-mails the Memo to subscribers every Monday evening (with occasional breaks; there are about 50 issues a year).

Subscriptions:

Individual subscriptions are $50 for the school year. Rates decline steeply for multiple readers within the same organization. See the website for these rates and information on paying by check or credit card.

Website:

If you go to http://www.marshallmemo.com you will find detailed information on:

- How to subscribe or renew
- A detailed rationale for the Marshall Memo
- Publications (with a count of articles from each)
- Article selection criteria
- Topics (with a count of articles from each)
- Headlines for all issues
- What readers say
- About Kim Marshall (including links to articles)
- A free sample issue

Marshall Memo subscribers have access to the Members' Area of the website, which has:

- The current issue (in PDF or Word format)
- All back issues (also in PDF or Word)
- A database of all articles to date, searchable by topic, title, author, source, level, etc.
- How to change access e-mail or password

Publications covered

Those read this week are underlined.

American Educator
American Journal of Education
American School Board Journal
ASCD, CEC SmartBriefs, Daily EdNews
Catalyst Chicago
Ed. Magazine
EDge
Education Digest
Education Gadfly
Education Next
Education Week
Educational Leadership
Educational Researcher
Edutopia
Elementary School Journal
Essential Teacher (TESOL)
Harvard Business Review
Harvard Education Letter
Harvard Educational Review
JESPAR
Journal of Staff Development
Language Learner (NABE)
Middle Ground
Middle School Journal
New York Times
Newsweek
PEN Weekly NewsBlast
Phi Delta Kappan
Principal
Principal Leadership
Principal's Research Review
Reading Research Quarterly
Reading Today
Rethinking Schools
Review of Educational Research
Teachers College Record
The Atlantic Monthly
The Chronicle of Higher Education
The Language Educator
The Learning Principal
The New Yorker
The Reading Teacher
Theory Into Practice
Tools for Schools