

Testimony of Thomas Kane, Professor of Education and Economics, Harvard Graduate School of Education, before the Senate Health, Education, Labor and Pensions Committee

April 15, 2010

Chairman Harkin, Ranking Member Enzi, Members of the Committee, thank you for inviting me to address the committee today. My name is Tom Kane. I am a Professor of Education and Economics at the Harvard Graduate School of Education. I am currently on leave from Harvard, working with the Bill & Melinda Gates Foundation on the Measures of Effective Teaching project, which I will be describing today. My testimony will emphasize the importance of better teacher evaluation systems, which everyone agrees are perfunctory and meaningless. These days, we all say “teachers matter.” If we began acting like we believed that, we would be telling teachers how they are doing, what they can do to improve and giving principals the objective data they need to make better personnel decisions.

Summary

Analysis of student achievement data over the last four decades has repeatedly confirmed what we all know: Without the right people standing in front of the classroom, school reform is a futile exercise. Everything else—educational standards, testing, class size, greater accountability—is background, intended to support the crucial interactions between teachers and their students.

And, yet, almost everywhere, teacher evaluations are meaningless and perfunctory. From the moment they are assigned their first classroom, teachers receive almost no feedback on their performance. The failure to meaningfully differentiate among teachers and their teaching practices has enormous costs: teacher performance plateaus after just two or three years on the job; principals grant tenure to virtually any teacher willing to remain after just two or three years; many of the best and brightest teachers abandon the classroom for other occupations and industries with better opportunities for growth. Meanwhile, we all lose: student learning is stunted; effective teachers are saddled with ineffective colleagues; the profession loses status; and the nation continues to slowly bleed economic productivity and competitiveness.

We need to begin building a teacher performance evaluation system that allows teachers to grow and allows principals to make better decisions. Given the complexity of teaching, there is no single statistic which will tell the whole story. Rather, we need to assemble a small package of indicators-- student achievement gains on state tests, objective feedback on classroom practice by trained external observers, student feedback on specific aspects of a teacher’s practice (e.g. were their comments on homework assignments helpful?, do they have multiple ways of explaining a given topic?, was time managed well in class?)—and put it in the hands of teachers and principals.

Of course, there needs to be some discipline to the search for “multiple measures of teacher effectiveness,” lest that system become “multiple excuses for teacher ineffectiveness.” To guide our own efforts at tool development, we have adopted the following two principles:

1. *Whenever feasible, the measure should include student achievement growth for all the students for whom a teacher is responsible;*
2. *Any other measures— for instance, those based on classroom observations, supervisor ratings, student evaluations, teacher assessments—must be demonstrated to help identify the teachers with the strongest student achievement growth . That evidence needs to be updated annually, based on the latest student achievement growth data, to guard against grade inflation and gaming.*

Although a few states have the key ingredients to start, most states would need to build the infrastructure to support such a system: creating a workable definition of “teacher of record” for each tested student; ensuring accurate data on teacher-student links at the state level; calculating student achievement growth for students and linking those to teachers; piloting new classroom observations and other non-test-based tools to be validated against student growth. States choosing to go down that path could do so within 3 years.

Introduction

Throughout four decades of education research, researchers have repeatedly confirmed three findings: First, there are huge differences in student achievement gains in different teachers’ classrooms. Year after year, some teachers lead their students to remarkable gains in academic achievement, while others lag behind. Second, the data suggest these teachers can be found throughout our education system, not just in wealthy suburban schools. Most schools—public and private, urban and suburban (and rural), high- and low-income—have such teachers sprinkled within their ranks. Third, a teacher’s effectiveness has only a weak relationship with his or her paper qualifications. Despite the focus on teaching credentials in state and federal law, a “highly qualified” teacher is little more likely to produce exemplary gains in student achievement than others.

And yet, almost everywhere, teachers receive meaningless and perfunctory feedback on their performance on the job. A recent survey of teacher evaluation systems in 12 school districts across 4 states found that most systems provide for only two possible ratings (“satisfactory” and “unsatisfactory”). In those districts, more than 98 percent of teachers received the same rating of “satisfactory”.¹

The failure to meaningfully differentiate among teachers and teaching practices has enormous costs: in the absence of feedback, teachers plateau after just two or three years on

¹ Weisberg, Daniel, Susan Sexton, Jennifer Mulhern and David Keeling. *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*. The New Teacher Project, 2009.

the job²; without objective evidence to support their decisions, principals grant tenure to virtually any teacher willing to remain after just two or three years³; without the feedback to help them learn, many of the best and brightest teachers abandon the classroom for other occupations and industries with better opportunities for growth. Meanwhile, we all lose: student learning is stunted; effective teachers are saddled with ineffective colleagues; the profession loses status; and the nation continues to slowly bleed economic productivity and competitiveness.

The No Child Left Behind Act (NCLB) required states to track performance and provide feedback to students and their parents, as well as to schools and districts. Regrettably, teachers *were* left behind. Teachers were simply obliged to have the right credentials.

The phrase, “teachers matter”, now appears frequently in the public discourse. Despite its popularity, the statement glosses over the truth-- teachers (and teaching practices) not only matter, they *differ*. Teachers deserve to know how their students’ growth in achievement compares to the students assigned to their peers. Moreover, they deserve to know what they could be doing to *improve* their practice. To do so, they need professional feedback from objective experts observing them in their own classrooms. They deserve feedback on how students perceive their classrooms. Obviously, they need to understand the concepts they teach. However, we also need to be sure teachers recognize the most common ways in which students will *misunderstand* the content they will teach, and have specific strategies for responding.

Defining “Effective Teaching”: An Evidence-Based Approach

Over the years, educators have proposed a number of alternative approaches to defining and recognizing effective teaching practice. For instance, the National Board for Professional Teaching Practices (NBPTS) has developed an application process and set of rubrics for scoring videos and essays submitted by teachers. In 1996, Charlotte Danielson published her *Framework for Teaching*, a general framework for evaluating teaching. Many states already publish their own set of standards for evaluating teachers.

Rather than choose one view of effective teaching as “the standard” against which all teaching everywhere is measured, we adopt a practical, evidence-based approach. A measure

² Kane, Thomas J., Jonah Rockoff and Douglas O. Staiger “What Does Certification Tell us about Teacher Effectiveness? Evidence from New York City,” *Economics of Education Review*, December 2008. Many other studies report similar findings.

³ Less than 3 percent of new teachers (with 1-3 years of teaching) report having involuntarily moved between schools or out of teaching. Luekens, Michael T., Deanna M. Lyter, and Erin E. Fox.. *Teacher attrition and mobility: Results from the teacher follow-up survey, 2000–01*. Washington, D.C.: National Center for Education Statistics, Department of Education, 2004 Deanna Lyter provided some additional tabulations from the School and Staffing Survey generate these estimates.

of effective teaching should be a summary of what we know about a teacher's impact on children. Whenever a teacher is working in a grade or subject where it is possible to track their students' achievement growth, then past performance should be helpful in identifying those teachers most likely to have a positive impact on children in the future. *But other, non-test based measures of a teacher's practice and skills may also help identify the teachers producing student achievement gains.* For instance, classroom observations may identify specific practices linked to student achievement gains. Student feedback on the quality of teacher comments on their submitted work, the pace of classroom lectures, a teacher's ability to provide multiple explanations for any given topic may also help identify those teachers. Therefore, a measure of effective teaching should be limited to the combination of two categories of measures:

1. **Direct evidence of student growth on an objective measure of student achievement**, whether on a state's end-of-year or end-of-grade assessment or some other externally scored measure of student work (such as a benchmark assessment). Such measures are often referred to as "value added" measures.

And

2. **Other non-test-based measures** (such as classroom observations, student feedback, assessments of teacher content knowledge or pedagogical content knowledge) **which have been shown to identify those teachers with exemplary student achievement growth.**

Under the above principles, teacher effectiveness would be measured by direct measures of student achievement gains and any other measure which can be shown to identify those teachers most likely to produce student achievement gains.

When measures of effective teaching have been proposed in the past, skeptics have on various sides of the debate have raised the following concerns:

"How will this work in the non-tested grades and subjects?" Many grades and subjects are not currently subject to mandatory state testing. By some estimates, only about a quarter of teachers currently work in grades and subjects where value-added estimation is feasible. The above framework suggests a way to incorporate those non-tested grades and subjects in the evaluation system. A district could supplement the testing currently required under NCLB with additional externally scored measures of student work and extend to additional grades and subjects. Moreover, if a state or district can provide evidence that a particular classroom observation protocol, or a particular student evaluation form helps identify teachers in the tested grades and subjects with demonstrable gains in student achievement, then that same protocol or form could be used in the non-tested grades and subjects to provide feedback to

teachers. In other words, one could assume that the process of student learning is similar in the tested and non-tested grades and subjects.

“Aren’t value-added measures too volatile to be used?” Because elementary school teachers may have only 15 to 25 students per year and middle and high school teachers 40 to 150 students per year, the value-added measures can fluctuate. Especially for sample sizes typically seen for elementary school teachers, a few particularly rowdy or attentive students can lead to changes in student achievement gains. However, to the extent that the non-test-based measures (such as feedback from classroom observations or student evaluations) are less subject to volatility or where those fluctuations are independent of the volatility in student test performance, the inclusion of these other measures will dampen the volatility and lead to more stable measures.

“Aren’t ‘multiple measures’ just a dodge, a way to avoid holding teachers accountable for student results?” Because of the long history of perfunctory teacher evaluations in education, many are skeptical of any measure which does not include student achievement directly. This is understandable. However, the above framework would provide some discipline to the search for non-test-based evaluation tools: if a teacher evaluation tool (as implemented in a particular locale, and not based on one unrelated study in the research literature) cannot be shown to identify those teachers producing exemplary student achievement gains, that measure could not be a part of the teacher evaluation system.

“If teacher evaluation is limited to student gains on tests and those factors related to student gains on tests, isn’t the logic circular? Won’t we end up just encouraging ‘teaching to the test’?” No. Virtually every rubric which external raters would use to score their classroom observation emphasizes the importance of teaching concepts, and the ideas underlying course material. The teaching of “rote skills” and simple procedures leads to poor scores on those rubrics. There’s good reason for that, since research in cognitive science suggest that without conceptual understanding, students find it difficult to remember their lessons and extend and generalize what they have learned. As long as the state test includes some items requiring conceptual understanding, a teacher’s scores from the classroom observations will remain “predictive” of their students’ achievement gains. Moreover, the juxtaposition of respectable test-based scores and poor non-test-based measures raises a possible red flag. Thus, the teacher evaluation system provides an opening for a frank discussion between the teacher and the supervisor about the nature of their instruction—highlighting possible “teaching to the test”, not disguising it.

Similarly, student evaluation forms, such as that developed by The Tripod Project (see the Appendix), can ask students to agree or disagree with statements such as: “My teacher tells us what we are learning and why”, “My teacher wants us to use our thinking skills, not just memorize things”, “My teacher has several good ways to explain each topic that we cover in this class”. Poor student responses on such items, especially when contrasted with moderate or high student achievement gains, would flag possible instances of “teaching to the test”, rather than hiding them.

The Measures of Effective Teaching Project

The Bill & Melinda Gates Foundation is supporting the rapid testing of new and existing tools for providing feedback to teachers. The *Measures of Effective Teaching* (MET) project involves nearly 3000 teacher volunteers in six school districts around the country (NYC, Charlotte-Mecklenburg in NC, Hillsborough County (Tampa) in FL, Memphis in TN, Dallas TX, Denver CO). The national offices and local affiliates of both the AFT and NEA actively helped recruit teachers for the project.

Research partners include the RAND Corporation, Harvard University, Stanford University, University of Michigan, Dartmouth College, Educational Testing Service, the Danielson Group and the University of Virginia. Private contractors, such as Teachscape, Westat and Cambridge Education are providing vital logistical support. The National Board for Professional Teaching Standards (NBPTS) is providing the scores of any sample members who applied to them for certification.

The study is currently focusing on grades 4-8 (math and English language arts) and Algebra I, ninth grade ELA and biology at the high school level. The findings of the project (and many of the tools) will be shared widely and made available to states and districts who want to use them. Data collection is under way and initial results will be available in Fall 2010.

A *Teacher Advisory Panel* (20+ practicing teachers, grades 4-9, from across the US) has been convened to provide feedback on the project design, to ensure that it captures what is most important to excellent teachers, is feasible in the classroom, and provides opportunities for teachers to improve their practice.

The project is collecting a variety of data:

- Student achievement gains;
- Classroom observations using innovative, low-cost digital video collection tools;
- Teacher reflections on their videotaped lessons;
- Student feedback;
- Teachers' perceptions of the quality of instructional support in their schools; and
- Teachers' ability to recognize and diagnose student misperceptions.

I discuss each in somewhat more detail below:

Student achievement gains: Student achievement will be measured two ways—using the mandated state tests as well as supplemental tests (the latter made up of open-ended, constructed response items to probe higher order conceptual understanding). The goal is to evaluate the widespread concern that those teachers posting large gains on the state tests are merely teaching test-taking skills at the expense of higher-order conceptual understanding.

An innovative approach to classroom observation: Meaningful observations require input from external observers. The project is exploring new ways to drive down the cost of having external observers provide feedback on instruction using digital video. Digital video will be used to record four lessons per year in each teacher's classroom. Scorers will be trained via a web-based certification regimen to score those videos using rubrics designed for classroom observations. (Several commonly used rubrics, such as Charlotte Danielson's Frameworks for Teaching, the CLASS measure from the University of Virginia will be used, as well as content-based rubrics for observing math, English language arts, and science classrooms). The goal is not simply to test whether a select panel of experts can identify effective teaching, but whether qualified scorers could do so after a finite course of training.

Teacher reflections on their videotaped lessons: Teachers will provide audio commentary and any relevant supporting materials to provide context about the videotaped lessons, and to share their self-reflections.

Confidential Student Feedback: Students will also provide feedback on their experiences in each classroom, their level of engagement, their perception of teachers' expectations of them, their perception of the quality of the feedback they receive from their teacher, etc. (An example of the questions on the student evaluation form developed by Professor Ron Ferguson at Harvard's Kennedy School of Government are in the appendix.)

Teachers' Pedagogical Content Knowledge: ETS is developing a new assessment to measure teachers' ability to recognize and diagnose common student misperceptions in their grade level and subject. That assessment, which builds on work started at the University of Michigan for assessing teacher's pedagogical content knowledge in mathematics, will be ready in the spring of 2011. (An example of several items from such a test are included in the appendix.)

Teacher surveys of the school environment: Teachers will complete a survey about working conditions and the instructional support they receive in school. (Representative items from the survey are included in the attached appendix.)

How Could a State Build Its Own System for Evaluating Teachers?

Below, I sketch out a three-year process by which a state could develop a new teacher evaluation system. This is merely intended as illustrative of the hurdles to be cleared. Some states could move faster; some slower.

Year 1:

Planning: The first task is to agree upon an approach to assigning a “teacher(s) of record” for each tested student by subject. A number of large school districts—such as Houston and Dallas in Texas, Hillsborough County in Florida— have shown that its possible to get accurate data on which students are assigned to specific teachers and to resolve the thorny issues which arise when students are taught by teams of teachers or when students receive special help in certain subjects. Indeed, the many districts that are using “benchmark” or “interim” assessments during the school year already have such links in place—that is how they know which teachers should receive the results for which students. Therefore, states and districts would need to plan how they would transfer such data from districts to states.

Choosing a Value-Added Model: There have been a number of different approaches proposed for estimating the impact of a given teacher on students’ achievement growth. For example, William Sanders of SAS has provided scores using the proprietary Education Value-Added Assessment System (EVAAS) to hundreds of school districts, including Houston in Texas and Hillsborough County in Florida. The Wisconsin Center for Education Research provides value-added reports to principals in Chicago Public Schools and New York City. Massachusetts, Colorado and New Haven, Connecticut have adopted an approach to measuring student growth proposed by Damian Betebenner at the Center for Assessment, which can be calculated for individual teachers. Each of these approaches involves a different trade-off between transparency and statistical reliability, which states and districts will have to weigh on their own.

Year 2:

Piloting Some Non-Test-Based Approaches to Teacher Evaluation: During the second year, any state intending to build its own measure of teacher effectiveness would begin to pilot some non-test-based measures, such as having external observers provide feedback on instructional practice in schools, or having students provide feedback on their own experiences. In order to test the predictive power of these measures for identifying teachers with exemplary student achievement growth, such piloting should include teachers in the grades and subjects where value-added estimates could be generated.

Calculate Value-Added and Test Predictive Power: The state would spend the second year implementing the definition of “teacher of record” developed during the first year and creating teacher-student links at the state level. It would be important in that process to

provide each teacher with the list of their students for whom they are responsible so that they could correct any errors in the data. Then, once student achievement data become available, a state could identify the subset of teachers for whom both value-added estimates and the other teacher evaluation data are available, and then identify which of the non-test based measures are demonstrably correlated with value-added. A state might adopt a minimum acceptable level of correlation in order to be accepted as a predictor of teacher effectiveness.

Year 3:

Year 3 would represent the first full year of implementation. In tested grades and subjects, teachers would be evaluated on the basis of both value-added and non-test-based predictors of student achievement growth. In non-tested grades and subjects, teachers would be evaluated on the basis of the validated predictors of value-added from the tested grades and subjects.

Year 4+:

As teachers and principals adapt to the new evaluation system, there may be the equivalent of “grade inflation”, with teacher evaluations becoming more compressed or rampant gaming of the system. Moreover, once an evaluation system has been scaled statewide, the new tools may not be implemented with the same fidelity as observed during the pilot phase. In those cases, the predictive power of some teacher evaluation tools may degrade over time. Moreover, new approaches to assessment student achievement may become available—particularly if states adopt new assessments to accompany the new common standards. As a result, the predictive power of the non-test-based approaches will need to be re-evaluated at regular intervals (e.g. annually) in order to ensure that the evaluation system continues to improve and evolve.

Bang for the Buck

How large are the potential payoffs to investing in teacher evaluation systems? Gordon, Kane and Staiger (2006) studied elementary teachers in Los Angeles Unified District, calculating value-added scores for a set of teachers who remained teaching for at least three years.⁴ They first sorted teachers into quartiles using their value-added during their first two years of teaching. They then observed the student achievement gains for the new crop of students they were assigned during their third year. They found that the average student assigned to a teacher whose value-added was in the bottom quartile of new teachers lost on average 5 percentile points relative to students with similar baseline scores and demographics.

⁴ Gordon, Robert, Thomas J. Kane and Douglas O. Staiger, *Identifying Effective Teachers Using Performance on the Job*, Hamilton Project White Paper 2006-01, Brookings Institution, April 2006.

In contrast, the average student assigned to a top-quartile teacher gained 5 percentile points relative to students with similar baseline scores and demographics.

They then simulated the effect of a new tenure review system, in which those in the bottom quartile of effectiveness during their first two years of teaching would not receive tenure and would have to leave teaching. Even taking into account the need to hire more novice teachers, they estimated that such a policy would raise student achievement by the time of high school graduation by roughly 14 percentile points.

Still, that figure may understate the potential impact of an improved teacher evaluation system. For instance, it assumed *no improvement* in the effectiveness of any teacher in response to the feedback. The gains resulted from better selection at tenure time alone. If teachers were to use the feedback to improve their practice, the impacts could be larger.

In contrast, a random assignment evaluation of a classroom size reduction in Tennessee found that schools could improve achievement by half as much—5 percentile points—by shrinking class size in early grades⁵. (As a cautionary note, these impacts were considerably larger than the impacts that were experienced following California’s classroom size reduction policy beginning in 1997. They may substantially overstate the actual impact of such a policy.) But class size reduction of the magnitude tried in Tennessee would be extraordinarily expensive: shrinking average class size from twenty-two to sixteen students per class would require nearly a 40 percent increase in the number of teachers and the amount of classroom space in those early grades! In other words, a policy of tenure reform is estimated to generate an improvement in student achievement three times as large as class size reduction—and would almost surely cost less than a 40 percent increase in instructional salaries in the early grades.

Conclusion

Ultimately, the success of education reform depends upon the skills of the 3.1 million teachers managing classrooms in elementary and secondary schools around the country. Everything else—educational standards, testing, class size, greater accountability—is background, intended to support the crucial interactions between teachers and their students. Without the right people standing in front of the classroom, school reform is a futile exercise.

Our educational system has never acknowledged that fact. If it did, we would be rigorously evaluating teachers during their first few years of teaching, and ensuring that only the most effective are granted tenure. We would be providing feedback to teachers on the specific areas where they are falling short, so that they could improve. We would be

⁵ Krueger, Alan B. “Experimental estimates of education production functions”. *Quarterly Journal of Economics* (1999) Vol. 114, No. 2, pp. 497–532.

identifying the most effective teachers and making every effort to retain them. We would be ensuring that the students who are most far behind have the teachers they need to catch up.

Successful education reform requires having the right people with the right skills in the classroom. We will never do so without a robust system for evaluating and providing feedback to teachers.

Thank you again for this opportunity to discuss this important issue with the committee. I would be pleased to serve as an ongoing resource to the committee as the results from the Measures of Effective Teaching project begin to emerge this summer and fall. I look forward to your questions.

Appendix:

Sample items from Student Evaluations

Source: Tripod Project & Cambridge Education

Sample Items: Elementary Student Survey

The teacher in this class encourages me to do my best.
Our class stays busy and does not waste time.
My teacher has several good ways to explain each topic that we cover in this class.
In this class, we learn to correct our mistakes.
My teacher pushes us to think hard about things we read.
My teacher wants me to explain my answers -- why I think what I think.
In this class, you must pay attention all the time in order to keep up.
My teacher gives us time to explain our ideas.
My teacher tells us what we are learning and why.
My teacher asks questions to be sure we are following along when s/he is teaching.
In this class, I stop trying when the work gets hard.
I have done my best quality work in this class.

Sample Items: Secondary Student Survey

The teacher in this class encourages me to do my best.
My classmates behave the way my teacher wants them to.
If you don't understand something, my teacher explains it another way.
My teacher explains difficult things clearly.
My teacher wants us to use our thinking skills, not just memorize things.
My teacher makes us think first, before he/she answers our questions.
In this class, my teacher accepts nothing less than our full effort.
My teacher makes lessons interesting.
I understand what I am supposed to be learning in this class.
My teacher knows when the class understands, and when we do not.
In this class, students take it easy, and do not try very hard to do their best.
In this class we have to think hard about the writing we do.
This class makes me a better thinker.

Sample Items from Teacher Working Conditions Survey

Source: The New Teacher Center

Teachers have time available to collaborate with colleagues.
Teachers are allowed to focus on educating students with minimal interruptions
Teachers are protected from duties that interfere with their essential role of educating students
Teachers* have sufficient access to appropriate instructional materials.
Teachers have access to reliable communication technology, including phones, faxes and email.
The school environment is clean and well maintained.
Teachers have adequate space to work productively.
The physical environment of classrooms in this school supports teaching and learning.
This school maintains clear, two-way communication with the community.
This school does a good job of encouraging parent/guardian involvement.
Parents/guardians know what is going on in this school.
The community we serve is supportive of this school.
Students at this school follow rules of conduct.
Administrators support teachers' efforts to maintain discipline in the classroom.
The faculty works in a school environment that is safe.
Teachers are recognized as educational experts.
Teachers are trusted to make sound professional decisions about instruction.
The faculty and leadership have a shared vision.
Teachers feel comfortable raising issues and concerns that are important to them.
Teachers are held to high professional standards for delivering instruction.
Teacher performance is assessed objectively.
The procedures for teacher evaluation are consistent.
An appropriate amount of time is provided for professional development.
Professional development deepens teachers content knowledge
Teachers are encouraged to reflect on their own practice.
State and local assessment data are available in time to impact instructional practices.
Teachers at my school are assigned classes that maximize their likelihood of success with students.

Study of Instructional Improvement/Learning Mathematics for Teaching

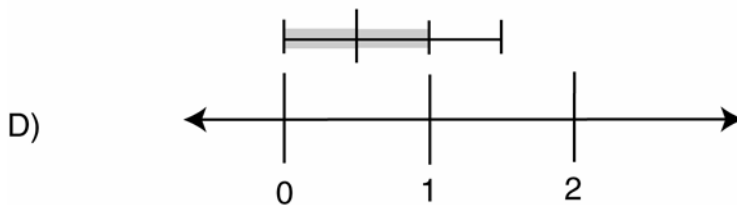
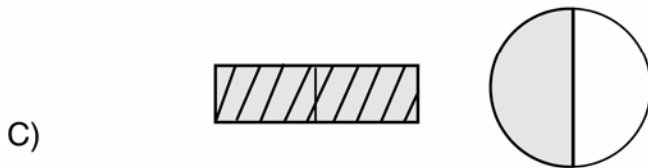
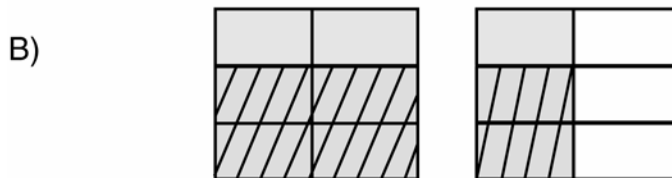
Content Knowledge for Teaching Mathematics Measures (MKT measures)

Released Items, 2008

ELEMENTARY CONTENT KNOWLEDGE ITEMS

6. At a professional development workshop, teachers were learning about different ways to represent multiplication of fractions problems. The leader also helped them to become aware of examples that do not represent multiplication of fractions appropriately.

Which model below cannot be used to show that $1\frac{1}{2} \times \frac{2}{3} = 1$? (Mark ONE answer.)

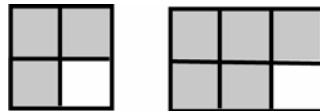


7. Which of the following story problems could be used to illustrate $1\frac{1}{4}$ divided by $\frac{1}{2}$? (Mark YES, NO, or I'M NOT SURE for each possibility.)

	Yes	No	I'm not sure
a) You want to split $1\frac{1}{4}$ pies evenly between two families. How much should each family get?	1	2	3
b) You have \$1.25 and may soon double your money. How much money would you end up with?	1	2	3
c) You are making some homemade taffy and the recipe calls for $1\frac{1}{4}$ cups of butter. How many sticks of butter (each stick = $\frac{1}{2}$ cup) will you need?	1	2	3

ELEMENTARY KNOWLEDGE OF STUDENTS AND CONTENT ITEMS

15. Takeem's teacher asks him to make a drawing to compare $\frac{3}{4}$ and $\frac{5}{6}$. He draws the following:

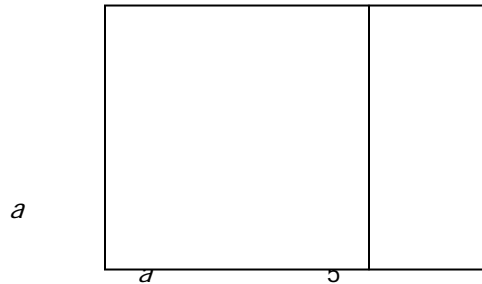


and claims that $\frac{3}{4}$ and $\frac{5}{6}$ are the same amount. What is the most likely explanation for Takeem's answer? (Mark ONE answer.)

- a) Takeem is noticing that each figure leaves one square unshaded.
- b) Takeem has not yet learned the procedure for finding common denominators.
- c) Takeem is adding 2 to both the numerator and denominator of $\frac{3}{4}$, and he sees that that equals $\frac{5}{6}$.
- d) All of the above are equally likely.

MIDDLE SCHOOL CONTENT KNOWLEDGE ITEMS

20. Ms. Whitley was surprised when her students wrote many different expressions to represent the area of the figure below. She wanted to make sure that she did not mark as incorrect any that were actually right. For each of the following expressions, decide whether the expression correctly represents or does not correctly represent the area of the figure. (Mark REPRESENTS, DOES NOT REPRESENT, or I'M NOT SURE for each.)



		Does not correctly		
	Correctly represents	represent	I'm not sure	
a) $a^2 + 5$	1	2	3	
b) $(a + 5)^2$	1	2	3	
c) $a^2 + 5a$	1	2	3	
d) $(a + 5)a$	1	2	3	
e) $2a + 5$	1	2	3	
f) $4a + 10$	1	2	3	

26. Mrs. Davies' class has learned how to tessellate the plane with any triangle. She knows that students often have a hard time seeing that any quadrilateral can tessellate the plane as well. She wants to plan a lesson that will help her students develop intuitions for how to tessellate the plane with any quadrilateral.

Which of the following activities would best serve her purpose? (Circle ONE answer.)

- a) Have students cut along the diagonal of various quadrilaterals to show that each can be broken into two triangles, which students know will tessellate.
- b) Provide students with multiple copies of a non-convex kite and have them explore which transformations lead to a tessellation of the plane.
- c) Provide students with pattern blocks so that they can explore which of the pattern block shapes tessellate the plane.
- d) These activities would serve her purpose equally well.

27. Ms. Abdul is preparing a unit to introduce her students to proportional reasoning. She is considering three versions of a problem that are the same except for the numbers used. Which version of the Mr. Short and Mr. Tall problem below is likely to be the most challenging for students? (Circle ONE answer.)

- a) A picture depicts Mr. Short's height as 4 paper clips and as 6 buttons. The height of Mr. Tall (not shown) is given as 6 paper clips. How many buttons in height is Mr. Tall?
- b) A picture depicts Mr. Short's height as 4 paper clips and as 7 buttons. The height of Mr. Tall (not shown) is given as 5 paper clips. How many buttons in height is Mr. Tall?
- c) A picture depicts Mr. Short's height as 2 paper clips and as 9 buttons. The height of Mr. Tall (not shown) is given as 5 paper clips. How many buttons in height is Mr. Tall?
- d) All three of the problems are equally challenging.