1. Questions(s) or issue(s) for dialogue at Learning Network Conference session:

The Teacher Knowledge Assessment System (TKAS) is an online system for administering the Learning Mathematics for Teaching (LMT) assessment questions. The LMT questions are designed to measure mathematical knowledge for teaching in the areas of elementary and middle school geometry, patterns functions and algebra, number and operations, proportional reasoning, data statistics and probability, and rational number. For each of these areas there are parallel forms suitable for use as pre- and post-session assessments. Many MSP projects have successfully used the LMT measures to study teacher learning and assess program effects.

In this session, we will demonstrate the capacity of TKAS. The system allows MSP users to create their own assessment plans by selecting content areas of interest and dates of administration. The LMT questions can be administered using either set forms or a Computer Adaptive Testing (CAT) option. After teachers complete the assessment, MSP program leaders can download teacher data directly from TKAS. The session will also cover new TKAS features that are currently under development. These include online training, an integrated analysis feature, and the capacity for users to select questions and build their own tailored assessment form.

Finally, we will discuss a research project that is being conducted using the TKAS platform. All TKAS users are required to complete a survey reporting on the content and design of their professional development program. This data will be used to conduct a larger-scale meta analyses of the effects of professional development on teacher learning.

2. Context of the work within the STEM education literature and within your MSP project:

Debates about what teachers need to know in order to teach and where they should learn that knowledge stretch back to the 19th century (Hill, Ball, Sleep, & Lewis, 2007). Empirical studies soon followed, and by the 1980s, there was accumulating evidence that teacher knowledge as measured by tests of verbal facility or basic mathematical skills was a prime contributor to student achievement (e.g., Boardman, Davis & Sanday, 1977; Ferguson 1991; Hanushek 1972; Strauss & Sawyer, 1986). Over this period there was also increasing public concern about the quality of the teacher corps. This led in the 1980s to the institution of basic-skills certification assessments for teachers, including those that focused on basic mathematical knowledge. It has also led to a number of alternative certification programs designed to increase the number of teachers with strong subject matter backgrounds, often at the expense of having these recruits complete formal teacher education coursework (e.g., Teach for America, American Board for Certification of Teacher Excellence).

Teacher educators object to the premise that strong subject matter training alone yields good teaching, and have strenuously protested these new tests and certification routes (Darling-
According to this view, measuring quality teachers through performance in college or on tests of basic verbal or mathematics ability may overlook key elements in what produces quality teaching. Chief among these elements, many propose, is profession-specific knowledge of subject matter. Shulman (1986) and colleagues are widely seen as having sparked this line of research by proposing the idea of “pedagogical content knowledge,” or knowledge that goes “beyond knowledge of subject matter per se to the dimension of subject matter knowledge for teaching” (p. 9, emphasis added). Components of pedagogical content knowledge, according to Shulman (1986), are representations of specific content ideas, as well as an understanding of what makes the learning of a specific topic difficult or easy for students.

The Learning Mathematics for Teaching (LMT) project’s work builds from Shulman’s observation by focusing squarely on the question of the knowledge needed for teaching, and how that knowledge can be developed in teachers. We have hypothesized a theoretical framework for defining and organizing this knowledge and developed, pilot-tested, validated and disseminated a set of measures for gauging teachers’ “mathematical knowledge for teaching” (MKT).

MKT includes all the mathematical knowledge used in carrying out the work of teaching, and includes many of the elements described by Shulman, including pedagogical content knowledge (Ball, Thames & Phelps, 2008). MKT differs from PCK, however, in that we more clearly articulate the characteristics of the subject matter knowledge needed for teaching. In this view, teachers need “common content knowledge,” or the mathematical knowledge that is used in common across many professions. Knowing a number between 1.1 and 1.11, that zero is an even number, or how to multiply 35 x 25 are examples of this kind of knowledge. Another dimension of subject matter knowledge is “specialized content knowledge”—or the mathematical knowledge and skill uniquely needed by teachers in the conduct of their work. In looking for patterns in student errors or in sizing up whether a non-standard approach works in general, teachers have to do a kind of mathematical work that others do not. Many everyday tasks of teaching require mathematical knowledge of this kind, including providing answers when students ask “why,” representing and linking representations of key mathematical ideas, and connecting new content to that already learned.

3. Claim(s) or hypothesis(es) examined in the work (anticipating that veteran projects will have claims, newer projects will have hypotheses):

The Teacher Knowledge Assessment System provides an opportunity to test two important hypotheses.

1- The first hypothesis and related set of questions focus on the use of Computer Adaptive Testing (CAT). The main advantage of CAT administration is that individual questions are selected based on a participant’s prior answers. This allows the test to adapt to a participant’s ability level by administering questions of appropriate difficulty. In theory, CAT administrations will be both shorter and provide more precise measures across the full participant ability spectrum. We will investigate two questions related to the CAT administration.

• Are CAT assessments substantially shorter than the analogous traditional fixed form assessments?
• Does CAT provide more precise estimates of participant knowledge across the full range of participant ability?
2- The second set of questions focus on the use of LMT assessments and associated surveys in TKAS to study professional development and teacher learning. When projects use TKAS to assess professional development, the PD provider and teachers involved in PD will complete surveys asking about the design of the professional development and about their own personal background. We will use this data along with both PD provider and teacher assessment scores to conduct a series of large-scale meta analyses to answer the following questions.

- What are the average effects for professional development programs on teachers’ mathematical knowledge for teaching?
- Do the effects of professional development programs vary by subject matter assessment and by program length? Do these effects vary by the average mathematics experience and pre-test knowledge of teacher participants?
- What features of professional development programs are consistently and most strongly associated with gains in teachers’ knowledge?
- Are particular program features more effective for teachers with different levels of mathematical knowledge, professional experience, and other teacher background characteristics?

4. Evaluation and/or research design, data collection and analysis:

To ensure the integrity of these instruments, the LMT project has conducted in-depth psychometric and validation studies to determine how the instruments performed. We have conducted studies in the following areas.

- **Studies to identify constructs tapped by the measures.** We determined the dimensionality of our items by subjecting data from ten large-scale pilot forms (n=450-1000) to factor analyses and related analytic techniques (Reckase, 1985). Knowledge of content and students (KCS) items form a factor separate from subject matter knowledge, as hypothesized by our theoretical formulation. And we found that common and specialized content knowledge are distinguishable from one another (Hill, Schilling & Ball, 2004; Schilling, 2007). We also examined teacher and non-teacher reasoning and found that teachers use knowledge of content and students unique to the profession; non-teachers rarely used this knowledge (Hill, Dean & Goffney, 2007).

- **Studies to determine reliability of measures.** Each of the large-scale pilots is scaled and scored using Item Response Theory (IRT) methods and software. Reliabilities range from .50 for a specific set of KCS items to .90 with more recent forms. Technical reports for many of the forms are available at www.sitemaker.umich.edu/lmt under “summary of technical information.”

- **Studies to determine elemental validity.** Elemental validity determines whether specific items tap the knowledge they are designed to gauge, or whether they tap construct-irrelevant variance (e.g., reading ability, test-taking skills) (Schilling & Hill, 2007; Schilling, Blunk & Hill, 2007). We used cognitive interviews to investigate elemental validity, and found strong evidence for the validity of subject matter items.

- **Studies to determine ecological validity.** Ecological validity refers to the relationship of the test to its environment (Schilling & Hill, 2007). To determine ecological validity, we conducted two studies. In the first we predicted gains in student achievement using MKT and controlling for other critical background variables at the teacher, student, and school level (Hill, Rowan, & Ball, 2005). We found a positive relationship between MKT and gains in student achievement in a sample of 125 schools and 600 teachers participating in school
improvement programs. In the second study, we administered the multiple-choice measures and collected a sample (n=9) of lessons from ten teachers. Using a video coding rubric specifically designed to capture the mathematical quality of classroom work (e.g., presence of teacher errors; richness of the mathematics; teacher skill in understanding and using student ideas), we coded these lessons and found correlations of .5-.8 between facets of classroom mathematical quality and multiple-choice scores (Hill, Blunk, Charalambous, Lewis, Phelps, Sleep, & Ball, 2008).

- Finally, we have used these measures to evaluate the efficacy of a program intended to improve teachers’ mathematical knowledge. We found that the program did contribute to detectable teacher gains between pre- and post-test scores (Hill & Ball, 2004; Hill, Ball & Schilling, 2008). Our current research using the TKAS system is designed to expand this work and provide both more nuanced and robust analyses of professional development programs.

5. Key insights (retrospective for veteran projects, prospective for newer projects) that have value for the Learning Network:

In her 2004 AERA presidential address, Hilda Borko called for more rigorous study of teacher professional development. While acknowledging the importance of close qualitative work, Borko pointed to a pressing need for studies designed to examine the contributions of professional development across multiple contexts and for programs with different design features. She noted that these studies “will require new data collection and analysis tools—for example, instruments to measure change over time in teachers’ subject matter knowledge for teaching” (Borko, 2004, p. 13). These arguments are echoed by the National Mathematics Panel Report. “Despite widespread beliefs about the qualities that make teacher education effective, the Panel did not find strong evidence for the impact of any specific form of, or approach to, teacher education on either teachers’ knowledge or students’ learning” (NMP, 2008, p.40). The panel goes on to argue that, “given the vast investment made in teacher education, knowledge about the effectiveness of different approaches is vitally needed. Well-conceived efforts to improve the outcomes of teacher education, to improve measures of those outcomes, and to implement better research strategies should be supported” (NMP, 2008, p. 41). They identify a number of critical questions, including how the duration, structure, quantity, content, pedagogy, and relationship to practice of professional development effects both teachers’ and their students’ learning (NMP, 2008, p.41).

These recommendations highlight two critical issues that need to be addressed by current research. First, they point to the need for large-scale studies. The majority of research into teachers’ learning has been small-scale qualitative studies (see for example, Simon & Schifter, 1991; Jacobson & Lehrer 2000; Wilson, Mattson & Lubienski,1996; see Wilson & Berne, 1999 for this argument more generally). While qualitative research and case studies play an important role highlighting issues with the quality of professional development or identifying features of programs that support teacher development and the dynamics of this learning, they cannot answer questions about how specific features of professional development programs are associated with teachers’ learning, for single programs contain multiple features, and thus it is difficult to tell which of these features are effective. In addition, the findings are often limited to the program or sample of teachers included in the study.

This first issue is directly related to a second. Until recently, researchers have had a limited set of tools at their disposal to study teachers’ knowledge and its growth. Large-scale survey studies have traditionally employed measures of teacher characteristics such as mathematics courses taken, degrees attained, certification, and teaching experience (Begle, 1979). Some have relied on teachers’ own assessment of the professional development or their own learning. In general, researchers have made use of such indirect or proxy measures of teacher knowledge. While
informative, these studies cannot provide guidance on how professional development programs are related to teachers’ actual knowledge and learning. One exception is a study conducted by Hill and Ball (2004) investigating the effects of the California Mathematics Professional Development Institutes on teachers’ mathematics learning. They examined the effects of 15 different institutes on teachers’ learning and found that the institutes had an overall positive effect on teachers’ learning. They also found that both institute length and mathematical focus (e.g., proof/mathematical analysis, communication, representation) were associated with growth in teachers’ mathematical knowledge for teaching. While this study was hampered by the use of data from a pilot version of the LMT instruments and by a generally small sample size, the results suggest that certain program features can have unique effects on teachers’ learning.

By using TKAS to gather data on professional development and teacher knowledge, we seek to contribute to research on professional development by continuing to make widely available measures of mathematical knowledge for teaching for use in both small- and large-scale survey research. We also propose conducting one of the first larger-scale studies to examine how specific features of professional development programs are associated with the development of mathematical knowledge for teaching. We will do this across multiple areas of mathematical knowledge, for teachers of varying backgrounds, and for programs with a variety of design features.

References


