THE RELATIONSHIP BETWEEN MATHEMATICAL KNOWLEDGE FOR TEACHING, MATHEMATICS INSTRUCTIONAL QUALITY AND STUDENT OUTCOMES

Heather Hill & NCTE Team, January 2013
In this talk…

- **Preliminary look at a large study designed to investigate linkages between resources, instruction, and student outcomes**
  - Examining Year 1 data only
  - Data literally hot off the presses – results may change (though not by much)

- **Thoughts on a sub-study in which we compared the instruction of high and low “value-added” teachers**

- **Discussion about how CCSSM, math reforms historically, and new teacher accountability may collide**
Part I

Sample & Instrumentation
Sample

- 274 teachers from 4 districts participated in a three-year study of mathematics teaching and learning
- Fourth and fifth grade teachers
- Two districts using Investigations; one using Everyday Math mixed with supplemental materials; one using Harcourt Brace
- Three separate states
- Basic premise: take measures of lots of stuff, correlate them all
  - Non-causal
Instrument: Mathematical Knowledge for Teaching Assessment

- Used the MKT measures developed by Learning Mathematics for Teaching project (Michigan)
- Contained a mix of items designed to capture the knowledge used in teaching, rather than just pure mathematical knowledge
- Administered a teacher survey with 32 MKT items
- Reliability of 0.85
Instrument: The MQI

- **MQI** is the Mathematical Quality of Instruction instrument
- Allows observers to evaluate the quality of the *mathematics* in instruction
- Provides separate teacher scores for different dimensions of the mathematical work teachers do
Dimensions of the MQI

- **Richness of the mathematics**
  Captures meaning of facts/procedures (Linking, explanations)
  Mathematical Practices (Multiple procedures, generalizations, mathematical language)

- **Errors and imprecisions**
  Captures teachers’ lack of clarity, language imprecisions, major mathematical errors

- **Working with students**
  Captures remediation and using students’ mathematical ideas
Dimensions of the MQI

- **Student Participation in Meaning-Making and Reasoning**
  - Captures student explanations, reasoning, task cognitive demand

- **Classroom work connected to mathematics**
  - Captures time off task, presence of mathematically vacuous tasks

- **Overall MQI**
  - Holistic judgment by raters of entire lesson quality
MQI logistics

- Uses video-recorded lessons
  - Expert raters
  - Our studies show that play-back is critical for errors dimension
- Three lessons per teacher, each lesson scored by two raters
  - Lessons must be spread out
  - Raters must be highly skilled, carefully trained \( n=30 \)

<table>
<thead>
<tr>
<th>MQI Dimension</th>
<th>Reliability estimates</th>
<th>Errors</th>
<th>Working with students</th>
<th>Student participation</th>
<th>Overall MQI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richness</td>
<td>0.80</td>
<td>0.75</td>
<td>0.68</td>
<td>0.82</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Instrument: Value-added scores

- Designed to compare teachers based on how much they contribute to student outcomes

  - Simplified example: My sister and I both teach 4th grade. Our incoming fourth graders scored, on average, 50 points at the end of 3rd grade. At the end of 4th grade, my students score 75 points, my sister’s students score 85 points. My sister would have a higher “value-added” score than me.

  - In practice: Many adjustments made, for student backgrounds, composition of class, schools
$a_{i,j,k,t} = A_{i,t-n} \alpha + S_{i,t} \beta + P_{j,k,t} \delta + T_{k,t} \gamma + E_{i,t} \rho + \nu_{i,j,k,t}$, where $\nu_{i,j,k,t} = \mu_k + \theta_{k,t} + \epsilon_{i,j,k,t}$

- Used a hierarchical linear model to fit (Raudenbush & Bryk, 2002)
  - $\alpha$ is a standardized state math scale score, current year
  - $A_{t-1}$ is the prior year’s math score, reading score, squares and cubics
  - $S$ is a vector of student characteristics
  - $P$ is peer (classroom) characteristics
  - $T$ could be teacher covariates (not used here)
  - $E$ denotes district fixed effects

  Kicks out teacher scores
Very early analytic strategy

- Correlate MKT, MQI and teacher value-added scores
  - Hope: There’s some signal there
  - Why: It’s commonly thought that teacher resources produce instruction, which produce student learning
  - Obviously more interactive, but that’s later analysis (!)
Part II

Results
**MKT vs. VAM**

- Much written about this (including some by me)
  - Hill, Rowan, Ball – MKT and related variables explained about 20% of teacher-level variance
  - Hill, Kapitula, Umland – Correlations range from 0.46 to -0.15, depending upon covariates included in VAM model
  - Here: Correlation of 0.11, significance level of $p = 0.11$

- Upshot: Weak direct effect
MKT vs. MQI: Does MKT predict the mathematical quality of instruction?

$r=0.44$
MKT vs. MQI by district

MKT by MQI 5
Year 1 By District

Graphs by district
Dimension-specific correlations with MKT

- Classroom work is connected to mathematics: 0.08
- Richness: 0.36***
- Working with students: 0.32***
- Errors and imprecision: -0.36***
- Student participation in mathematical meaning-making and reasoning: 0.26***

*** = significant at less than p < 0.001
**Upshot**

- MKT significantly related to all major dimensions of MQI
- Relationship is weaker than other studies
- Other contributors to MQI
  - Curriculum materials, other teacher resources (e.g., beliefs), other resources in the setting, students themselves
    - Part of the project: Sorting some of these out
  - Measurement error
MQI and teacher VAM scores

$r=0.24^{***} \ (P < 0.001)$
MQI and teacher VAM by district

MQI 5 by VA
Year 1 By District

[Graphs showingscatter plots for Year 1 districts 11 to 14, with teacher MQI score on the y-axis and teacher VA score on the x-axis.]

- MQI Lesson Code: Whole-Lesson MQI, 5 Point Scale
- Fitted values

Graphs by district
MQI dimensions and VAM

- Classroom work connected to mathematics: 0.18**
- Richness: 0.15*
- Working with students: 0.17**
- Errors -0.11~
- Student participation in meaning-making and reasoning: 0.15*
Upshot, part II

- We cannot right now tell a causal story.
- Modest but significant relationship between overall MQI and VAM scores.
- Specific dimensions related to VAM as well.
  - Weaker – lots of noise generally.
About that noise…

- Sampled three highest/lowest VAM teachers from each district
- Groups of four people each watched all the video from each teacher (including Year 2 video, if available)
- Made a guess as to whether, based on the visible instruction, the teacher was a high value-added or low value-added teacher
  - Used not only MQI but other aspects of instruction, like classroom climate, efficiency, pace of instruction
Results

- Out of 17 teachers completed to date, we have been wrong 6 times.
- In many cases, we had difficulty seeing the teacher as either high or low (strong/weak) based on the instruction – down the middle.
- A team of experts is NOT particularly good at picking out high and low-VAM teachers at the extremes.
- Not unusual – MET report shows very low correlations between value-added and observer ratings.
- Non-expert raters (less training, less subject matter knowledge) may make the situation worse.
Before moving on: Any questions about this study?

☐ We’ll discuss teacher evaluation next, but first, any questions about this study?
Current reforms in teacher evaluation and mathematics teaching

- In all Race to the Top States, and in many NCLB waiver states, teacher evaluation must have both VAM and observational components
  - Recent Measures of Effective Teaching reform calls for weighting 1/3, 1/3, 1/3 (last=student surveys)
  - Most of those observation instruments are content-general (Framework for Teaching, Marzano)
- This is being placed in a context in which...
  - Many have been working to make mathematics teaching and learning more ambitious
  - Common Core State Standards in Mathematics continue this trend
Questions

- Will new teacher accountability levers help or hurt efforts to extend mathematics reform/implement the CCSSM?
  - Please vote!

- How can we mitigate some of the potential downsides of the implementation of teacher accountability alongside CCSSM?
Any questions? Email:
Heather_hill@harvard.edu