FIRSTMATH Meeting: June 7th, 2011

Introductions

Teresa’s Presentation:

Overview of TEDS-M Study

(Powerpoint will be uploaded to the FIRSTMATH website)

Overview of FIRSTMATH

Question: Saudi Arabia: how can I know particular knowledge comes from school context? TT: That’s a challenge for us. Look at patterns in induction. Cannot test this in a causal manner.

Edward Silver (History of TEDS-M and FIRSTMATH type studies)

We don’t understand something so well as when we have to teach it to someone else (students). Preparation and reacting to students. Folk wisdom – something powerful and important about trying to teach something to someone else. It is not only important for other person learns but its important for what you (the teacher) learns. Long standing interest – possibly a century ago or longer (in some countries) – relationship between mathematics knowledge and effectiveness as a teacher. A belief: a good teacher of mathematics knows a lot of mathematics. Knowledge extensive but flexible and deep. Compelling as an idea. Impossible to challenge it as a statement of belief. Difficult to establish empirically the relationship between mathematics knowledge and effectiveness of teaching. Going back to 1950s – syntheses of this work. Very weak association btw mathematic knowledge and teaching effectiveness in many studies (some small ones do, but most don’t). Counter-intuitive. FIRSTMATH study in same family tree as earlier studies in mathematics knowledge and teaching effectiveness. Broad scope – chance to produce empirical evidence to give us more leverage on this relationship. And, put it in broader context (policy environment, teacher prep territory, etc). Great deal that can be contributed by FIRSTMATH for people in education and math education. Policy-makers subscribe to notion that more mathematics is better than less. Often doesn’t yield a solution for policy-makers. I can talk from US context: in USA fair amount of skepticism about teacher preparation and its relationship to successful teaching. Studies done in recent years – try to track teachers from various teacher preparation pathways into first years of teaching – try to understand what happens to teachers who come from different pathways. In the US, math teachers have a high attrition rate, disproportionately high. Desirability of a math/sci background for other professional activities – have more choices. Still something in FIRSTMATH – may give us a clue, why teachers might seek alternative pathways. US – we don’t have a coherent organized system of teacher education. We have pre-service teacher education, we have informal pathways, etc. Not a
system of continuing professional education – random set of experiences. No
standardized/organized system…in USA no certainty what teacher will teach or do.
Teacher preparation people feel they have to do everything, cant be sure what will
happen subsequently. FIRSTMATH offers a sense to understand how system is organized
by looking internationally – from a US perspective, we have a lot to learn from this.

[Story]

FIRSTMATH – operates at levels of policy/system and level of teachers and their
knowledge. Lot of interest (recently) in research on mentoring and induction programs.
Professional introduction in teaching. New teachers are expected to do what experienced
teachers do. Expected to teach same schedule. In some countries, more of a gradual
introduction. When contrast with what happens with doctors and lawyers, expectations of
what people spend time doing when they start out. Something about profession of
teaching – inside of that, alternatives that might help us enrich the set of possibilities for
thinking about induction. Does vary across countries. Deep and long traditions of
different kinds of practices. Both at level of knowledge and how knowledge
development interacts with teaching context. One might presume knowledge would
grow/knowledge might dissipate. Problem – lack of coherence in US across preservice
and inservice dimension, and lack of coherence across school settings. Britton & Paine –
qualitative look at induction. Getting teachers on the job training, did teachers learn
something important. Question for us: whether scope of things to be studied in firstmath
– whether we’ll be likely to capture the most imp things that beginning teachers might
learn. Things that are missing. Would be interesting to think about – what’s missing. Is
there something more subtle, in a diff space, that might be important. Subject matter
knowledge and other knowledge.

-----

TT: Question, clear definition of entire population. How define unit, teachers? District,
more? Expand more, and have more teachers. How define the entity from which we’ll
draw teachers. Using representative samples makes this important. Use regional samples.
Question by funders: can we actually implement this study? Can we give them a
knowledge test. Can we find a rep sample of beginning teachers? How do we manage
to do this, how do we manage to inspire teachers and people to contribute to collect
data (like in TEDS-M).

-----

Lori Stratychuk

Methodological Considerations

Approach for creating surveys
Step 1: envision the final report (take a step back)
Goals

Primary goal (layman’s version): representative sample

What makes a representative sample?

Representativity (2 criteria)

First: Every member of the target pop has a non-zero chance of being selected (sampling)

Second: Every member of the target pop has the same chance of responding (operations)
To be discussed at a later date…

Frame – Optimal Qualities

Available (lists that exist, if have access to them. No duplicates)

Complete

No duplicates

Up-to-date (make sure current in schools)

Contains all desired information

Unique identifiers

Contact info

Coherent (if compiled from multiple sources)

Ideal Frame: an example

France has REGISTRY of teachers
   A complete list of every teacher

Contains both previous information (education) as well as current (employment)

List Frames

List of desired members of target population
Usually kept for other purposes
   e.g., professional associations, etc
Since they exist for other reasons, they:
Are not always in the desired format
Do not contain ….

May need to combine numerous list frames to make a single list (e.g., public and
catholic school system in Ontario. In addition to usual issues, there is also issue of
coherence between various lists.

Question (Chile) representative sample – every one has chance of being chosen. Can be changed, each member has a chance. Can change it.

Lori: People have diff chance of responding, you need to impose some different model. Calibrate things.

(Chile): Problem of having list of teachers. Sampling has two stages. Choose first school and build from that.

Lori: some countries don’t have a list of schools. This is where we’d have to select a certain area. Do extra work to make list of schools.

Question (Silver): How to think about – representative of what? If thinking about US, where in current economic situation the newer teachers are likely to be found (high frequency new teachers). Most likely to be school districts with high rates of turn over. Relatively few teacher vacancies. Economic situation led to repression in that regard. To get a rep sample of new teachers who find jobs – it isn’t rep of teachers teaching in US. May be diff in diff countries. Complicates comparisons, across countries.

Lori: Not going to be rep of US.

(Bulgaria): new teacher or new teachers in-service.

TT: Novice teachers, inservice, may define in different way (than TEDS). Diff responsibilities of a teacher. A novice teacher. Things will change, depending on countries. Defined as a teacher who has full responsibility for a class.

(Saudi Arabia): since novice teachers are from 0 – 5 years of experience, are we putting all one or two years in one group?

TT: It’s a cohort.

(Saudi Arabia): During sampling, have to keep this in mind.

TT: independent samples, those are cohorts.

(Saudi): What about those who graduate from school of education (but didn’t go to school immediately after graduation, working elsewhere).
Lori: student clients (don’t be greedy). Lovely to have teachers fit nicely into cohorts. Number crunching, may have to group 0 to 5. Some countries may not fit nicely into these. TEDS – pick people as finishing school. People have math degrees stopped being teachers. What happens to teacher who decide no longer be teachers? In this study, TEDS finished, we didn’t ask permission to continue study. FIRSTMATH is retrospective, current teachers and look back. In Canada know their credentials etc. teachers have a vested interest in knowing retrospective issues. Teachers who are currently teaching. Have to be beginning teachers.

(Pat): First five years of teaching, not first five years after graduation?

TT: first five years of teaching. Eligible to be certified and licensed to teach. People in US get license to teach (after 3 years) and then are beginning teacher. Then in this study they are okay. In US teachers are not paid in terms of coursework. They have no idea how many math courses or math ed courses…or even how long their student teaching was. Ask them to bring transcript, can get data. They don’t know it in their head.

Lori: how define novice teachers 0 – 5 years. Differences upon differences. In Canada teachers know educational background. See if teachers willing to provide transcripts.

(Pat): Its diff to say provide transcripts than to bring transcripts. Difficult to gain human subjects approval to gain transcripts. Could count credits/count courses they could check. Go home and take transcript with them. Some belief data entered was correct. Whereas just ask them in US (I’d have no belief they have it correct).

(Namita): Defining representative in a certain way, its useful if its common across all countries. If keep impact/end goals in mind, poss for representative to be diff defined across countries. In India, study has to be representative of novice teachers. Government…no way of tracking novice teachers. Practical reasons in particular states not representative of teachers across country.

Lori: this speaks to how good our list is. If have a bad list, you have a bad list. Ask principal or teachers themselves. They’ve been here for two years, teaching for several years at another school. Based on quality of list info available, may have to change practices. Study meant to be representative. What about qualitative stuff – one country they are nice to new teachers and another isn’t nice to new teachers. Want to put some thought into results. Aware of differences, not going to be rep of countries, will be rep of novice math teachers in that country and some non-rep in some regions/areas because of hiring practices..schools.

(Chile): it is not clear what is a novice teacher?

TT: developed proposal, all kinds of things we can do. Have to define your parameters. Including other teachers. Change character of study. Teachers in first, second, third, forth five… just going beyond. First we had minimal years. Five is where most people drop out of profession. Its when people decide to stay in teaching or to leave. Depending on what
find out, have very few teachers. Switch from six to six. Could include more teachers (what pat and tt talked about).

Lori: definition of novice teachers is up for discussion. Subject matter question. Comparing old teachers with new teachers: let’s envision final report. Valid to be collecting information on novice teachers.

(Silver): what are the advantages from a standpoint persp of thinking of doing a split of three years or less (three to five years). Is there likely to be detectable differences in a year by year stratified sample. Enough expectation and variability between two years and three years, etc., would that allow you to avoid some of the sampling problems (when consider timing).

Lori: create studies of just women and men. If you think just zero or one. Could consider these diff groups if think they are very different. Maybe zeros are different than ones. All things didn’t know…compare the really new teachers to not new teachers. Subject matter decisions.

TT: based on daughters reports, helped train new teachers. Amazing changes from time teacher began to teach from first six months to a year. The school year, daughters stopped reporting. Would be interesting what happened to person in second year. Teacher was at msu in internship year. Our teachers behave like second year teachers not like first year (principals say) they have benefit of learning on the job and students helping them.

Lori: Their may be some methodological repercussions…census, sampling. Based on research interests. Depend on final report.

(Saudi Arabia): depends on date/time conduct study. At end of year it will not be zero.

Lori: something to keep in mind. If teacher happens in January. Timing of surveys, each country has own calendar year, second year construction, also going to be little let’s keep these in mind. If differences unavoidable. We have to change this. Have reference point (in statistics Canada) always keep track of things done relative to different points.

TT: we will have to do first stage (two phases): beginning of school year. Survey people before finish school year. Study has to be implemented in one school year.

Lori: north to south in country, richer as head south. Weird results…if doing sampling, be aware of any of issues, doing it over eight months, make it be as random as possible. Randomly would like to go all over the country. Quantative and qualitative issues.
Target Population and Definitions

What is a “school”?
What is a “math” teacher?
   Teaches: only math, mostly math, some math, this year? Previous year?
   Previous five years?

What is a “novice” teacher?

Career Progression for Canadian Teachers (Olden days)
   Completion of studies to daycare employment to supply teaching to term position to full-time permanent position.

Frame contents

If (school) lists have information available for both: math teacher/not, novice/not

If (school) lists don’t…

Review of landscape

Why questions were asked
   How responses will affect the viability of the project

Landscape: Sources

   Allows coordinating staff to review information without always having to contact each country…

Landscape: List (Frame)

Determine:
   Availability
   How current
   ETC

Landscape: Anticipating the future

   Need to know not only the number of novice math teachers now, but will need to know attrition, growth..

Need to compare “apples to apples”

Never assume anything
Landscape: ISCED/Grades/Ages
May focus in on certain grades/ages to enhance comparability issues

Career Progression
Definition of “novice” may be influenced by “natural” career progression in the country.
Example of Canada, pattern over time.

TT: Powerpoints will be posted on website.

Lunch Break

Kiril Bankov: FIRSTMATH Mathematics and related teaching knowledge of primary and secondary school teachers.

How methodology can be implemented for FIRSTMATH.

FIRSTMATH builds on the hard work of TEDS-M
TEDS-M studied the level and depth of mathematics and related teaching knowledge attained by future teachers…

Links between two studies, research question corresponds between TEDS and FIRSTMATH.

Mathematics and related teaching knowledge – bridge between TEDS and FIRSTMATH

Shulman’s Conceptualization of Knowledge for Teaching – 3 dimensions (subject matter knowledge—all knowledge of future teachers); (curricular knowledge); (pedagogical content knowledge).

Revised for TEDS-M – Mathematics Content Knowledge (MCK) and Mathematics Pedagogical Content Knowledge (MPCK).

Classification of Mathematics Content Knowledge –

Subject matter domain (adapted from TIMSS)

Cognitive Domain

Curricular Level
MCK Content domains and topics (number, algebra, geometry, data)

MCK Cognitive subdomains and processes (knowing, applying, reasoning)

Curricular levels (Novice, intermediate, advanced)

Question (Saudi Arabia): can you propose one example of what you mean by “intermediate”?

Kiril: Imagine primary teachers. TEDS-M assess those who will be primary teachers. Expected to teach from grades one to four. Go to novice. Items recommended for five and six go to intermediate. For grades 9 and so on.

TT: primary teachers prepared to teach 8. Knowledge to teach curriculum that’s classified as advanced.

Silver: There’s no parallel notion of understanding prior grids. Upper secondary teachers, there is no advance that involves understanding of topics teaching in grades five and six, which is the symmetric version of this.

Kiril: for some teachers yes, other countries not. More or less problematic, in terms of country (classification).

Categorizing Mathematics Pedagogical Content Knowledge

Sub-domains (curriculum, planning, enacting)

Content (numbers and operations, etc)

Mathematical Curricular Knowledge

Establishing appropriate learning goals, knowing different assessment formats, selecting possible pathways and seeing connections within the curriculum, identifying the key ideas, knowledge of math curriculum

Knowledge of Planning for Math teaching and learning (pre-active)

Planning or selecting appropriate activity, choosing assessment formats, predicting typical students’ responses, including misconceptions, planning appropriate methods for representing mathematical ideas…

Enacting mathematics for teaching and learning (interactive)

Analyzing or evaluating students’ solutions or arguments, analyzing the content of students’ questions or responses…

Item formats
Multiple choice items (MC)
Complex multiple choice items (CMC)
Constructed response items (CR): completely correct, partially correct, incorrect.

Primary items by domains and blocks: five blocks

   Domain: algebra, geometry, number, data, curriculum and planning, enacting.
   124 items in test for TEDS-M.

TT: when we were designing this test and the items, we did curriculum analysis quickly, then a more detailed one. We tried to see how domains rep in curr at primary and secondary levels. Can see higher items in curr. You will notice in data, there are very very few items. Items decided for data couldn’t work with. Interesting finding of study, couldn’t come up with good items. Teacher educators best for coming up with items, we couldn’t figure out very well how to design items.

(chile): you have repeated the items?

Kiril: you will see some of the released items, some are kept confidential.

Chile: You will have three groups of teachers?

Kiril: for teds m, two groups.

TT: blocks in tests to cover domains.

Kiril: two blocks arranged in three diff booklets. Each student answers only one. Complete block design.

TEDS-M Released Items

   Primary (24 MCK Items, 10 MPCK items)
   Secondary (23 MCK items, 9 MPCK items)

Items for FIRSTMATH

   Released items from TEDS-M.
   Newly developed items
   Items proposed by the participating countries

Example Item – TEDS-M Primary Level

Example Item – TEDS-M Primary Level

Example Item – TEDS-M Primary Level

Example Item – TEDS-M Primary Level (MPCK)
New Item – Primary Level
New Item – Primary Level
New Item – Primary Level (MPCK)
New Item – Primary Level

Example Item – TEDS-M – Secondary Level (Geometry)
Example Item – TEDS-M – Secondary Level (Constructing a proof)
Example Item – TEDS-M – Secondary Level (MPCK)
Example Item – TEDS-M – Secondary Level (Rational Numbers)

New Item – Secondary Level (MPCK)
New Item – Secondary Level (MCK) (Graph, option of whether options represented by graph)
New Item– Secondary Level (MPCK) (geometry).
New Item – Secondary Level (MCK/MPCK) (geometry)

Item and Instrument Development

The processes for item development that were used in TEDS-M will be used to design the FIRSTMATH items…

Item development

In firstmath we will design items that each country expert team will contribute…

Field Trial

Then questionnaires will be piloted on what is called “the field trail”
During these processes, some items may be eliminated or slightly changed
Items remain will become part of the final study instruments.

Question (Italy): general consideration: if this study has to do with how teachers teach math or how do they teach as teachers? From the items, esp pedagogical content knowledge and mathematical content knowledge. As teachers or as math teachers. How teachers teach in first five years, should include how they are disciplined, but also more general pedagogical competences. In Italy, beginning teachers: teachers tell them their issues are more general pedagogical problems not a third dimension. How teach as teachers in first five years.
TT: we tried in teds m, had full set of questions, never got them to work. There is a questionnaire, ten or nine countries, we didn’t get it to work well. Beatrice Avalos published book about these findings…we couldn’t meet the strict standards of IEA for how items should behave, they are mostly open ended. Was a very high area. Problem of definition – general pedagogy. Was extremely difficult to measure. Decide to concentrate on mathematics and mathematics pedagogy. What teds has contributed is giving a common language. For general pedagogy we need to know the same work. We’re still waiting to see results of general pedagogy. We’re trying to approach this in a new way here.

Kiril: what we did at beginning with TEDS-M for general pedagogy I liked a lot, sorry it didn’t work. There wasn’t a commonly understood as good pedagogical activities across countries. Not acceptable in all contexts.

South Africa: would firstly the primary and secondary teachers, probably two different (instrument). Within that there’s a general set of questions about maths knowledge, depending on which one you categorize at one of three levels. How would you categorize….someone teaching grade three and grade seven, they are able to master content up to a certain grade? Is that logic?

Kiril: Item three common items, not there for some comparison among items. Teds m no intention to compare these levels.

TT: separating primary and secondary

South Africa: If somebody can answer content only up till level teach, if two grades higher, how teach.

Kiril: using categories for item development, controversial. Any items can be justified different ways. Just for preliminary purposes. What about item selection…

Thailand: use release items?

TT: not use,

Pat: is the breakdown for firstmath for primary and secondary, where does the grade stop for primary and the next grade starts for secondary. That seems to not be consistent. A sixth grade teacher would be able to do item but not a second grade teacher. Where do the grades break? Two groupings (primary/secondary), where does break between two happen?

TT: your question makes me think. Don’t know yet how countries configured. At end, end up with six diff groups. Four groups for primary, first two groups for lower primary and other two groups for higher primary and two groups for lower secondary….if we were going to think of these six groups as parallel to something that exists out in field. Reason why came up with six group, that’s how they think of knowledge teachers have. Conceive of lower primary teacher. Conceive of higher secondary teacher. In lower primary teacher they conceive of two different levels. And a higher level, then they higher primary…..we may try to think of those six groups,
where primary begins, ends, etc. we did those groups after we understood complexity of those systems. A test for primary teachers and secondary teachers. Division of groups, opportunities to learn. Where teacher educ programs…tests measure outcomes. Could say test in lower is fine. Way to resolve is have each grade (have) as primary, mid, secondary. Very few beginning teachers. If only look for beginning teachers teach eighth grade, we aren’t going to have enough people. At what point do you apply the secondary test/the primary test. I don’t think we want to design six diff tests for six diff groups.

Pat: do you need to span the entire spectrum, should firstmath should not go from grade one to grade twelve. Should firstmath only look at fourth and fifth grade. Maybe you do it by content. Let’s define it by content children are taught. Related content (num, alg, geo), seems to be content taught to eight, nine, and ten year olds. In US it’s three four and five. In another country maybe its diff grades. You do it by content. Content taught in same cluster. Teachers teaching higher level of content. You got points you could be doing. If just do primary/secondary, challenge to think of teacher of first graders and teacher of fifth graders. What kind of math teacher teaches. We could be messing up connection Ed was talking about, OTL from colleagues and from prof dev and from own thinking and reflection. And not if you’re asking me questions that I don’t know. Knowledge test score would be lower, but your pedagogical knowledge test score would be higher. The items they got aren’t related to what they’ve been teaching for last two years. Maybe we shouldn’t think about whole spectrum. Given what you learned in TEDS about six groupings, why go back to two?

Kiril: we have to think about this. We don’t have answers at moment.

TT: tomorrow Mark Reckase comes. He may have good ideas about this, measurement expert.

Namita (India): Planning to test teachers, same as students.

TT: did it in sri lanka. We would like to be able to use the tables timss develops and measuring knowledge. Question came up by reviewers of proposal: can you measure for students. Many logistic questions here. time teachers were answering tests students were answering tests. Some concern in research team that this could draw the energy out of FIRSTMATH. Now thinking about how test students. Tests would only be for teachers who selected in sample. Mark Reckase would ask: why are you asking? What’s your question? Same test, what learning from that? Serious question from funders, why aren’t making that link.

Sandra: OTL questionnaires decide….in addition to asking teachers to answer questions, you could have opportunities to teach kinds of questions. If haven’t taught it, have that as a resource might be helpful. Surveying in opportunity to teach framework. Issue of grade-level issue.

Wendy: Small bit of experience with computer adaptive tests. Instead of making cutoff in primary/secondary what if have one test and put people at diff starting points. Depending on whether getting questions right or wrong, moves them up or down. Measure knowledge further on, perhaps we could do one big test, no one takes all items.
TT: terrific, I would love to do that. Problem: whether we should do computers to do tests. Another country have access? Bring teachers to one room to answer? Control testing situation? Brings a lot of logistic issues. Studies move things forward. Tests should be the next step.

NEA: what year looking to do that? By 2014 in USA, the US students are going to have to take standardized tests to the core on computers. There should be computers in place to address concern. That’s happening right now, that we hope they’re funded, setup situations in schools where might have equipment, connectivity, can test in schools. It could be by time your ready to do testing that it will be dressed.

TT: 2014 is a possibility if we work really hard.

South Africa: two points. One is whether we test students and my response, what research questions in respect to this study? What question you want to answer. Lori said don’t be greedy. Second point: instrument about maths teacher knowledge. Does force us to make some demarcation which group of teachers want to look at. Paper and pencil instrument. Computer may not be applicable in all countries. Paper and pencil test and take a teacher that teaches grade one, versus a teacher at grade eight, it doesn’t seem to be a fair instrument to give content knowledge to all teachers. As design goes forwards, make decision about cutoff point.

TT: computer test, what the IEA has done…what we’re trying to do is convince IEA this would be a good idea to undertake. IEA has in the past developed same test for computers and paper/pencil. Comparison see how works. Differently, these are beginning teachers…assume have knowledge to teach for primary teachers. If you’re trained as primary teacher expected to teach to grade six (in Mexico). Assumption is badly placed. Revise assumptions. Depends on assumptions countries have. Mexico it’s from seven to nine. A lot of teachers in secondary schools, expected to teach to high school knowledge. From middle school to primary school supposed to teach these…

Silver: I think the issue people are raising: while that’s true, that was a reasonable assertion to make (TEDS) – its all fare game to throw math at you. Point others have made, by the time you look at a fifth year teacher if they’ve spent five years teaching first grade, feels a little different throwing them upper level content. While they’ve all had opportunity to learn it initially, have had more opportunity to learn, work on, dig into mathematics, at particular level they’re teaching, diff sort of testing situation. Diff sort of game, we don’t know, how much of a shock to this system this could be…all of a sudden ask to answer questions that’s well beyond what they learn. It seems we already would’ve expected that. If there are other choices, if test could be designed to be more grade level experienced. Maybe there would be more … more narrow content band rather than broader band.

TT: TEDS driven by curriculum of teacher education programs. Test has to be driven by common core standards.

Silver: This is an argument, if what we want to learn, what mathematics and other related knowledge teachers acquire on the job. Chances they have to learn on the job are more intimately
connected to mathematics that they’re teaching. Question is what will we learn about the question that ….

TT: what they bring with them and what learn on job (two parts)

Silver: if you look at that teacher in strata six, the foggier novice teacher, you don’t know what they bring. You will guess what they brought. What learn in teds-m, presumably had a lot of variability. You don’t know what that teacher brought. You will know what that teacher has been doing, wont know if you had testing them once they had left their program. Will have something like that, earlier you test the more you’ll have of that, the later the less you’ll have. If in fact, people are influenced by experiences teaching.

Toni: Wide grade band to have a test that’s represented of bands. Learn more about teachers of fifth grade. Learn less about PCK. Seems like a wide grade band.

Pat: you kind of are aligning items from grades. Items connected to what is taught. How do teachers with five years of experience. Assumption want to investigate, does teaching matter. That will get confounded by the items. Tested on many items about content they learned five or six years ago. Disadvantaged teacher who is fifth year teacher without classroom experiences, not alike, fifth year teacher who does have five years of classroom experiences. Compounded by domain, narrow that domain to be more closely aligned, then you have a better measure. Just raising as a concern. Maybe you don’t want to do entire span. Maybe want to do splits. Maybe gaps in between, wont have teachers in there as a purpose.

David Evans (WB): these points very interesting, particularly points if data looked like in particular ways. We assume all is the same. If we don’t have any data on what teachers are teaching. Seeing differences in content across years, esp given variation across countries. A fifth year teacher spent a lot of time or less time teaching. If narrow it, and try to test what? What teacher teaching last two years. Broader view – gives us a broad array. If they’re still certified to teach across sixth year. As long as test isn’t ten items long, we’ve learned how much they learn of first grade content, second grade content. Definitely a lot of value of testing teachers across the years.

Jouni: Why we’re speaking. Not comparing individual teachers…not on individual level. We’re trying to measure potential group of teachers has. And of course variation. What is there opportunity to teach, opportunity to learn, make some analysis. Makes it a bit complicated. I don’t see a crucial problem here. its interesting to know, if it matters if you have some experience or not. Lots of group, covering a lot. Need more teachers for that. Im worried about sampling design, six different cohorts, because it makes it very complicated. We need a lot of teachers for that. Do we really need that? Then the sampling, we have to know how many years, and on what level they’ve taught.

TT: I’ve had this conversation with Jean Dumais/ and Lori, he says you get teachers…forget about cohorts, just get beginning teachers, have them all there, ask them question, less complicated.

-----

Coffee Break

----

Use of Videos and Questionnaires (Joe Tobin)

Possible Uses of Videos in Firstmath: joetobin.net

TT: Introduction, work on videos, to learn about blackbox in classrooms, just get right in there and understand classroom dynamics. Preschool in three cultures.

Plan for presentation:

Not a math educator
Some thoughts on an ethnographic component to study
Four ways to use videos in this study
Technical issues
Two examples of classroom videos
Discussion

Ethnographic method

Method: Video-Cued Multivocal Ethnography (the “PSin3C” Method)

Video as a tool both to show what goes on in classrooms and to elicit teachers’ thinking behind their practices.

Show what goes on in classrooms, and to elicit thinking behind the practices of teachers. We don’t see teachers pedagogy in video, we see behaviors recorded on screen. The thoughtfulness, planning and thinking. Watch a decontextualized question. Want to combine video with interviews with teachers. Get them to talk about the things that they did. Might respect them more if we heard the thoughtfulness behind the approach.

Pre-School in Three Cultures – videos as cues. And sequel.


Chinese preschool teachers emphasize say living how to give and receive constructed feedback what get in school. Video can lead to rich discussion. Showing to other teachers in other countries.

Ethnographic Approaches

View beginning teachers as informants rather than as research subjects or Ns. They know a lot about what is helping them and holding them back as effective teachers.

If one of our main research questions is what is helping new teachers be effective or keeping them from being effective (that is, induction), we can ask them in this study…

Ethnographic questions to ask new teachers about:

- Preparation, school context, inservice professional development, career and life. Understand more going on than just pedagogical content knowledge…etc.

Silver: how universal the framing of questions you’ve raised? Career and life balance issue is same across cultural boundaries?

Joe: begin with question without prior assumptions. Ask how’s it going for you as a new teacher. Then start to introduce questions that follow the hypotheses and assumptions of the study. If you ask the question…large study need to do open-ended ethnographic question. Unstructured anthropological world vs. structured comparative world. How many initial assumptions that guide study come out of a North American/Anglophone perspective. Important for people in other countries and be open to rethinking everything.

Video-cued ethnographic interviewing

Participant observation. Show video ask it and stop it. Why didn’t you answer that. Way of doing participant observation with video.

What’s the best way to help students understand how to add fractions with different denominators? You can show them a fourth grade teacher teaching about adding fractions in a certain way and ask them, what do you think about that approach?

What you’re after, what teachers are thinking with pedagogical moments.

Four approaches that might be used:

1. case studies (longitudinal) begs for falling people across time. Could be a subset over a longer period of time. Seen from first year to fifth year…what learn or didn’t learn in teacher preparation program. Have three math teachers at these level. Could combine thicker case study data with big statistical data to have it makes sense, breathe life into it.
If have teachers videotaping selves teaching, a friend could tape it, could ask teacher what they’re going to teach. Afterwards could do a debrief. Planning to do a review. Without that introduction and debrief you lose thoughtfulness.

2. interviewing cue – could edit a lesson down (5-10 min) and shown to beginning teachers in each country and used as cues for eliciting reactions about their beliefs about effective mathematics teaching. This can be qualitative, quantitative, or both. Videos could be put on a website and beginning teachers asked to respond to an online questionnaire. Videos could be put on a website and beginning teachers asked to respond.

3. visual examples – what looks like in each country (math teaching), at three grade levels. These videos will be useful for providing concrete examples that can ground discussion of key issues across the national sites. They can also be used in presentations of findings from the study. And the can be re-edited for use in teacher preparation. Question: would videos of beginning teachers look very different than videos of experienced teachers in each country or will national differences push aside differences between new and experienced teachers?

4. observational analyses of enacted practices – could be qualitative or quantitative analyses of teachers’ classroom practices. The question is whether such observational analyses should be done by a classroom observer or using video. Video is good for microanalysis but takes a lot more time and money. Is repeated viewing needed to make judgments or can it be done in real time?

**Technical Issues:**

One camera or two?

Fixed camera position and focus or moving?

Audio recording

Whole lesson or edited

Made by video expt, researcher, or practitioner

Formats for capturing, editing, sharing.

Examples: Germany Vignette (scripted classroom lesson)  
Saudi Arabia (not edited or scripted)

Can see differences, two extremes. Problem with one camera shoot vs. multicamera shoot. Subtitles are an issue.

Videos are time-consuming and potentially expensive.
TT: Very few things have been decided. If people think we want to all embark on some video project, there are a number of choices. Video? Yes or no. series of options of many uses of video. 9 per country. Nine teachers per country, select them. Representative sample.

DISCUSSION ON VIDEO

Wendy: adding videos along with interviews with teachers. Could add valuable data cant capture just in surveys. We are videotaping right now some lower primary teachers, not beginning teachers, but teachers becoming math specialists, exploring how to efficiently analyze hundreds of video tape. We have fifty video tapes, we will be getting more. Asked all teachers who have been involved, they’re supposed to videotape twice in fall and twice in spring. Loaned flip cams with tiny tripods. Out of thirty five teachers got twelve or thirteen different formats.

Joe: could make a couple of template videos, videos of fourth grade twelfth grade lessons and those would be come style everyone would use.

South Africa: presentation has two dimensions. Supplement survey. Other aspect: how use technology like videotaping for analysis. Talk about that? Timss study if one used videotaping as technology, if analysis included in there, could include.

Joe: there are a lot of uses of video in research. Video in learning sciences recently came out. 1) first question is why video? If video is used for observational analysis, then specific software can use. Lot of software packages to use coding, atlas TI with video. Some of qualitative stuff. Some dedicated video ones. Can put codes tied into segments of video. Helps for doing analyses and for data recovery. So have big archive of videos, see all times teachers dealt with quadradic equations, wait time, posing of questions, develop categories you want. Timss videos were used more to show examples of classroom practices that could be intuitively correlated. Timms videos could show you something about teaching strategies. Jim stigler dev sophisticated video software.

Namita: described different uses of video. Kinds of video you would need across different types. We would require different videos of different classrooms.

Joe: maybe is the answer. It could be done in diff ways. 1) case studies of teachers in each country. E.g., three fourth, sixth, and eighth grade teachers in each country. Have protocols for how videos are made. After one month, eight months. How define it. We’d have nine videos. 2) or have videos scripted or edited that would be a list…designed in the form of a question. Used to elicit teachers thinking, watch videos and respond to them. We might have the same lesson acted out by actors in different countries. We could have the same lesson from one country subtilted.

In my own research, use video as cue, then re-edit it, same video footage…when ask teacher why didn’t intervene, robs children of figuring it out on own. Used for curriculum later.

Silver: from the point of view of the firstmath project roles, the case study use of video seems very compelling as a complement for what can be used in surveys, etc. I like this idea of video-
cued…getting good items to assess pedagogical content knowledge. This technique might be a way to get at pedagogical content knowledge at a diff way rather than having it written in a survey. Haven’t had much success writing like this. One challenges, it would be hard to have the same videoclip play well across countries. Another way – go to neutral format. People who have done work on cartoon presentation, characters portrayed not specific enough to know if American or not, can imagine any classroom where kids are configured in classroom. Chasen and herbst have done a lot around teaching of geometry (around this), practical rationality around teaching of geometry….cartoon like format might be easier to edit than doing it across countries. Whatever expressions teachers use across countries. Use whatever display. Can experiment with items, get at items that are hard to test. Imagine you could do some interesting things, measuring PCK in video cued format. Areas complement stuff that’s already there in TEDS-M.

Pat: when you write paper/pencil pedagogical content items. You struggle finding out what right answer should be. What is the intended answer that gets full credit. How document that that should’ve been the intended answer. Maybe that’s not the question if you do something like this: how do diff countries interpret situation as to what intended action or interpretation should be? Are there differences across countries regarding if a teacher should of done something or what should do? Open it up to a different kind of analysis. Challenge of writing pedagogical items…can dream them up, but cant figure out right answer. This would be for a different reason. How different nations interpret these things differently.

Joe: two diff uses. Video items as opposed to pencil and paper. Video is more evocative rich contextual cue than a hypothetical situation. Other possibility getting inside heads of teachers. Cue to get them to talk about their own journey through induction. It would be good to have case studies beginning teachers what their struggles are. One use: pedagogical beliefs and knowledge through video. Treating teachers as informants through own struggles. With or without video, thinking about teachers we should ask them. Half way through first year, how’s it going for you? It could be relatively small, it doesn’t have to be that expensive. Seems odd not to do that.

TT: how many people think this is possible to do in own countries/contexts?

Brazil: acceptability of it…with or without camera. More difficult to get in with camera. Could go into friendly schools, it will be so interesting. In our case, they don’t agree much.

Joe: in brazil, do teachers make videos as part of teacher training?

Brazil: not a current practice.

Joe: harder and harder to get into classrooms, concerns about human subjects, children images on internet…cameras are all over. Harder for researchers but more common on youtube. Teachers motivated to do it as something else. Help us as researchers.

Brazil: if it comes from a particular school, its well accepted, state secretariat it becomes more difficult. See what is going wrong. Teachers who are struggling less likely who are struggling than those who think they’re very good.
Sandra: one quick point – one way of approaching teachers teaching. Have teachers conduct interviews, esp if given script for interviewing. See kids in classrooms, watching teachers conduct interviews can be helpful.

Joe: practical problems – use cartoon version (wouldn’t be hard to do permissions), case studies of teachers and struggles, could be done without video. If goal is to interview teachers as they go through inductions overtime. Wouldn’t require videotaping in classrooms. One is more case study of induction, other is videocued data collecting.

South Africa: the video itself as an analytic tool.

Joe: want to measure something about teachers, beg and end of year. Get into sampling purposes. Better to use video or observational analyses, then don’t have technical issues.

Chile: we use video with variation of teachers. Use for evaluation. Somebody videotapes them. Required to be videotaped. Inservice teachers videotaped. Each year videotaped. Use video for research, project with Finland, using video. Use video.

Joe: parts of china – teachers required to bring video of lesson and present it to other teachers/experts. Everyone criticizes! Teachers would do a demonstration lesson. Some are moving to bring into a video. National variation in acceptability of video.

Jouni: it’s possible to use video for research purposes. We have to know the way and purpose how we use it. If we make that kind of material, commercial purposes, we must be sure of it in advance. Have to ask students if old enough, have to get permission from parents, detailed info for purposes. They have to know. I don’t know all those legal issues. People upcoming, much more sensitive now a days. Student would’ve been happy because of this internet problem. People becoming sensitive today, may cause some difficulties today. Shouldn’t be impossible.

Saudi Arabia: There’s no rules about these things. In case of video, I personally asked teacher if I could videotaped, and he agreed. Good thing we don’t have to go through all these things, and writing things.

TT: If MSU is lead center, we would need IRB approval, we would have to show teachers some forms, very formal. And people can say no, and that’s it.

Italy: I’ve used informal contacts, and received permission from parents contacts. Had permission from head teacher and informal contacts are easier. Easier to get inside through informal contacts then coming from above.

TT: we have irb, videos not seen as real risky as long as people agree. We are exempt. Our review board that is. To make the videos as form of data collection for representative sample, in some places that might not be possible. In brazil, people would not agree to do this.

Brazil: to do it systematically difficult.
South Africa: go through same roots of ethnics clearance (ministry of education), one can undertake it, have to fill in a lot of forms of what do with video.

Thailand: have to have personal contacts in school, have an honesty form to sign.

India: ethnographic case study, extremely useful. Something we would like to do. In general, purely quantitative studies only go so far...we can get under skin. Video part – video cued interviewed testing sounds interesting, really would have to be placed in context with whatever alternatives you have. Video-cued...you give a qualitative example and then ask them. I don’t know if that’s a useful route. In general I don’t think its difficult to get into classrooms and do video. Something exciting and useful for organization. Engage in project we will learn how to use it. I don’t think its impossible to do. We did a lesson video for this. Through informal contacts, get in nine classes, videotaping apple, and route to get there.

TT: what about other regions in brazil, same case, difficult?

Brazil: a little bit. Can have problems with public attorneys, have a movie. Its only nine or ten teachers, it might be okay. Its possible.

Joe: Not desirable to make all teachers in study. Should there be a subset of study that would be case studies and would videos be useful. If countries are going to participate in case studies and qualitative measures. People in big data sets aren’t good with case studies and vice versa. Mixed methods is tough to get in study, two diff paradigms. Able to bring in ethnographic, could come up with explanations with surprising ethnographic findings...some of the countries could do some of the case studies. Get people who have experience with case studies/interviews.

Honduras: start to use Japanese methodology in 1990. Gov’t took videos of teachers using that methodology, lot of material of this. We have material of teachers in preservice induction (videos).

Bulgaria: very attractive and interesting (video). If theres another person, not licensed, another licensed. Its not real.

Joe: maybe too early to make decisions, but we thought we’d do it early to do some thinking.

-----

TT: Barbara Bruns presentation coming up.

-----

Short break
Barbara Bruns (World Bank)


Brazil – teacher incentive reforms. Controversy around bonus pay around world. Brazil – quality and quantity of data on outcomes available. You can know what learning happens with standardized measure across country, what children learning. Wanted to understand channel through which incentives operated...how incentives changed teachers. How incentives induced subject to do something differently. Underresearched, how incentives change behavior at school level. How measure how teachers behaving at school level. Commissioned colleague at Harvard.

What makes teachers effective in the classroom? What do we know: having material in classroom not enough, computers not always enough, having a teacher in classroom is not always enough. Education is process of transforming all inputs into learning. All investment in education is wasted. Depends on final transformation of inputs …

Know from research in US – teacher effectiveness varies enormously. Huge variations, some cover half the career curriculum. Variety in how effectively teachers teach. Little data to support that intuitive observation. Scholars have started quantifying this. Across subjects, across grades, etc.

Every economic study correlated student learning result...weak correlations. Years of teaching experience, first few years of teaching experience. Spending levels also. Cannot find correlation between what is spent on education and what students learning.


Classroom- ten observations at each class hour, coded on standardized grid. Ten observations of each class taken at regular intervals and coded using a standardized grid.

Data get – teachers use of instructional time, student/teachers...engagement. Share of class time students are visibly not engaged.

What are some of the advantages...teaching methods. You can code for question/answer method. A lot of things it doesn’t capture. Advantages it has, of things do observe, quite high reliability, applied in OECD countries and developing countries. Easy to mount data collection round. We’re finding more and more evidence of robustness of approach…

Coding grid... various teaching activities listed

With each class, ten snapshot observation. Get basic demographic information, availability of materials.
Teacher is anonymous, no record made. Know grade and subject, but don’t know them. Aren’t interested in individual teacher, interested in representative snap shot of how system is performing on any single day.

Classroom observations in various places throughout Brazil…

Results: five main conclusions

1) Brazil is far from OECD benchmarks (85 percent of every class hour should involve learning activities…brazil has no more than 65 percent). Teacher shouldn’t be off task. Teacher should be on task. Learning time should be spent on task.

Teacher arriving late or leaving early…teacher has to be in class room for learning to take place.

2) strong correlation between teachers use of time…more time on learning is highly correlated with learning outcomes.

Germany: individual or school level? More about culture?

We also have classroom level data. Cultural difference here. what is socioeconomic status of these schools? Potential circularity in causality. Seems to be a link between timeline tasks.

Results statistically significant.

Correlation between use of time in schools…observations taken before one year tests. Captured what was going on in schools at moment for determining achieving targets.

3) more time on learning predicted schools’….

4) teacher practice varies greatly across schools.

5) and within schools. Large variance in practice across different classrooms in the same school.

East Asia – culture of observation more typical. Expectation teachers work will be viewed and shared. Don’t see this in USA or Latin America mostly.

New results from Mexico (DF): teacher time on task in math classes is 62 percent, slightly above average (60 percent). But imp diff across grades.

Math teachers are most likely to use blackboard (in Mexico). Math teachers are more likely to have large group students off-task (five or more). Share of tiem students are involved or visibly not engaged.

Summing up – stallings method classroom observations potentially powerful tool…started with research interest. Operational: benchmarking sys performance.
Identifying where most excellent teaching happening. Secretary of education has no way of getting objective, systematic data. Can pinpoint classrooms where students are never off class. Sixty percent of time students not engaged.

Gives statistical platform for finding where best practices are.

Bring together best practices, how to do things. Work with the system to get practices noted can improve performances.

Give statistical platforms. Why are those teachers doing such a great job. Easy to go back and videotape them. Test math competency. What makes teaching effective in classroom. Teachers with best classroom management and delivery mechanisms.

TT: what are your thoughts/questions misgivings or whatever….what do people think of this as a method for data collection for our research?

Mexico: Implementing this process of classroom observations. Build capacity doing supervision. Good results in mexico, have three hundred in class supervisors. For one thousand five hundred public schools. They were trained and what they said, they learned how to engage in classroom once they go inside for close surveys. To get into a classroom at least one hour has change in dramatic ways they do observations. Supervisor got big surprises. They found teachers who have routinary practices. Found teachers who aren’t able to control groups. But they also found teachers with big knowledge and can manage groups. We now can say we have good evidence of what is the good and what is the bad practices that are are not delivering results. Make note, this is a demanding job for the supervisor. But from my position for my person in charge of educ in mexico city, I am convinced this is a very good tool to get into heart of system which are the classrooms. This is a good tool to focus on. Teaching specific math or other courses.

TT: divide classes into ten sections, people do snapshot, do people also do verbatim of what’s going on, how does he have data on content teachers are having.

Barbara Burns: Don’t get data about content that teachers teaching. It’s really about the methods that are used and materials used. That’s the big trade off, in order to evaluate what kind of material in qualitative way…we ask observer to write down what s see. We have ability to go back and audit sample of sheets and no inconsistencies.

Mexico: in case of mexico, in demographic, we ask of teachers – what is the perception of holdings taught during this classroom. That is beyond the method. In case of mexico, there are supervisors familiar with what should be taught, also have an idea. Important to realize to be in classroom at least one hour. Study records what’s happening. That’s independent of study recording.

Barbara: feedback we’ve gotten every place we’ve done this, by third day of training, I’ve never seen things before the same way I see them now. Not really in a disciplined way collect data.
Mexico: in the case of Mexico, there are no formal hours and he changes subjects as he wishes. In secondary there’s a tiered scheduled. In secondary you can clearly say, observe maths or sciences. In primary way they did it, has to go to a classroom and see what is happening. They know that eighty percent of classes in primary math were in Spanish.

We face a challenge when we have these changing subjects in primary. We modified which subject teacher was teaching. The method allows us to check how much time was lost when changes done by teacher.

Saudi Arabia: how many inside visits by supervisor and who observing?

Barbara: observers are usually supervisors in the education system. Someone who plays a role in system. This is a technique that can enrich those peoples execution of their normal job. The protocol they use…each school try to observe six. Observe six diff classes during that day. Sit through six hours of classes non-stop and make these obligations. They observe each teacher once. The objective is to get a snapshot of entire system.

Chile: in chile want to do same thing…we didn’t have money. We asked ministry of education about relation with teachers. Each video with teacher, class of forty five minutes. Took randomly some piece of video and observe each kind of thing.

Italy: observation of whole sample or some non-rep sample of teachers as a complementary view and see what they are doing in diff areas. Not just see how respond in certain areas. How do they behave in different issues. Don’t know if possible to do this as a rep sample or a small number of teachers.

Barbara: we try to do a sample that’s representative and that captures changes in variable in system over time.

Italy: if we want to have the full picture, this could be a good complementary method. Videos and interviews would help understand PCK…teachers will not talk about PCK or how teach math, or talk about general issues how control classrooms, their career, their salaries. In a way we have a lot of choices…

Namita (India): discussing classroom observation…way of deepening understanding what novice teachers do in classrooms. Question: is it possible to get more disaggregated data? Observation: large scale study in India that shows that teacher absence is not influenced at all by any sort of incentives. The only variable we can say, if varies across teachers (novices and veteran). As a policy tool those variables are useful.

Barbara: out of school tutoring of kids? I want to understand how behavior is changed as well as incentive. That result isn’t too representative of how teachers behave…

Namita: disaggregated data rather than splitting time on tasks?
Barbara: yes, we could look at methods, we haven’t had time to look at data. Teachers in their first five years. There’s a lot we could do we haven’t been able to do yet. Its all about observable interaction, how well keep students engaged, can get a lot of illumination.

Pat: if you wanted to do observations, want to think about, are there aspects of teachers and structural moves. Instructional moves that keep kids engaged in academic pursuit, whether teacher is present or not, other kinds of things you might want to think about peculiar to mathematics. Did work in Baltimore city, in some early visits, one of big issues, we need to target our intial professional development at teachers. Shifting notions of what it meant to teach math in big ways. Later we would worry about details. Spending an awful lot of math time with children copying…wasting time with copying. How to have mathematics instruction, everyone gets bored with copying. They weren’t learning anything. Lots of classroom disruption, hard to keep kids on task. Once you get beyond issues where there’s no teaching going on in room. In a lot of places, there is teaching ogin on. What is it about math instruction want to say components and say if present or not. Not getting too fine grained and not getting into lots of nuances. When go to observe a math instruction. Kind of mundane change way principals look at math. Teachers ought to look at a few problems working on. Should not demonstrate how to solve problem, they should ask kids. They couldn’t tell if it was a good or bad question, but they could tell it was a question, when child gives an answer, teacher should say: how did you do that? That was it. It wasn’t elaborate. Helped teacher and principal focus on that. Second or third grade tteacher in same room and get thinking about it. In forth or fifth grade room, how did you do it, how do you know you’re right? Begin wanting to make sure getting instruction right. We want to move beyond that. We don’t want to make it too nuanced, how good of a question to pursue childrens understanding. May not be possible, wont know how good data is. Is there a simpler way of doing this. Is there simpler ways to do observations, that could give us a little more info about mathematics. Want to know on task, if there’s instruction, we could probably go a little further.

Joe: are those things you described, effective strategies. True universally?

Pat: just an example of this city, what they wanted to focus on. They wanted to change math instruction. Teacher wanted to listen to what kids thought. Lets get kids talking about math, and teachers listening about that. Way to shift landscape of what was norm of what was happening in those classrooms. You have to use something that everyone can understand and handle.

TT: what Barbara has presented is interesting, we lack methodology, information we can take to scale, imp for this research. Pat says this is generative. Question is important – how do we keep the views, everything that gives this…also make it method collect some valid relevant math data without making it extremely complicated. Have someone in classroom for an hour. Question could be written down. Could be a way math relavant information is collected. We are trying to figure out…whether we can use it this way for study. Make it more part of math study.

Mexico: how much time used in copying and how much time used in q a session. Copying is a bad practice, how much time changes across grades, more secondary than primary. Method is very informative.

TT: presentation informative, what is possible.