What Do We See? Real Time Assessment of Middle and Secondary Mathematics Teachers' Pedagogical Content Knowledge and Sociomathematical Norms[1] Preliminary Report

Billy Jackson	Lisa Rice	Kristin Noblet
Saint Xavier University	University of Wyoming	University of Northern Colorado
bjackson@sxu.edu	lrice4@uwyo.edu	Kristin.Noblet@unco.edu

Abstract. The article reviews efforts to develop an observation protocol to assess the pedagogical content knowledge (PCK) and sociomathematical norms (SMN) that middle and high school teachers may develop over time as part of their participation in a master's program for secondary mathematics teachers. We observed each of 16 teachers in real time using the instrument, before involvement in the project and again after one year. Aspects of the protocol measure four critical components of PCK including curricular content, discourse, anticipatory, and implementation knowledge as well as some sociomathematical classroom norms. We present preliminary quantitative and qualitative analysis of the observations and discuss various challenges faced in the instrument development and its relation to similar protocols used by others previously.

Key Words: Pedagogical content knowledge, sociomathematical norms, inter-rater reliability, teaching moves

There have been several approaches to measuring the pedagogical content knowledge (PCK) of practicing teachers. Indeed, Hill, Ball, and Schilling (2008) and Hauk, Jackson, and Noblet (2010) have documented their development of written instruments designed to assess aspects of PCK. Both groups have developed theoretical frameworks for PCK that have similarities and some differences. One of the principle differences is that the Hill, Ball and Schilling linear model seeks to measure each of their proposed categories of PCK as distinct from each other, while Hauk, Jackson, and Noblet take a non-linear approach that presumes measurement overlap among categories.

Hauk, Jackson, and Noblet discuss PCK in terms of four components: curricular content, discourse, anticipatory, and implementation (action) knowledge. *Curricular content knowledge* is "substantive knowledge about topics, procedures, and concepts along with a comprehension of the relationships among them *as they are offered in school curricula*" (p. 2). *Discourse knowledge* "is about the culturally embedded nature of inquiry and forms of communication in mathematics (both in and out of educational settings" (p.2), and as such includes knowledge of mathematical syntax as a sub-category. *Anticipatory knowledge* "is an awareness of, and responsiveness to, the diverse ways in which learners may engage with content, processes, and concepts" (p. 3). *Implementation* or *action* knowledge "includes knowledge about how to adapt

[1] This material is based upon work supported by the National Science Foundation under Grant No. DUE0832026. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

teaching according to content and socio-cultural context and *enact in the classroom* the decisions informed by content, discourse, and anticipatory understandings" (p. 3).

Both groups' written assessments use multiple choice items and are limited in measurement of action knowledge. Implementation knowledge is more challenging to assess as this type of knowledge requires actions executed in the classroom (i.e., teacher moves). That is, the written assessments could not test for this type of knowledge because it requires that the teacher act "in the moment." At best, any written item could only gauge what a teacher *might* do in certain situations (e.g., see Ball, Hill, & Schilling, 2008).

In order to validate their written instrument, Ball and others (Learning Mathematics for Teaching (LMT), 2006) developed another instrument aimed at quantitatively measuring aspects of elementary and middle school teachers' classroom practice. Ten K-8 teachers who had taken a PCK test were videotaped for 3 times prior to, during, and after participation in professional development. Over the course of a year, a team of mathematicians, mathematics educators, mathematics teachers, and non-specialists analyzed the videos for various aspects of mathematics and mathematics teaching present in each lesson. A rubric was developed containing several items and video reviewers trained for and then coded each 5 minute segment of each lesson for X different categories of teacher move or classroom interaction. Each category had four possible codes: Present and Appropriate (PA), Present and Inappropriate (PI), Not Present and Appropriate (NPA) and Not Present and Inappropriate (NPI). LMT team leaders noticed early on a wide variability in how individuals coded lessons based upon the individuals' own professional backgrounds, and so to help ensure inter-rater reliability, the lessons were all recoded in pairs. A glossary describing each category (column) in the observation rubric was written, with each description giving some detail on when each code should be assigned during a segment.

Theoretical Perspective

Our research blends the Hauk et al., framework for PCK and the LMT instrument designed by the research team at the University of Michigan. We take the view that the teacher actions or moves (or the absence thereof) in the LMT protocol can be observed in the classroom, and that such actions or moves can be described (at least approximately) in a predetermined coding format *independently of the researcher involved*. Now, this is not to say that two different researchers may not observe and record different things (as frequently happened with the team at the University of Michigan and for our team) for a given segment, but, like the LMT tool, for an observation we would expect overall variation between observers to be minimal.

We use here the typologies of Hauk, Jackson, and Noblet. The reason is that any particular move that a teacher makes in the classroom can demonstrate multiple facets of PCK simultaneously, and hence we take their view that the strands of PCK are interwoven during instruction. Also, Hauk, Jackson, and Noblet make cultures in the classroom an explicit part of their definitions, which in turn may be part of teacher decisions to make certain moves in response to them.

The research questions for the work reported here are: How might we track the effects of professional development through changes in observed PCK and SMN? If traceable, how might professional development be designed to foster particular classroom moves through changes in PCK and SMN? Work on both of these questions continues, and we will primarily address the first here but some attention will be given to implications of current results for the second.

Methods

The research team at the University of Michigan point out in their technical report that there is a need to develop an instrument for doing observations in real time (Learning Mathematics for Teaching, 2006, p. 20). In order to address this need, we examined their observation protocol in some depth and determined which items were most appropriate given our focus on observing in secondary mathematics classrooms in real time (the LMT work was in grades K-8). Their protocol contained over 30 categories. To streamline for real-time observation we shortened to a protocol containing 20 items. Some of their categories were replaced or condensed in our version. For example, in the LMT version, the researchers created columns for the following: selection of correct manipulatives, and other visual and concrete models to represent mathematical ideas (their column II.e on sheet 2) and multiple models (column II.f on sheet 2). In our version, these two columns were condensed into the column that we titled multiple representations, which could include all of the things that the LMT team was looking for in II.e and f..

Great care was taken in finding an appropriate length of a segment to be viewed during the class. The team started with the 5 minute length that the LMT used for recorded sessions, but it soon became clear that a "5 minute on, 5 minute off" strategy in which the researcher would observe for 5 minutes and then record tallys on the protocol during the next 5 minute interval would result in 5 or fewer codings per class period for each category. Eventually, the team agreed upon observing for 3 minutes, and then recording for 3 minutes.

After the team started using the protocol, we began to reexamine the glossary that the LMT team had developed. We found that trying to use the instrument in real time created new challenges with respect to inter-rater reliability. In particular, the words "explicit" and "inappropriate" leave much room for interpretation even in the definitions provided by the LMT team. Though we used many of the same column categories and indentifying language as they did, we also saw it was important to craft definitions and create a new glossary. The idea was to create an instrument with sufficient examples and non-examples for each category that it could act as a coding book: a guide to the intended viewpoint of the protocol and how to observe through a particular lens. The eventual goal is to have an instrument that is terse but of sufficient detail that individuals can observe classrooms after a short calibration training paired with a practiced observer.

For example, while our glossary continues to be refined, we felt a need to be, well, more explicit about what "explicit talk about a topic or subject" meant. Currently, our glossary description of this category is: any utterance from student or teacher in which a topic or subject is stated verbally or in writing or by reference to a clear written or verbal precursor familiar to people in the room. In-vivo exemplars have been included in our glossary to demonstrate categories. For example, during one 3-minute segment, the teacher presented the Fundamental Theorem of Algebra. The exercise the teacher assigned called for students to find a polynomial of lowest degree with real coefficients that had certain prescribed roots. At one point, an exercise asked for a polynomial with roots $3i$, 4, and 5. The teacher produced a monic degree 4 polynomial with these 3 prescribed roots, and a student asked why it was necessary to have $-3i$ as a root when this number was not contained in the list. The teacher responded that since $3i$ was a root, its conjugate $-3i$ also had to be a root. The student again asked why this must be true when $-3i$ was not listed, and the teacher replied "because conjugates are always roots." The researcher coded this particular segment as NPI in the explicit talk about ways of reasoning column due to the teacher's not addressing directly the student's question (e.g., the idea that the

requirement that the coefficients of the polynomial be real was connected to the need for the use of conjugate roots).

Each column of the protocol was assigned a quadruplet of the form (*c, d, a, i*) where the values of *c, d, a, i* were determined by the research team based on the descriptions of the categories Curricular Content, Discourse, Anticipatory, Implementation and the glossary description of the category represented by the column. The possible values for *c, d, a,* and *i* were 1 if a particular kind of knowledge was present in the observable category and a value of 0 otherwise. Research team members spent a significant amount of time on coming to a consensus on implementation knowledge and trying to understand how it actually gets demonstrated in the classroom. One challenge in defining this particular type of knowledge is that the other three are interwoven with it so much that at times it can be difficult to "tease apart" implementation knowledge from say anticipatory knowledge. After much discussion, we began to understand that implementation knowledge had to meet **both** criteria given in the definition by Hauk, Jackson, and Noblet (i.e., satisfying only one of the two pieces was not enough). This categorical inductive coding left some of the columns without non-zero alignment to any PCK codes. In reviewing what was left uncoded, it was apparent that all of these were related to the establishing of sociomathematical norms. One such example is the column titled "instructional time is spent on mathematics (>75%)" in which a segment being marked as PA indicates nothing in particular about a teacher's knowledge of teaching mathematics, but rather indicates something about what the teacher and students treat as acceptable time to spend on mathematics instruction, fitting Yackel and Cobb's (1996) classic definition of sociomathematical norm. One particularly interesting column titled "encourages diverse mathematical competencies" has a unique feature: we determined that this column loaded heavily on PCK by assigning it a quadruplet of (1,1,1,1) (and hence having all four components of PCK) as well as being a sociomathematical norm. The item loads in discourse because of the communication about the mathematics that occurs between a teacher and student or among students when the item is present, and it loads on curricular content knowledge as a teacher must know about the connections among different procedures and solutions that students may use in solving problems. This previous statement also shows that a teacher will demonstrate anticipatory knowledge in this item's presence as she must be aware of how the students may interact with the problem at hand in order to encourage the competencies (i.e. curricular content and anticipatory knowledge overlap for this category). The teacher then uses her curricular content and anticipatory knowledge to adapt her teaching in response to the diverse competencies that arise as well making choices for her instruction in encouraging these competencies, thereby demonstrating her implementation knowledge. The item is also a SMN since the presence of the item in a segment is illustrative of a shared meaning between teacher and student of what diverse competencies in the classroom are.

Results

As indicated above, the research is currently in the data analysis stage, which will be complete by January, and a summary of results will be offered at the conference.

Questions (a handout of the protocol will be provided to the audience)

1. If the goal of observation of teaching is basic research about the nature of teacher enactment of PCK and SMN for secondary mathematics instruction, what are the implications of the realities of classroom practice for the revision of the protocol?

2. If the goal of the observation is evaluation of the impacts of a professional development (PD) program in which the teacher has been participating (PD focused on PCK growth), what are the audience suggestions for the revision/streamlining of the protocol?
3. How might such a protocol be used to help pre-service and practicum teacher candidates to think about and prepare for teaching?

References

Hauk, S., Jackson, B., Noblet, K. (2010). No Teacher Left Behind: Assessment of secondary mathematics teachers' pedagogical content knowledge. *Proceedings for the 13th Conference on Research in Undergraduate Mathematics Education (electronic)*. PDF available at sigmaa.maa.org/rume/crume2010/Archive/HaukNTLB2010_LONG.pdf

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*, 372-400.

Learning Mathematics for Teaching (2006). *A Coding rubric for Measuring the Mathematical Quality of Instruction* (Technical Report LMT1.06*)*. Ann Arbor, MI: University of Michigan, School of Education.

Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in Mathematics Education*, 27, 458-477.