Teachers' Knowledge Development During Lesson Study: Impact of Toolkit-Supported Lesson Study on Teachers' Knowledge of Mathematics for Teaching[1]

Rebecca Perry, Catherine Lewis, Shelley Friedkin & Elizabeth Baker
Mills College Lesson Study Group

24 March, 2009 (Draft)
Paper presented at AERA, 2009, San Diego

[1] This material is based upon work supported by the National Science Foundation under Grants DRL-0633945 and DRL-0723340. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Lesson study is one professional development approach for "learning from practice" (Ball, 1996; Ball & Cohen, 1999). During lesson study, teachers formulate long-term goals for student learning and development; collaboratively work on "research lessons" to bring these goals to life; observe, document and discuss student responses to these lessons; and revise the lessons (and the broader approach to instruction) in response to student learning (Fernandez & Yoshida, 2004; Lewis, 2002; Stigler & Hiebert, 1999). These four activities – planning, observing, analyzing student learning, and revising instruction – constitute a cycle of collaborative inquiry centered on classroom instruction, making the lesson study process consistent with many of the qualities U.S. researchers suggest are effective for teacher professional development (Darling-Hammond & McLaughlin, 1995; Goldenberg & Gallimore, 1991; Hawley & Valli, 1999; Loucks-Horsley, Hewson, Love, & Stiles, 1998; Putnam & Borko, 2000).

Since it first came to the attention of North American educators in 1999, lesson study has rapidly spread to schools and districts in North America; in 2008, more than 650 K-12 schools and 104 universities in North America reported ongoing lesson study efforts (Lesson Study Research Group, 2004; Lewis, Perry, & Murata, 2006; Mills College Lesson Study Group, 2008). However, the evidence that lesson study supports teachers' learning and subsequent instructional improvement is still scant (Lewis, Perry, & Hurd, 2009).

Several qualities of lesson study pose substantial challenges to researchers. First, lesson study groups typically determine the topic and focus of their own inquiry, so researchers cannot easily use pre-specified measures. Second, each teacher participating in a lesson study group is likely to learn something different, depending on their current knowledge; for example, there is evidence that a novice teacher and an experienced teacher watching the same lesson may each take away knowledge relevant to their own current challenges (Lewis, Perry, & Hurd, 2004). Third, lesson study's impact may occur in different domains of teacher development, including teachers' knowledge, practice, identity, sense of efficacy, motivation to learn, connections with colleagues, and habits of attending; this makes it necessary for researchers to cast a wide net and to understand how different dimensions of development interact over time – for example, how changes in connections with colleagues might subsequently support teachers' knowledge development.

In this paper, we present the preliminary results of a three-year study designed to conceptualize and measure teachers' learning and development as they engaged in a lesson study cycle supported by lesson study "toolkits." The "toolkits" were developed by the project and tested by elementary and secondary lesson study groups as they conducted lesson study on one of two mathematical topics: area of polygons or proportional reasoning. The toolkits were developed in response to our earlier research on naturally-occurring lesson study groups, which suggested that the teachers' mathematical learning was enhanced when groups had access to high-quality mathematics materials or expertise. For the toolkits, we located research and research-based instructional materials relevant to each of the mathematical topics, and suggested a process and discussion questions for using these resources in the lesson study cycle. In other words, we connected existing resources, including research-based articles, video, instructional materials, and student work and assessment items, to the tasks of lesson study. The toolkits were designed to build lesson study practitioners' own interest and understanding of each mathematical topic as they considered its teaching-learning, and to involve them actively as researchers (for example, by interviewing students and bringing the data to the lesson study group and by solving and discussing mathematical tasks in order to predict student thinking). We hoped that the toolkits would also build a desire on the part of lesson study group members to use research, and to place their own investigations within existing research.

We chose the topics of polygon area and proportional reasoning because of their foundational nature, the evidence that US elementary and middle school students do not understand these topics in a way that will support later mathematical learning, and the frequent requests we received from lesson study groups for resources on these topics. Measurement (the strand that encompasses polygon area in the NCTM standards and many others) is the only content area in which U.S. fourth grade students did not perform above the international average in a recent international study, and prior research suggests that the situation is even worse for eighth grade students.[i] Elementary students may not be gaining the strong conceptual knowledge of area they need for later facility with geometry, data representation, and other topics. A recent National Research Council report describes proportional reasoning in this way: "[it is] the capstone of elementary school arithmetic and the gateway to higher mathematics, including algebra, geometry, probability, statistics, and certain aspects of discrete mathematics. Nevertheless, US seventh and eight graders have not performed well on even simple proportion problems… On the 1996 NAEP, only 12% of eighth-grade students could solve a problem involving the comparison of two rates…."[ii]

Conceptualizing Teachers' Development During Lesson Study

Our conceptualization of teachers' development during lesson study is based on existing research on teachers' learning and development and on qualitative analyses of both North American and Japanese lesson study cases performed over multiple years (Clarke & Hollingsworth, 2002; Lewis, Perry, Hurd, & O'Connell, 2006; Lewis, Perry, & Murata, 2006; Lewis & Tsuchida, 1998; Perry & Lewis, 2008; Wang-Iverson & Yoshida, 2005). As shown in Figure 1 and described below, we conceptualize teachers' learning and development in three broad categories.

(1) Development of Knowledge for Mathematics Teaching

The first broad category is knowledge development, including content knowledge, pedagogical content knowledge, pedagogical knowledge, knowledge of student thinking and knowledge of curriculum (Hill, Ball, & Schilling, 2008; Seago & Goldsmith, 2006; Shulman, 1986). This category subsumes several types of knowledge that have been separated elsewhere; for example, Hill, Ball, & Schilling, 2008 differentiate subject matter knowledge that teachers share with other professional users ("common content knowledge") from subject matter knowledge specific to teaching ("specialized content knowledge"), and they differentiate subcategories of pedagogical content knowledge, including knowledge of curriculum, knowledge of content and students and knowledge of content and teaching. Although these researchers have made considerable headway in identifying the subcategories of knowledge needed to teach mathematics, we focus here on the umbrella category of knowledge for teaching mathematics as a useful place to begin. The following responses from teachers from the current study illustrate the range of knowledge included in this umbrella category. Responses were written at the end of the lesson study cycle, in response to a written prompt asking what they had learned about mathematics from the lesson study cycle.

> Using the area of a rectangle to solve other areas is the basis for most simple formulas. [ID323]
>
> There is a relationship between the area of a rectangle and the area of a right triangle. [ID324]
>
> Nesting concepts together is a more beneficial teaching approach than isolating. Teach how to find area of a rectangle and right triangle together so students can apply similar approaches to both. [ID326]
>
> Area is not to be taught quickly or is not to be done in a few days. The idea of area seems simple, but students struggle with strategies of finding area if they can't count the inside easily. The idea of area is not strongly developed prior to fifth grade. [ID 327]

(2) Development of Personal Dispositions

The second broad category is the development of personal dispositions that support learning. This category includes, for example, the teacher's identity (beliefs about one's responsibilities as a teacher), sense of efficacy (the belief that changes in instruction can affect student learning), beliefs about student intelligence and the extent to which it is malleable or fixed (e.g., Dweck, Chiu, & Hong, 1995), attention to student thinking and curiosity about it, and an inquiry and experimentation stance toward practice.

Identity acts as a filter that shapes how teachers invest their time and effort (Collopy, 2003; Spillane, 2000). Instructional improvement is extremely effortful work, and identity shapes where and how teachers invest effort – for example, teachers may invest time in improving their teaching of reading but not mathematics if they consider reading more important to student development. Teachers may see adaptation of the textbook as within or outside their purview. Lesson study may offer teachers opportunities to expand or modify the kinds of learning and work they see as central to teaching. For example, in our prior research, the comments of a kindergarten teacher following a 2-week mathematics lesson study workshop suggested she had expanded her ideas about the mathematics she was responsible for learning:

> I think a way I'm going to change is... As a Kindergarten teacher, I was always very focused on the standards. Of course, that was only the Kindergarten state standards. ...And I always thought "I like teaching Kindergarten because... I know enough. I don't need to learn any math. I know enough because I teach these five year olds." … when I saw that first grade example [of a lesson planned by Japanese teachers], they weren't thinking first grade math in their heads. I mean, they knew the standards all the way up. In their heads, they were probably going as far as they got in math. I mean, to me, it would be like high school math. And that's what they were thinking. ...I feel like I've been teaching with such a narrow perspective. Like "This is all I need to know to teach them."..." I really didn't understand the first week [of a two-week summer workshop] why we kept spending an hour or two on geometry. It was like "Who cares, I'm not going to teach this in Kindergarten." And then I realized, "No, I need to know the whole picture." ....I always thought "I know enough." ....I feel ashamed that that is the way I've been thinking.

Teaching efficacy is the belief that one's teaching can influence student learning. Studies indicate that efficacy is associated with desirable outcomes including student achievement, teacher retention, commitment to teaching, commitment to reform oriented education, and willingness to experiment with teaching innovations. Puchner & Taylor (2006) document a lesson study case in which teachers improved their sense of efficacy through collaborative re-design and re-teaching of a lesson that had been "disastrous" in their individual teaching the prior year. Questionnaire measures of teaching efficacy often include items like "By trying a different teaching method, I can significantly affect a student's achievement" (Cohen & Hill, 2001).

Attention to student thinking has also been found to be pivotal to teachers' learning to improve instruction (Franke, Carpenter, Fennema, Ansell, & Behrend, 1998; Seago & Goldsmith, 2006; van Es & Sherin, 2008). Certain types of professional development may help teachers attend to the mathematical thinking in student responses, rather than to focus only on evaluating whether responses are correct or incorrect (Schifter, 2001). The opportunities to study student thinking and make sense of it with colleagues that are provided by lesson study may stimulate teachers' attention to student thinking and curiosity about it.

A stance of inquiry and experimentation toward practice enables teachers to ask questions about practice and to experiment with changes. Such a stance can enhance teachers' capacity to learn from practice, curriculum, colleagues, and research, and a stance of inquiry and experimentation may be learned in part from participation in a lesson study culture centered on inquiry into practice.

(3) Development of Learning Community

The third broad category is development of a professional learning community, including capacity and commitment for joint learning among colleagues (Lichtenstein, McLaughlin, & Knudsen, 1992). Just as the dispositions of an individual teacher, such as curiosity about student thinking, may shape the teacher's learning over time, qualities of the professional community may shape its members' learning over time. For example, teachers may be able to learn more in professional communities where they have many opportunities to observe and analyze teaching, where the community norms emphasize continued learning and self-improvement, and where there are effective shared tools for study and discussion of practice. Lesson study may change the teacher learning community via three basic pathways, listed below:

a. Changes in Norms: Lesson study may shape the beliefs, values, expectations, etc. that broadly characterize the professional community. For example, a professional community may shift from norms of "privatized" teaching to norms of collaboration. Where teachers once felt hesitant to ask advice or observe colleagues, these may become widely expected and valued. Where teachers once felt it was fine for all teachers to "do their own thing," they may begin to feel responsible for each other's practice and for a coherent approach across grades.

b. Changes in Participation Opportunities

Participation opportunities are ongoing structures such as shared discussion of student work, collegial observation, negotiation of a shared lesson plan, etc. Participation may provide catalysts for individual knowledge integration, and may build tools and norms of the community.

c. Changes in Tools

The teacher community can provide tools that enable (or constrain) individual teachers' learning. For example, the lesson plan template that asks teachers to anticipate student responses, the protocol for observation of student thinking during the research lesson, and the agenda for the post-lesson discussion may provide means for teachers to develop their individual and collective knowledge.

Toolkit Overview

The toolkits were developed with the goal of supporting teachers' learning about the target mathematical content areas and at the same time supporting the development of the broader personal dispositions and teacher community structures just described. The toolkit materials included research articles, curriculum materials, classroom videos, "thought-revealing mathematical tasks," information on student learning trajectories, and student work that exemplified common solution strategies and misconceptions. [iii] These materials were intended to expand the knowledge available during the lesson study cycle, particularly during the *kyouzai kenkyuu* (curriculum study) that occurs during the first part of the cycle prior to lesson planning. Along with tools to support lesson study, the research and instructional materials were organized into binders that included suggested activities and discussion questions for using the materials.

The toolkits consisted of four primary activities. First, teachers individually solved several mathematical problems (taken from the public-release items from the National Assessment of Educational Progress, the Mathematics Assessment Resource Service and other sources), examined student work and response data for the tasks, and discussed their own solutions and their ideas about student responses. Next, teachers individually developed and then shared concept maps about the learning of the mathematical topic, focusing on the sequence of understandings students need to develop and the tasks and experiences that help them develop these understandings. Next, teachers explored more mathematical tasks and/or curriculum materials designed to help them build and refine as a group their concept map of the important student understandings and how they are developed. Excerpts of the Japanese mathematics curriculum and trajectory were included as one example of a coherent trajectory; it was hoped that teachers would refine their thinking about the teaching-learning of the topic by comparing the Japanese materials to familiar curricula. Teachers were also encouraged to read and discuss articles that had been selected to highlight particular important ideas about the target topics. Finally, in part four of the toolkits, lesson study groups selected a particular area of interest to study more closely, planned, conducted and discussed a research lesson on this topic, and drew out the implications of the research lesson for their teaching and learning. (See http://www.lessonresearch.net/nsf_toolkit.html for abbreviated versions of the two toolkits.)

Research Sample and Methods

This paper focuses on toolkit-supported lesson study by 16[2] teacher groups from across the U.S (and on comparison data from teachers who participated in pre- and post-assessments but not in lesson study). In some locations we were able to recruit two groups of teachers both willing to be assigned to either one of the toolkit-supported lesson study conditions (polygon area or proportional reasoning). In this case we randomly assigned one group to a topic, assigned the other group to the other topic, and used each group as an experimental group for one toolkit topic and a control group for the other topic. In locations where two lesson study groups did not volunteer, we recruited individual teachers to take the pre- and post- assessments only. The resulting sample of 109 teachers with matched pretest and posttest data is thus organized into three groups: those in the experimental condition for the proportional reasoning toolkit (N=8 groups and 44 teachers); those in the experimental condition for the area of polygons toolkit (N=8 groups and 37 teachers); and comparison teachers who received only the assessments (N=28 teachers). Figures 2 and 3 summarize sample and group characteristics, including a brief description of each group's target area of study.

The following data were collected from all groups: teacher surveys before and after the toolkit-supported lesson study (or at comparable times during the 2007-08 school year for the comparison teachers); artifacts produced or used during the lesson study cycle; brief meeting reports filled out at the end of each meeting; and end-of-cycle reflections on what was learned from the lesson study cycle. Additional data were collected from four sites chosen as "intensive study sites" (ISS): video- and audio-recordings and observational notes by one member of the research team.

The teacher surveys included four types of items aligned with the conceptual framework:

1) Multiple choice or fill-in-the-blank items related to teachers' knowledge of mathematics for teaching. These items were drawn primarily from the Learning Mathematics for Teaching project at the University of Michigan (Hill, Rowan, & Ball, 2005) and the University of Chicago School Mathematics Project (The University of Chicago School Mathematics Project, 2006);
2) Additional short-answer items written by our team, focused on teachers' knowledge of the two mathematics topics addressed by the toolkits (polygon area and proportional reasoning).
3) Likert-scale items on teachers' attitudes toward mathematics, teaching and learning, collegial work and lesson study activities (drawn, in part, from other survey instruments, e.g., Center for Research on the Context of Teaching, 1991); and
4) Fill-in-the-blank, multiple choice or short answer items on teacher demographic characteristics and group lesson study activities.

Most items in the first three categories were included on both the pre- and post tests, although presented in a different order between the two tests. The assessment instruments underwent several checks to strengthen their validity and reliability. For

---

[2] The study is ongoing in a few remaining sites. To date, research in 16 out of 20 intended sites has been completed.

example, with input from consulting mathematics experts, we iteratively reviewed original assessment forms provided by the other (LMT and UCSMP) projects in relation to the toolkit contents. We narrowed the selected items to include only knowledge that we expected teachers might encounter in their toolkit-supported lesson study work. For example, the LMT Winter, 2007 (pilot test) form of the proportional reasoning instrument included 35 question stems with 73 items; from these, we selected only 11 question stems with 24 items. We recognized that this process of reducing the total number of items related to a given topic area might reduce the reliability of the assessment in relation to the original form. Other items adopted from existing survey instruments (e.g., those on teachers' attitudes about students or teaching and learning or on teacher motivation) were included so that we could use scales related to our overall theoretical framework that had been tested in other research (e.g. Center for Research on the Context of Teaching, 1991).

About half of all items on both surveys were focused on teachers' knowledge of mathematics for teaching (55% of the total on pretest and 47% of total on posttest). Of the 53 items related to teachers' mathematical knowledge, 34 items were drawn from established instruments (all were multiple choice or fill-in-the-blank items). Of these 34 items, 24 were related to proportional reasoning and 10 were related to area of polygons. The remaining 19 items (from 3 question stems) developed by our group were mostly short-answer items, 16 of which were related to area of polygons and three to proportional reasoning. The two question stems related to polygon area asked teachers to describe what students should know about each of eight topics related to area of polygons and when they should learn it. As shown in Figure 4, the third stem presented a currency exchange situation and asked teachers to explain whether it was proportional and why and to use a graph, equation, and table to represent it.

Lesson study artifacts, meeting reports and end-of-cycle reflections provided additional qualitative information on each group's activities and teachers' perceptions of meeting effectiveness and self-reported learning. Two reflection forms are discussed here. A brief report completed at the end of each lesson study meeting asked teachers to describe whether and how the meeting contributed to their learning about mathematics and its teaching and learning, and to name the specific resources that contributed. The form also asked teachers to rate (on a 5-point Likert scale) aspects of the meeting such as whether it was a productive use of their time, increased their understanding of student mathematical thinking, and motivated or inspired them. This data source varied by group and teacher since some groups met more frequently than others (resulting in more opportunities to use the reflection form) and some teachers responded in more depth.

A second reflection form, completed at the end of the lesson study cycle(s), asked teachers to describe what they learned from the entire cycle about four aspects of the work: mathematics, curriculum, students, and learning with colleagues. Again, this source of data varied by group and teacher: Some groups conducted two lesson study cycles and completed two forms; other groups who began their participation in the study prior to a redesign of the toolkit materials did not complete the form at all.

In addition to teachers' written reflections, we also gathered focus group comments from teachers in the district with the largest number of participating sites (7). Teachers from the two topic areas met separately. Transcribed comments from those video-recorded meetings were an additional source of data.

Analysis

Analyses reported here focus on changes in the constructs included in the conceptual framework, from pre- to post- intervention for the toolkit-supported lesson study groups, and over a comparable time period for the comparison teachers. Analyses proceeded in four ways, dictated by the nature of the data, as follows.

(1) Quantitative survey data on mathematical knowledge for teaching

Teachers' responses to the standardized assessment items were scored as either correct (1) or incorrect (0) and summed to create two pre-intervention and two post-intervention scores (one for each toolkit topic area) for each participating teacher. Information on the items and sources for the two subtests is shown in Figure 5. These scores were then standardized and subjected to statistical tests of mean differences (t test and analysis of variance) and ordinary least squares regression analysis to assess the predictive value of the toolkit intervention on the posttest scores.[3] For example, we conducted separate (mixed-between) ANOVAs using the posttest knowledge measures (of proportional reasoning and area of polygons) as dependent variables, controlling for pretest scores, to determine if there was an improvement due to time (pretest-posttest), group (lesson study with relevant content toolkit or not), and the interaction of group by time.[4] For each of the two toolkit topics, we used the posttest subtest score as the dependent variable, toolkit group as a factor, and pretest subtest score as a covariate.

(2) Quantitative survey data on teachers' personal dispositions and learning community

These data derive from a question stem ("Please indicate how well each of the following statements describes your attitude…") followed by 25 items that teachers rated on a 5-point Likert scale. Scales were formed based on prior published research, theoretical considerations, and principal components factor analysis (varimax rotation) to identify items that defined common factors, using data from teachers with both pre- and post-assessment data. Three reliable[5] scales were identified by factor analysis. Below we include a brief description of each along with the internal consistency reliability (Cronbach's alpha) of the scale; scales are shown in the appendix. Scale scores were subjected to statistical tests of mean differences by time (pre- vs. post-) and group (toolkit-supported lesson study vs. comparison). We conducted separate analyses of

---

[3] We recognize the limits of ordinary least squares regression analysis for data such as ours, where teachers are organized into groups and thus not independent observations. Hierachical Linear Modeling (HLM) analyses necessary to address the nested (teacher within group) nature of the data are in process and will be reported at a later date.

[4] Four separate ANOVAs were conducted, varying both the group factor (i.e., two separate ANOVAs for: (a) proportional reasoning toolkit compared to teachers who did not receive that toolkit and (b) area of polygons toolkit compared to teachers who did not receive that toolkit) and the dependant variable (posttest score on the proportional reasoning teacher knowledge measure; posttest score on the area of polygons knowledge measure).

[5] DeVellis (1991) recommends an alpha below .60 as unacceptable; .60-.65 undesirable; .65-.70 minimally acceptable; .70-.80 respectable; and .80-.90 very good. (An alpha of .85 is recommended when developing a scale to compare groups on some property.) From http://faculty.css.edu/dswenson/web/scaledevstat.html 2/9/09.

these measures for the combined toolkit-supported lesson study groups compared to the comparison group (two groups) and for the three separate groups (proportional reasoning toolkit group, area of polygons toolkit group, and comparison group). We found similar results when the toolkit groups were combined. In the interest of providing the most complete picture of each group, we present results in this paper for the three groups separately.

(a) *Interest in and Enjoyment of Mathematics Teaching and Learning*. This scale includes 6 items pertaining to teachers' interest in mathematics learning and teaching. The reliability of this scale was .79 on both the pretest and posttest.

(b) *Expectations for Student Achievement*. This 6-item scale was used in teacher surveys administered by the Center for Research on the Context of Teaching in several of their projects (Center for Research on the Context of Teaching, 1991). The scale included items related to teachers' sense of efficacy, expectations for student learning, and students' attitudes and habits. The alpha reliability was .70 on the pretest and .65 on the posttest.

(c) *Collegial Learning Effectiveness*. This scale includes 4 items on perceived learning about mathematics and its instruction from colleagues. The scale reliability was .71 on the pretest and .64 on the posttest.

In addition to these three scales, we also identified several individual items that were of theoretical interest in relation to the conceptual framework. These included items related to teachers' perceptions of their own knowledge of the two topic areas, interest in learning more about the two topic areas, collegial learning, and student effort and achievement.[6]

(3) Short-answer assessment data

For the problem related to proportional reasoning shown in Figure 4, we gathered all matched pre- and post assessment results (N=58 complete pre and post sets, of which 18 were from proportional reasoning toolkit-supported teachers and 40 were from comparison teachers).[7] While some research documents students' over-reliance on linear methods (Dooren, De Bock, Janssens, & Verschaffel, 2008), our preliminary review of the responses to the currency exchange task suggested that the responses of these teachers (56% of whom teach at middle or high school) emphasized more abstract methods of representing proportional situations, such as setting up proportions as two equal ratios, cross-multiplying, and checking for equal products. We wanted to ascertain whether teachers' work with the toolkits supported awareness of other representations as well. Using proportional reasoning "big ideas" from the mathematics education literature (e.g.,

---

[6] We explored an additional scale of 5 items related to teachers' attitudes about student and teacher effort in relation to students' mathematics intelligence and achievement and whether students' math intelligence is fixed (or changeable). While we felt these items cohered theoretically, the low scale reliability (.43 on the pretest and .36 on the posttest) suggested otherwise. We are continuing to explore the relationship of these items to our conceptual framework.

[7] Fifty teachers in 8 groups from one district received an earlier form of the pre-assessment prior to the inclusion of the open-ended items, hence the lower number of teachers than for the previously-discussed measures.

(Cramer & Post, 1993; Cramer, Post, & Currier, 1993; Lamon, 2005; Langrall & Swafford, 2000; Lo, Watanabe, & Cai, 2004; Van de Walle, 2004), we developed a coding scheme to identify the big ideas about proportional reasoning that teachers mentioned in their open-ended responses. After two rounds of coding system refinement, two coders independently coded the data with inter-rater reliability of 73%. Further discussion enabled the two coders to reach 100% agreement on teachers' use of the criterion ideas about proportional reasoning shown in Figure 6. We then compared the number of criterion items used on the pretest and posttest.

One of the problems related to area of polygons asked teachers on both the pretest and posttest to "list briefly what, if anything, you want students to understand about each of the following 8 topics by the end of elementary school (if you are an elementary teacher) or by the end of the most relevant course (if you are a secondary teacher)." The eight topics included ideas related to area of polygons, including area of rectangles, area of parallelograms, congruent triangles, etc. For purposes of this manuscript, we focused our attention on teachers' responses to the "area of rectangles" prompt. We again developed a coding scheme to identify the big ideas about area of rectangles from the research literature; the coding system is shown with examples in Figure 7. We coded and counted the ideas newly added by teachers to their posttest responses that had not been present in their pre-test responses. Two coders independently coded the data, with initial reliability just under 80% agreement. Discussion resulted in 100% agreement on teachers' use of the coding categories.

(4) Teacher reflection data

The reflection questions asking teachers to describe what they learned (during a meeting or during the lesson study cycle as a whole) provided a window on teachers' self-perceived learning during lesson study. We initially developed a coding system based on the four major item categories laid out by Hill et al. (2008): common student errors; students' understanding of content; student development sequences, and common student computational strategies. We then added to this framework additional categories derived from our theoretical framework and inductively from reading the responses, including: development of an inquiry stance (or "learning how to learn"); desire/ motivation to continue learning and lesson study and collaboration; and barriers to lesson study or collaboration that might inhibit teachers' learning opportunities. Our intention in starting with the Hill et al. (2008) categories and adding inductively was to ascertain whether or not the learning reported by teachers fell within the conceptualization of mathematical knowledge for teaching embodied in the LMT, from which many of our assessment items were taken. We wondered whether the learning that teachers reported during lesson study was similar to that assessed by the LMT, or whether we might pick up different sorts of learning, such as ideas about how to teach the given topic, changing goals or expectations for students, or ideas about student capabilities.

A particular difficulty in coding the open-ended reports of learning was what to consider "learning." Some individuals responded to the prompt asking what they learned by recounting their group's activities. For example, teacher 327 mentioned on one meeting reflection that "our discussions are primarily about the actual lesson, what the flow of the lesson will be, what is expected from our students, and possible snags (timing, pace, materials)." This comment could be interpreted as simply a recounting of what the group was focused on during the meeting, or could be interpreted as the teacher learning something important (yet mostly unnamed) about lesson flow, expectations for students, and possible lesson difficulties.

We began coding for six primary constructs related to our conceptual framework: (1) learning of mathematical content; (2) learning about curriculum; (3) learning about student thinking and pedagogy; (4) development of an inquiry stance (or "learning how to learn"); (5) desire/ motivation to continue learning and doing lesson study/ collaboration; and (6) barriers to lesson study or collaboration that might inhibit teachers' learning opportunities. Using these six categories, we engaged in multiple rounds of sentence-by-sentence coding of teachers' written data, reliability checking, and elaboration of the coding guide for different lesson study groups. We were not able to achieve above 70% reliability between two coders, struggling, for example, with whether and how to categorize as learning teachers' statements like the following:

> "We talked about a lot of things, especially finding the area of parallelograms. There were several different ways to achieve this. Ways I hadn't thought of." [ID 323]
>
> "The lesson study has been very helpful in that it helps teachers to think about other concepts/ understandings that lead up to more complex problems." [ID 324]

Given our inability to achieve coding reliability, we honed in on one code designed to capture teachers' learning about mathematical content and about student understanding of content. At the time of manuscript preparation, we had just completed coding and had achieved an overall agreement of 93% reliability. Although results by toolkit group have not yet been quantified, we provide some illustrative examples of what teachers reported learning about content and student understanding of content.

Results

As noted in the introduction, we are interested in the impact of lesson study on three domains of teachers' learning and development: knowledge for mathematics teaching; personal dispositions; and development of the professional learning community.

Knowledge for Mathematics Teaching

Figures 8-10 show standardized pretest and posttest scores for each toolkit group versus the comparison group for the two content subtests (one focused on area of polygons, one on proportional reasoning).[8] The results indicate that teachers in both toolkit conditions showed small, statistically non-significant gains on the content related to their toolkit topic, while the comparison groups did not show gains. T-tests indicated that there were no significant differences at pretest between toolkit groups and their corresponding comparison groups and no posttest difference between the polygon area toolkit group and its comparison group. At posttest, the proportional reasoning toolkit group scored about half a standard deviation higher on the area of polygons subtest than its comparison group ($t=2.04$; $df=70$; $p=.045$), a difference attributable to the drop in mean subtest score for the comparison group at posttest.

[8] The final subtest scores were standardized in keeping with the terms of use of Learning Mathematics for Teaching Project standardized test items.

A second, more subjective, source of data is teachers' self-rated knowledge of the toolkit topic. Figures 11-13 show pre-post scores on teachers' responses to two Likert-scaled items "I have a strong knowledge of… (a) area of polygons; and (b) ratio, proportion, and rate." Teachers in both toolkit lesson study groups showed significant increases in *self-rated* knowledge of the toolkit topic they had studied. In other words, despite weak evidence of knowledge gain on the standardized measures, teachers believe they gained knowledge of the mathematical topic studied during toolkit-supported lesson study. The open-ended mathematics items developed by our group and the teachers' written reflections provide two additional data sources that may shed light on this apparent paradox.

Figure 14 shows data from the currency exchange problem developed by our group, focusing on pre- to post- changes in use of ten "big ideas" about proportional reasoning identified from the research literature. As the figure shows, teachers in the proportional reasoning toolkit condition increased by 10% or more use of seven of these big ideas in their solutions to the problem. Comparison group teachers increased their (by at least 10% or more) use of two of these ideas. The following ideas were mentioned at least 10% more frequently on the posttest than on the pretest by the teachers in the toolkit-supported proportional reasoning lesson study groups: constant rate/ ratio; multiplicative relationship; equivalent ratios; y=mx form of the equation; linear relationship/ straight line on a graph; line through 0; and evidence of equal cross products. Comparison group teachers increased their reference to two of these ideas: constant rate/ ratio and equivalent ratios. The comparison teachers were also less likely than toolkit-supported teachers (by at least 10%) to mention on the posttest the two big ideas of (1) unit rate and (2) linear relationship/ straight line on a graph. These results suggest that, at least for this task, the toolkit-supported lesson study may have increased teachers' flexibility in demonstrating a proportional relationship.

Figure 15 shows the data from the open-ended assessment item on area of polygons, providing the percentage of teachers who included new ideas at posttest related to each of the three overall categories: area calculation (e.g., different ways to calculate rectangle area); meaning of area (e.g., that it is the space inside and how the formula is derived); and connection to other topics (e.g., the connection between triangle and rectangle area). The data suggest that teachers in the area of polygon groups expanded their thinking about area as a result of their toolkit-supported lesson study work. Teachers assigned to the area of polygons toolkit increased their use of all three categories of ideas from pretest to posttest. The increase was particularly large for the category of ideas about the *meaning* of area or measurement, which the meeting artifacts suggest was a major focus of attention in several groups; 52% of teachers assigned to the area of polygons toolkit condition and 27% of comparison teachers added ideas in this category at posttest, a statistically significant difference.

Teachers' qualitative and focus group responses also suggest types of learning about content during lesson study that fall outside the types of knowledge measured by the standardized MKT assessment items, as the following examples for each topic illustrate. For example, the following open-ended responses from teachers in the proportional reasoning groups report a range of different kinds of learning about how students' reasoning builds from the elementary years through high school, about teachers' recognition of their own areas of weak knowledge, and about areas of difficulty for students and instructional implications.

Proportional Reasoning

Another thing that was interesting was that trying to map out which concept relates to the other concept and what concept leads to another concept that still has to do with proportional reasoning. That was interesting. It was an eye-opener for me. The fact that you can start as simple as counting by hands and then once you go up to Algebra 2 or let's say Algebra, you start dealing with slopes and lines and that's still a representation of proportional reasoning. Normally, if I were to teach proportional reasoning, it's what's in the book. Now when I teach it, I try to link it more towards slopes or more towards the development of slope and linear equations so that students can see where they're going to and what they need. [ID unknown]

We need to have teachers explore what's not proportional and what is proportional. Because I think that as elementary teachers with six units of math in college, and some of us had those six units of math 30 years ago, that the math content isn't there. And so I know that one way to teach it and that's the way I'll teach. I don't understand it. …. And until we say that understanding math is important, our kids are never going to grasp it in fourth grade and take it to fifth grade. So I think if we just stopped and said this is proportional and this is not. And I can't remember if the Japanese [curriculum] did this in fifth or sixth grade, but there were five examples – the candle burning and the sister's ages. Some of those were proportional and some of them weren't. So we need to build teachers' content knowledge. I have a degree in math and I know that my content knowledge for teaching understanding in fifth and sixth grade has grown more as a teacher than it ever has taking those college classes at __ State. Because I learned how to plug in formulas at __ State. But no one taught me understanding at __ State. [ID 154]

Since I'm still developing my knowledge and understanding of proportion and ratio, I was surprised to learn the importance of fractions, algebra, patterns, multiplication and addition, tables, and graphs in solving problems involving ratio and proportional reasoning. [ID 369]

Area of Polygons

Today we discussed how to find the area of a parallelogram. Everyone seemed to find a different way to find the area. Some of the ways are as follows:
1. Counting full and partial squares.
2. Making a much larger rectangle from the parallelogram and identifying a smaller rectangle or two triangles. The area of the smaller rectangle or two triangles was subtracted from the area of the larger rectangle.
3. Making six equal triangles, finding the area of one and multiplying by six.

After discussing our solutions and reviewing student work, we all agreed that a parallelogram can be found by base x height which is the same as the area of a rectangle (length x width). [ID 324]

Teacher: We figured out that we needed to start with measurement and how students would develop gradually from there - the understanding that area implies something more than just figuring out a formula - then we were able to work that out in the lesson…. We found out that not all but the majority of students had grasped the concept. [Researcher asks: "Students were able to work out the area without using the formula?...How could

> you tell they were not using a formula?"] We started out with different units of measurements and just gave out the idea of finding out what, what would cover the area of the polygon. By the end of the lesson the students were able to translate that into standard measurement units and that gave us the understanding that they had grasped the concept. [ID unknown]

In addition to the preceding reflections, which focus on teachers' understanding of the mathematics and of student thinking about the mathematics, many teachers reported that they had learned about instruction. While coding is ongoing, the following quotes hint at the areas of learning about mathematics teaching that were not picked up on the standardized MKT items. Some common themes focused on having students learn from and discuss multiple solution strategies; teaching conceptual ways to solve problems prior to teaching procedural methods (e.g., "teaching the formula for area last"); using visuals or manipulatives differently to highlight mathematical concepts; and the importance of using proper mathematical vocabulary.

> During today's meeting, I learned that in order for students to truly own the concept of finding the area of polygons, they must do so much more than simply learn (and memorize) the various formulas. Students must have experiences creating and manipulating different shapes to understand and identify the different properties of polygons. Once students are able to identify and classify the shapes, they must then be immersed in the vocabulary (and continue to be throughout). [ID 326]
>
> "I learned how powerful it is to consider and present all the solutions. [ID 313]
>
> The most important thing I learned was how affective [effective] simple visuals can be for certain students to learn the concept. [ID 313]
>
> Observing the students is where I find I learn most. Considering ways to modify what's being said by the teacher is challenging and interesting. A little change can promote so much more learning. [ID 316]
>
> Focusing/ starting with the student poster enlargement was very helpful for me in both grounding my thinking in real students as well as allowing me to think about the (more abstract) mathematical underlying ideas (proportionality). [ID 300]
>
> Today I learned about a personal shortage area. I found that I have little understanding about what mathematical skills are necessary for success on a given problem, and what previous math experiences will help students access the problem. [ID 302]

Development of Teachers' Personal Dispositions

Our theoretical framework suggests that a second domain influenced by lesson study is the development of personal dispositions conducive to learning and improvement of teaching. We had three measures of this domain. The first is interest and enjoyment in learning mathematics, a 6-item scale composed of items such as "I enjoy teaching mathematics," "I actively look for opportunities to learn mathematics," and "I am interested in the mathematics taught at many grade levels," as well as items tapping whether teachers want to learn more about the mathematical topics studied in the toolkits.

Results are shown in Figures 16-17 and they indicate that teachers in both toolkit groups reported pre-post increases in interest/ enjoyment that did not reach significance; comparison teachers' ratings were about the same on the pretest and posttest. Teachers in the proportional reasoning toolkit group showed significantly higher interest/ enjoyment in learning mathematics than the comparison teachers at posttest (t=2.37, df=68, p=.021) but not at pretest.

The second measure is expectations for student achievement, a scale that includes attitudinal items such as "Most of the students I teach are not capable of learning material I should be teaching them" (reversed) and "My expectations about how much students should learn are not as high as they used to be" (reversed). The findings, illustrated in Figures 16 and 18, again show that teachers in both toolkit groups reported statistically non-significant increases in expectations for student achievement in both toolkit groups; comparison teachers' ratings were about the same on the pretest and posttest. Comparison group teachers showed significantly higher expectations for student achievement at pretest than did area of polygons group teachers (t=2.27, df=63; p=.027), a difference that had disappeared by posttest.

Findings for the third measure, a single item that indicates how strongly teachers agree or disagree with the item "by trying a different teaching method, I can significantly affect a student's achievement" are shown in Figures 16 and 19. These figures reveal that teachers in the proportional reasoning toolkit group significantly increased their agreement with this item from pretest to posttest.

The overall pattern of the findings suggests the toolkit-supported lesson study positively shaped teachers' personal dispositions with respect to learning and improvement.

Development of the Teacher Learning Community

Since this study included a single cycle of lesson study and observations were conducted only at the four intensive study sites, opportunities to collect data on the development of norms, tools, and teacher learning structures were limited. However, we administered one scale, Collegial Learning Effectiveness, to tap teachers' perceptions of the effectiveness of the professional learning community. It included items such as "I have learned a lot about student thinking by working with colleagues" and "I find it useful to solve mathematics problems with colleagues." Results for this scale are shown in Figures 20 and 21. Results of independent sample t-tests showed that both toolkit groups rated the collegial learning effectiveness significantly higher at posttest than at pretest (for the proportional reasoning group, t=3.69, df=43, p=.001; for the area of polygons groups t=4.78, df=36, and p=.000). Both toolkit groups reported significantly greater collegial learning effectiveness than did comparison teachers at posttest (proportional reasoning group compared to control group: t=3.66; df=70; p=.000; area of polygons group compared to control group: t=2.70, df=63, p=.009), a difference that was not significant for either group at pretest. The qualitative data give a taste of some of the kinds of collegial learning that were important to teachers:

> The toolkit gave a lot of guidance, to give us examples… But the real gift came from the colleagues. And the more diverse the group, I think, the more that was gained. Not just grades…, but styles and different types of students and the

> whole gamut… There were a lot of stereotypes of "oh, you come from that school therefore your kids must be that way" and vice versa, and all of those were just blasted away, especially when we walked into each other's classrooms. …A GATE 4th grader can gain the same concept as a 6th grader if it's … taught in this style. That's what I came away with from lesson study. The colleague interaction and processing time. [ID unknown]

Discussion

Improvement of instruction is "steady work," demanding that teachers continue to learn over many years, in order to develop the extensive knowledge of subject matter, students, and pedagogy that enable in-the-moment decision-making during instruction (Ball, Hill, & Bass, 2005). Any study of teachers' learning during a single school-year is thus just a small window into a continuing process. Our theoretical framework conceives teacher learning as occurring over time through changes within and across three broad and interconnected domains of development: Knowledge; Personal Dispositions; and Teacher Learning Community. Because we see development of all three domains as essential to broad, long-term instructional improvement, and therefore essential to attend to in any professional learning intervention, we included (modest) measures of each domain.

In the domain of teacher knowledge, we used: short-answer and multiple-choice items selected from standardized measures of teachers' mathematics knowledge for teaching: open-ended mathematics tasks developed by our group; self-rated knowledge of the toolkit topics; and teacher reflections on what was learned from the lesson study cycle (in response to written prompts and discussion prompts). The standardized measures of teachers' mathematical knowledge for teaching did not indicate an impact of the toolkit-supported lesson study on teachers' learning. However, teachers reported that their knowledge of the mathematical topics had increased, and our other measures suggest that this indeed may be true. On the open-ended proportional reasoning and area of polygons tasks, teachers in the toolkit groups added more key mathematical ideas to their solutions from pre- to posttest than did the comparison group. The written reflections, although not fully analyzed and discussed here, also suggest learning about a number of aspects of each mathematical topic and its teaching and learning that were not captured by the standardized MKT assessment items. For example, teachers reported learning that students have difficulty mentally structuring the space inside a rectangle into square units; this finding has important implications for a teacher's design of instruction (Driscoll & Mark, 2007) but is not tapped by any standardized test item we located.

In the domain of personal dispositions, we studied three measures focused on teachers' interest and enjoyment in continuing to learn mathematics, their beliefs about student capacity to learn, and their belief that they student achievement can be affected through changes in instruction. These measures provide some evidence suggestive of increased interest and efficacy in the toolkit-supported lesson study groups relative to the control.

With respect to the third domain, teacher learning community, self-reported effectiveness of collegial learning showed significantly greater increases among teachers participating in toolkit-supported lesson study than among the comparison teachers..

These results suggest several implications for the study of teachers' professional learning during lesson study and more broadly. First, if we were to base recommendations about the effectiveness of toolkit-supported lesson study solely on the results of the standardized items used to measure mathematical knowledge for teaching, we would conclude that it is not effective to have teachers engage in toolkit-supported lesson study. However, the open-ended tasks and teachers' focus group and written reflections suggest that important knowledge relevant to teaching may indeed have been gained. The failure to detect impact using the standardized items is likely due to the (relatively) small number of items we used (10 for polygon area and 24 for proportional reasoning), a decision guided by both time considerations (we wanted to keep the whole assessment under two hours) and the lack of additional items directly relevant to the toolkit content. Were we to design the study again, we would probably elect to find or develop and test a larger set of items directly targeted on toolkit content.

Second, we think that the tentative theoretical framework for teacher learning proposed here deserves further discussion, refinement and use. The impact of the toolkit-supported lesson study shows up in all three domains of development – in teacher knowledge, personal dispositions, and learning community. While the current measures are modest, the domains merit further consideration and measure development. Focus on all three domains may help us keep our eye on the full range of teachers' development (both as individuals and as part of a teacher learning community) that is likely to fuel their future learning. From this perspective, we think one of the most intriguing findings in the study is toolkit teachers' expressed desire to learn more about the topic (polygon area or proportional reasoning) that they have already studied for an entire lesson study cycle. Likewise, there is evidence in the toolkit-supported lesson study groups of increased enjoyment and interest in learning and teaching mathematics. These changes strike us as potentially important elements of identity (as learners of mathematics) that may shape and fuel future learning. An oft-heard complaint about standardized testing of students is that it misses important qualities (such as motivation) that will shape future learning. We do not want to make the same mistake for teacher learning.

Finally, our theoretical model poses the question of how the domains interact over time to build teacher learning. Unfortunately, our current results do not allow us to answer this question. In future research, it will be important to ascertain whether, for example, the self-reported increases in enjoyment of learning mathematics lead to additional participation in learning activities, and whether the increases in reported effectiveness of learning with colleagues lead teachers to sustain and increase collegial learning. For more than a decade, professional development specialists have called for a shift from one-shot trainings to ongoing, job-embedded professional learning. Lesson study embodies many of the characteristics identified in these calls: it is practice-embedded, ongoing, engages teachers as learners, and draws on multiple sources of knowledge (from colleagues, careful observation of students, and research). However, our theoretical model of teacher learning over time and methods for tracing the interactions between knowledge, personal dispositions, and resources from the teacher learning community need much more development to take us beyond pre-post snapshots like those presented herein.

Appendix: Survey Scales

Collegial Learning Effectiveness (4 items; Alpha=.71 on pretest and .64 on posttest)

c) I have learned a lot about student thinking by working with colleagues.
d) Working with colleagues on mathematical tasks is often unpleasant (reversed)
l) I have learned a great deal about mathematics teaching from colleagues.
n) I find it useful to solve mathematics problems with colleagues.

Expectations for Student Achievement (6 items; Alpha=.70 on pretest and .65 on posttest)) - scale derived from the 1991 CRC Teacher Questionnaire

f) No matter how hard I try, some students will not be able to learn aspects of my subject matter (reversed).
k) My expectations about how much students should learn are not as high as they used to be (reversed).
o) Students who work hard and do well deserve more of my time than those who do not (reversed).
q) The attitudes and habits students bring to my classes greatly reduce their chances for academic success (reversed).
w) There is really very little I can do to ensure that most of my students achieve at a high level (reversed).
x) Most of the students I teach are not capable of learning material I should be teaching them (reversed).

Interest/ Enjoyment in Learning Mathematics (6 items; Alpha=.79 on pretest and .79 on posttest)

a) I enjoy teaching mathematics
e) I like solving mathematics problems.
j) I actively look for opportunities to learn more mathematics.
t) I would like to learn more about area of polygons
s) I am interested in the mathematics taught at many grade levels
v) I would like to learn more about ratio, proportion, and rate.

References

Ball, D. L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. *Phi Delta Kappan, 77*(7), 500-508.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey-Bass.

Ball, D. L., Hill, H., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(3), 14-17, 20-22, 43-46.

Center for Research on the Context of Teaching (1991). *Teacher Survey: 1991 Questionnaire*. Stanford, CA: School of Education, Stanford University.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*, 947-967.

Cohen, D. K., & Hill, H. (2001). *Learning policy: When state education reform works.* New Haven: Yale University Press.

Collopy, R. (2003). Curriculum materials as a professional development tool: How a mathematics textbook affected two teachers' learning. *The Elementary School Journal, 103*(3), 287-311.

Cramer, K., & Post, T. R. (1993). Making connections: A case for proportionality. *Arithmetic Teacher, 60*(6), 342-346.

Cramer, K., Post, T. R., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. T. Owens (Ed.), *Research ideas for the classroom* (pp. 159-178). New York: Macmillan Publishing Company.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 8*(76), 597-604.

Dooren, W. V., De Bock, D., Janssens, D., & Verschaffel, L. (2008). The Linear Imperative: An Inventory and Conceptual Analysis of Students Overuse of Linearity. *Journal for Research in Mathematics Education, 39*(3), 32.

Driscoll, M., & Mark, J. (2007). *Fostering geometric thinking (FGT) toolkit*. Portsmouth, NH: Heinemann.

Dweck, C., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions. *Psychological Inquiry, 6*(4), 267-285.

Fernandez, C., & Yoshida, M. (2004). *Lesson Study: A case of a Japanese approach to improving instruction through school-based teacher development*. Mahwah, NJ: Lawrence Erlbaum.

Franke, M. L., Carpenter, T. P., Fennema, E., Ansell, E., & Behrend, J. (1998). Understanding teachers' self-sustaining, generative change in the context of professional development. *Teaching and Teacher Education, 14*(1), 67-80.

Goldenberg, C., & Gallimore, R. (1991). Changing teachers takes more than a one-shot workshop. *Educational Leadership, 49*(3), 69-72.

Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Hammond-Darling & G. Skyes (Eds.), *Teaching as the learning profession: Handbook of policy and practice.* San Francisco: Jossey-Bass.

Hill, H. C., Ball, D. L., & Schilling, S. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuirng teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372-400.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Lamon, S. (2005). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers* (Second ed.). Mahwah, NJ: Erlbaum.

Langrall, C. W., & Swafford, J. (2000). Three balloons for two dollars: Developing proportional reasoning. *Mathematics Teaching in Middle School, 6*(4), 254-261.

Lesson Study Research Group (2004). [LSRG maintains a central database of U.S. lesson study group] Retrieved July 19, 2004, from http://www.tc.columbia.edu/lessonstudy/lsgroups.html

Lewis, C. (2002). *Lesson Study: A handbook of teacher-led instructional change*. Philadelphia, PA: Research for Better Schools.

Lewis, C., Perry, R., & Hurd, J. (2004). A deeper look at lesson study. *Educational Leadership, 61*(5), 18-23.

Lewis, C., Perry, R., & Hurd, J. (2009). Improving mathematics instruction through lesson study: A theoretical model and North American case. *Journal of Mathematics Teacher Education, 12*(in press).

Lewis, C., Perry, R., Hurd, J., & O'Connell, M. P. (2006). Lesson study comes of age in North America. *Phi Delta Kappan, December 2006*, 273-281.

Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher, 35*(3), 3-14.

Lewis, C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: Research lessons and the improvement of Japanese education. *American Educator*(Winter), 14-17 & 50-52.

Lichtenstein, G., McLaughlin, M. W., & Knudsen, J. L. (1992). Teacher empowerment and professional knowledge. In A. Lieberman (Ed.), *Changing contexts of teaching, NSSE 91st Yearbook: Part 1*. Chicago: Chicago Press.

Lo, J. J., Watanabe, T., & Cai, J. (2004). Developing ratio concepts: An Asian perspective. *Mathematics Teaching in Middle School, 9*(7), 362-367.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and math*. Thousand-Oaks, CA: Corwin Press.

Mills College Lesson Study Group (2008). Locate lesson study groups Retrieved 7/30/08, from http://www.lessonresearch.net/locate_ls_groups.pdf

Perry, R., & Lewis, C. (2008). What is successful adaptation of lesson study in the U.S.? *Journal of Educational Change*, DOI 10.1007/s10833-10008-19069-10837.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4-15.

Schifter, D. (2001). Learning to see the invisible: What skills and knowledge are needed to engage with students' mathematical ideas? In T. Wood, B. S. Nelson & J.

Warfield (Eds.), *Beyond Classical Pedagogy: Teaching Elementary School Mathematics* (pp. 109-134). Mahwah, NJ: Lawrence Erlbaum Associates.

Seago, N., & Goldsmith, L. T. (2006). Learning mathematics for teaching. In J. Novotna, M. Kratka & N. Stehlikova (Eds.), *Proceedings of the 30th Conference of the International Group fo the Psychology of Mathematics Education* (Vol. 5, pp. 73-80). Prague: PME.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Spillane, J. P. (2000). A fifth-grade teacher's reconstruction of mathematics and literacy teaching: Exploring interactions among identity, learning, and subject matter. *The Elementary School Journal, 100*(4), 307-330.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Summit Books.

The University of Chicago School Mathematics Project (2006). *Algebra/Geometry Readiness Test: Parts One and Two*. Chicago: University of Chicago School Mathematics Project.

Van de Walle, J. A. (2004). *Elementary and middle school mathematics: Teaching developmentally* (Fifth ed.). Boston, MA: Pearson Education, Inc.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education, 24*(2), 244-276.

Wang-Iverson, P., & Yoshida, M. (2005). *Building our understanding of lesson study*. Philadelphia: Research for Better Schools.

---

[i] North Carolina Mathematics and Science Education Network (2003). *The TIMSS international program*. Chapel Hill, NC: The University of North Carolina Center for School Leadership Development. Retrieved from http://www.unc.edu/depts/msen/statement/timss.html on July 15, 2005.

Gonzales, P., Guzman, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D. & Williams, T. (2004). *Highlights from the Trends in International Mathematics and Science Study: TIMSS 2003*. National Center for Education Statistics, US Department of Education. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005005 on July 15, 2005.

NCTM (2005, March). TIMSS results show US math moving in right direction. *News Bulletin*, March 2005. Retrieved from http://www.nctm.org/news/articles/2005_03nb_timss.htm on July 15, 2005.

[ii] National Research Council. (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

[iii] Cobb, McClain, K & Gravemeijer, K. (2003) Learning about statistical covariation. *Cognition and Instruction*, 21(1), 1-78.

Fosnot, C.T, Dolk, M. (2001) *Young Mathematicians at Work: Constructing Number Sense, Addition, and Subtraction.* Westport, CT: Heinemann.

Franke, M. L., T. Carpenter, et al. (1998). Understanding teachers' self-sustaining, generative change in the context of professional development. *Teaching and Teacher Education 14*(1), 67-80.

Heid, M. K., Blume, G. W., Zbiek, R. M., et al. (1998-1999). Factors that influence teachers learning to do interviews to understand students' mathematical understandings. *Educational Studies in Mathematics, 37*(3), 223-249.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996) A Longitudinal Study of Learning to Use Children's Thinking in Mathematics Instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Schorr, R. and R. Lesh (2003). A modeling approach for providing on-the-job teacher development. In R. Lesh and H. Doer (Eds.) *Beyond constructivism: A models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp.141-157). Mahwah, NJ, Lawrence Erlbaum Associates.

Lesh, R. Hoover, M., Hole, B., Kelly, A., Post , T. (2000). Principles for developing thought-revealing activities for students and teachers. In (A. Kelly&R. Lesh, Eds.) *Handbook of research design in mathematics and science education*. Mahweh, NJ: Lawrence Erlbaum Associates pp. 591-645.