Conceptualizing and Assessing Curriculum Embedded Mathematics Knowledge

Ok-Kyeong Kim
Western Michigan University

Janine T. Remillard
University of Pennsylvania

Improving Curriculum Use for Better Teaching Project (ICUBiT)

Paper presented at the Annual Meeting of the American Educational Research Association in New Orleans, LA, in April 2011. Do not cite without permission.

This content of this paper is based upon work supported by the National Science Foundation under Grants Nos. 0918141 and 0918126. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

ABSTRACT

As part of a 4-year project exploring teachers' curriculum use and the kinds of capacities (knowledge, abilities, ways of understanding, acting) needed for effective curriculum use, we developed a framework for and a tool to measure teacher knowledge required for understanding the mathematics underlying tasks, instructional designs, and representations in elementary mathematics curriculum materials. In this paper, we describe the methods and procedures used to undertake this work. In particular, we elaborate the framework with some sample assessment items. We also discuss issues and challenges that we faced in this conceptual and developmental work.

Perhaps more than ever before in recent U.S. history, educational leaders and policy makers are currently looking to mathematics curriculum programs to foment change in mathematics teaching and learning. In response to pressure to raise students' achievement scores and improve teaching, school districts are relying on the adoption of a single curriculum program as the primary strategy for influencing the teaching of mathematics. The expectation that curriculum materials can influence instruction is predicated on the assumption that teachers use them in ways that are true to the expectations of the curriculum designers. Research, however, has revealed substantial variation in how teachers use mathematics curriculum materials (Remillard, 2005).

Fewer studies consider the kind of knowledge and capacities necessary for teachers to use curriculum resources productively and fruitfully. Brown (2002) proposed the term *Pedagogical Design Capacity* (PDC) to characterize an individual teacher's ability "to perceive and mobilize existing resources in order to craft instructional contexts" (p. 70). PDC includes both teacher knowledge (subject matter and pedagogical content knowledge) and the ability to act with and on that knowledge. Researchers have yet to articulate the dimensions of PDC for the purpose of measuring it, studying it, and developing it in teachers.

In this paper, we identify a specific kind of teacher knowledge that we believe is an important dimension of PDC and describe our efforts to conceptualize and measure this knowledge. *Curriculum embedded mathematics knowledge* is the mathematics knowledge required to understand the mathematics underlying tasks, instructional designs, and representations in mathematics curriculum materials. A central assumption of our work, and an ongoing principle in our decision making processes, was that reading mathematics curriculum resources to guide instruction calls on a kind of mathematical knowledge that involves identifying and grasping the mathematical meaning and potential of tasks that is not always made explicit. We think of this knowledge as a part piece of mathematics knowledge for teaching (Ball, Thames, & Phelps, 2008) that is highly situated in nature, and ultimately a component of PDC.

Theoretical Perspectives

Teachers' use of curriculum materials is a relatively new area of study. (See Remillard, Herbel-Eisenmann, & Lloyd, 2008.) A commonly held assumption is that teachers are conduits of curriculum (Clandinin & Connelly, 1992), simply delivering the written curriculum to the student to produce learning. An alternative way of looking at teacher-curriculum use suggests a dynamic relationship between teacher and curriculum. Curriculum resources and teachers interact with each other: curriculum materials "talk" to the teacher through various features and the teacher interprets and interacts with these features (Remillard, 2000, 2005). As they prepare lessons, teachers read and interpret the suggestions in the materials by using their mathematics and pedagogical knowledge, previous experience, personal philosophy and beliefs.

The importance of understanding how teachers use curriculum materials is emphasized by Stein, Remillard, and Smith (2007) in their review of the research on the impact of curriculum materials on student learning. The role the teacher plays in shaping students' experience with curriculum materials is captured in their framework, which includes phases of curriculum use. Over the course of curriculum use, the written curriculum (printed materials) is transformed into the intended curriculum (teacher plans for instruction) and then into the enacted curriculum (actual implementation of the lessons). Using this framework, Stein, et al. argue that curriculum materials alone do not shape opportunities for student learning. Student learning opportunities are influenced, first, by how teachers interpret and use curriculum materials to plan instruction and, second, by how the teacher's plans are enacted in the classroom with students. In order to understand how teachers use curriculum materials to design instruction, we examine teacher knowledge as one of the most significant factors that influence teacher curriculum use.

Teacher Knowledge for Mathematics Teaching

Those who study teaching have long identified *teacher knowledge*, or the specialized knowledge that teachers develop, hold, and use, as central to understanding the practice of teaching. Distinct from general theoretical or content knowledge, teacher knowledge is conceptualized as situated within the work of teaching (Ball & Bass, 2003; Ball, Lubienski, & Mewborn, 2003; Carter, 1990; Shulman, 1986). This work involves constant assessment and decision making based on knowledge of content, pedagogy, curriculum, and students. Some scholars have proposed dimensions of teacher knowledge that are not associated with particular content areas and that focus on pedagogical decision-making and the assessment of students. Carter, for example, argued that practical knowledge includes “knowledge teachers have of classroom situations and the practical dilemmas they face in carrying out purposeful action in these settings" (p. 299).

The most influential work on teacher knowledge has focused on knowledge that is content specific. In the late 1980s, Shulman and colleagues introduced the construct of *Pedagogical Content Knowledge* (PCK), describing it as a specialized form of content knowledge focusing on the "aspects of content most germane to its teachability"

(Shulman, 1986, p. 9) and "is enriched and enhanced by other types of knowledge, including knowledge of the learner, knowledge of the curriculum, knowledge of the context, knowledge of pedagogy" (Wilson, Shulman, & Rickert, 1987). Teachers demonstrate *Pedagogical Content Knowledge* as they transform content knowledge into a form that is understandable to students; this act of generating possible forms of representation and choosing between them is referred to as *pedagogical reasoning*. In this sense, we see PCK as a central component of PDC.

More recently, Deborah Ball and her colleagues have introduced the concept of mathematics knowledge for teaching [MKT], which is arguably a component of PCK for mathematics. MKT focuses on understanding of subject matter that enables teachers to carry out key tasks of teaching, including using representations to model mathematical concepts, assess representations and tasks in curriculum materials, provide students with explanations, and assess their solutions (Ball & Bass, 2003; Hill, Sleep, Lewis, & Ball, 2007). Their work has focused, primarily on developing tools to measure this specialized mathematics knowledge.

Measuring Knowledge for Mathematics Teaching

Despite general agreement that specialized mathematics knowledge is necessary for teaching, the field has been limited in its capacity to measure this knowledge, primarily because of the way it is embedded in practice and intertwined with other forms of knowledge. In this study, we draw on work from recent efforts to capture and measure this specialized knowledge.

Through extensive work, Ball and colleagues have developed and validated multiple choice items to assess *Content Knowledge for Teaching Mathematics* [CKTM] that target elementary and middle school teachers, assessing various content areas as well as pedagogical content knowledge (Hill & Ball, 2004). These multiple-choice items make it possible to gather a large sample data in a reliable way. Still, the assessment remains distinct from a teacher's actual practice. This study used excerpts from various curriculum materials to develop an assessment to measure teachers' understanding of the mathematics content needed for teaching. By using curriculum materials in the assessment, we tried to develop a measure that addressed teacher knowledge tied into teaching practice.

Some researchers have used cognitive interviews to measure teacher knowledge (e.g., Ball, 1990; Ma, 1999; Tirosh & Graeber, 1989, 1990). Ma's interviews of Chinese and U.S. teachers are well known because they allowed her to contrast the "profound understandings of mathematics" and "knowledge packages" held by expert Chinese teachers to the segmented pieces of knowledge common among U.S. teachers. Whereas the other kind of measures attempt to quantify teacher knowledge, cognitive interviews examine teacher knowledge in depth and qualitatively. In this study, during multiple rounds of pilots we interviewed teachers to gather data on teachers' thinking on the mathematics embedded in curriculum materials and to refine the items.

Curriculum Embedded Mathematics Assessment

The *Curriculum Embedded Mathematics Assessment* [CEMA] is a tool to measure teachers' understanding of the mathematics underlying tasks, instructional designs, and representations in elementary curriculum materials. The development of CEMA was guided by two questions: (1) *How are mathematical ideas represented and embedded in various features of elementary curriculum programs?* and (2) *How are these ideas* *interpreted by elementary teachers?* To begin with, we analyzed various curriculum materials to examine content coverage and its sequence and identify key representations used in each curriculum. We also examined common and unique features that were used in the curricula. Using various representations and features in the curricula, we developed items to assess teachers understanding of the mathematics that they teach.

The rationale for developing this assessment is that the mathematical understanding required to interpret the important mathematical ideas embedded in curriculum-based tasks is an important dimension of mathematics knowledge for teaching that has not been explicitly studied. Moreover, 'reading' these tasks and interpreting them for the purpose of designing instruction is a central task of teaching. Thus, curriculum-based tasks offer a potentially fruitful vehicle for exploring critical components of teachers' mathematical knowledge.

Our work on the CEMA aimed at developing a *proof of concept* of the specialized knowledge needed to read the mathematics in curriculum materials and its relationship to Mathematics Knowledge for Teaching (Ball & Bass, 2003). As a result, the CEMA we developed should be understood as a prototype of a new tool to measure teachers' understanding of mathematics embedded in curriculum resources.

The assessment uses excerpts from five different elementary mathematics curriculum materials around which questions (items) about the mathematical intent and purpose are formulated. Each excerpt is followed by 4-7 items. The five programs are: (1) *Investigations in Number, Data, and Space*, (2) *Everyday Mathematics,* (3) *Math Trailblazers*, (4) Scott Foresman *Mathematics*, and (5) *Singapore Mathematics.* The first three are NSF-funded programs; the fourth is commercially developed; the fifth is from

Singapore and has gained popularity in the U.S. over recent years. In order to narrow the content focus, we chose the tasks related to the number and operations and algebra strands from these programs.

CEMA: A prototype of a new tool to measure teachers' understanding of mathematics embedded in curriculum resources (tasks, representations, teachers' guides, etc.).
Guiding questions: (1) How are mathematical ideas represented and embedded in various features of elementary curriculum programs? (2) How are these ideas interpreted by elementary teachers?
Aim: Developing a *proof of concept* of the specialized knowledge needed to read the mathematics in curriculum materials and its relationship to Mathematics Knowledge for Teaching (Ball & Bass, 2003).
Structure: Excerpts and associated items – excerpts from five different elementary mathematics curriculum materials around which questions (items) about the mathematical intent and purpose are formulated (4-6 questions per excerpt).
Five programs used: (1) *Investigations in Number, Data, and Space*, (2) *Everyday Mathematics,* (3) *Math Trailblazers*, (4) Scott Foresman *Mathematics*, and (5) *Singapore Mathematics.*
Content focus and grade level: Number and operations and algebra strands in grades 3-5.

Figure 1. Essence of CEMA

Methods and Procedures

The CEMA development team was composed of two mathematics educators, a statistician, a mathematician and mathematics educator, and four research assistants. The two mathematics educators had taught several years in elementary schools and worked with preservice and inservice elementary teachers. The mathematician had taught middle school and served as an education program leader. The statistician had expertise in psychometrics. All four research assistants had mathematics teaching experience and had used mathematics curriculum materials, primarily in middle and high schools.

The process used to develop the CEMA involved six steps: Conceptualization; Item development; Piloting and revision; Expert review and revisions; Development of online assessment; and Field testing and analysis. These steps are summarized in Figure 2. The procedures taken were not completely linear in a sense that the results of later work led the team to revisit earlier work in many cases.

A. *Conceptualization*

- Review and analyze 5 curriculum programs to identify key features.
- Develop initial items using each of the 5 programs, focusing on its key features.
- Develop a conceptual framework of the kind of knowledge to measure.
- Identify 4 dimensions as components of the knowledge.
- Interview teachers about curriculum reading to test and further refine the conceptual framework.

B. *Excerpt and Item Development*

- Identify excerpts from each curriculum and develop items for each excerpt.
- Determine 8 excerpts with open-ended items for pilot.
- Conduct pre-pilot with teachers and graduate students to ensure readability.
- Revise the 8 excerpts and associated items based on the pre-pilot results.

C. *Pilots*

- Conduct the first pilot with 26 teachers (surveys).
- Revise the excerpts and items, and create multiple-choices for each item based on the first pilot
- Conduct the second pilot with additional 27 teachers (survey + interview)
- Revise the excerpts and items based on the second pilot

D. *Expert Review*

- Send the excerpts and items to math educators, mathematicians, and psychometricians for review
- Revise the excerpts and items based on the external reviews

E. *Online Assessment*

- Group excerpts and items into two and develop two online survey links
- Format items appropriate for online response
- Test the online assessment with 13 classroom teachers

F. *Field Test and Item Analysis*

- Conduct a field test with 154 teachers online
- Analyze the items using teacher responses (in spring 2011)

Figure 2. Conceptualization and Development of CEMA

Conceptualization

Our first step was to conceptualize the domain of knowledge we were hoping to measure. Doing so was particularly difficult because little research exists on the mathematical knowledge needed to understand curriculum embedded knowledge. We began the work by reviewing the five curriculum programs in order to identify unique features and representations incorporated in each program as well as characteristics common among them. Concurrently, we identified excerpts from the five programs in the selected content strands that seemed mathematically fruitful and that covered a range of subtopics, and we developed one or two items for each excerpt. We drew on our analysis of the texts and their content, our preliminary experiences creating items, and our experiences teaching and working with teachers to propose a set of dimensions. These dimensions were further conceptualized and refined through the item development process.

Four dimensions. The framework contained various aspects in this specific kind of knowledge that we were trying to assess, which we referred to dimensions. Based on previous literature and our experience in creating items, these dimensions were narrowed down to four major aspects of *Curriculum Embedded Mathematics Knowledge* as follows:

1. **Mathematical ideas** – Knowledge related to the mathematical ideas embedded in a particular task or student work; the ability to identify the mathematical point of a task or lesson.
2. **Surrounding knowledge** – Knowledge of how a particular mathematical goal is situated within a set of ideas, including the foundational ideas that undergird it and the future ideas that can be developed from it.
3. **Problem complexity** – The ability to assess relative complexity and difficulty of a variety of mathematical ideas or tasks. The ability to categorize and order by

increasing difficulty. The ability to identify possible points of confusion for learners associated with a given task.

4. **Connections across representations** – The ability to make connections across representations of the same mathematical idea, including narratives and symbolic representations.

When we discussed items we developed, we characterized each of them in terms of the kind of knowledge that was required to answer the item correctly. In order to specify the kind of knowledge we tried to assess, we categorized them in terms of major dimensions of the knowledge required, while eliminating less substantial ones. This process required multiple rounds of a careful examination of the excerpts used, the mathematics embedded in them, and the kind of knowledge needed to recognize the mathematics. It is important to note that we do not think of these dimensions as identifying subscales because it is not always possible to assign a single dimension to each item. In fact, these dimensions are closely related to each other. (This issue is further explained in the forthcoming example and in the item development process.)

Literature connections. As we characterized the four major dimensions, we reviewed related literature and compared different characterizations put forth by various researchers. That led us to see similarities and differences among them and evaluate our framework in progress in terms of what was missing.

Figure 3 illustrates similarities and differences of the main literature that we used (Ball, Thames, & Phelps, 2008; Grossman, 1991; Shulman, 1986; Sleep, 2009). For each of the characterizations, we identified major categories and their specification. Then, we compared them to see how these different characterizations fit into each other. Shulman, and Ball, et al. attend to content knowledge for teaching entirely. Shulman's content knowledge in teaching has three major categories: subject matter content knowledge, pedagogical content knowledge (PCK), and curricular knowledge. Based on Shulman, Ball, et al. further specify mathematical knowledge for teaching (MKT) as subject matter knowledge and pedagogical content knowledge, including knowledge of curriculum in PCK. In contrast, Grossman and Sleep each focus on one particular area of teacher knowledge. Grossman's main focus is on pedagogical content knowledge, which includes conceptions of the purposes of teaching (e.g., "teacher beliefs about what is important for students to know, understand, and appreciate about specific content, and their understanding of the interrelationship of topics within a subject," Grossman, 1991, p. 209) and knowledge of curriculum. Note that both Grossman and Ball, et al. see knowledge of curriculum as part of PCK.

<table>
<tr><td>Researcher</td><td colspan="7">Characterization of teacher knowledge</td></tr>
<tr><td>Shulman's content knowledge in teaching</td><td colspan="3">subject matter content knowledge</td><td colspan="2">pedagogical content knowledge</td><td></td><td>curricular knowledge</td></tr>
<tr><td rowspan="2">Grossman's pedagogical content knowledge</td><td rowspan="2" colspan="3"></td><td colspan="4">Pedagogical content knowledge</td></tr>
<tr><td>knowledge and beliefs regarding student understanding</td><td>knowledge of instructional strategies and representation</td><td>conceptions of the purposes of teaching</td><td>knowledge of curriculum</td></tr>
<tr><td rowspan="2">Ball, et al.'s mathematical knowledge for teaching (MKT)</td><td colspan="3">Subject matter knowledge</td><td colspan="4">Pedagogical content knowledge</td></tr>
<tr><td>common content knowledge</td><td>knowledge at the mathematical horizon</td><td>specialized content know knowledge</td><td>knowledge of content and students</td><td>knowledge of content and teaching</td><td colspan="2">knowledge of content and curriculum</td></tr>
<tr><td rowspan="2">Sleep's mathematical purposing</td><td rowspan="2" colspan="4"></td><td colspan="3">Mathematical purposing</td></tr>
<tr><td>knowledge of content and teaching</td><td colspan="2">knowledge of content and curriculum</td></tr>
</table>

Figure 3. Comparison of various characterizations of teacher knowledge

Drawing on such previous work, Sleep articulates Ball, et al.'s knowledge of content and teaching (KCT) and knowledge of content and curriculum (KCC) by analyzing the mathematical knowledge demands of identifying mathematical goals and using them to design instruction, which she refers to as *mathematical purposing*. In fact, Sleep describes the relationship between other components of MKT, specialized content knowledge in particular, and mathematical purposing, and yet her major contribution is the elaboration of KCC, which has not been clearly defined previously. She suggests specific "types of knowledge, reasoning, and dispositions included in the domain of KCC" (p. 225), which include:

- Knowledge of foundational mathematical learning goals and the instantiation of those goals at particular grade levels; and the ability to determine how they can be worked toward in particular activities
- Knowledge of productive curricular trajectories through the mathematical terrain for different topics; connections across this content; the typical order in which it is taught; what is assessed
- The ability to specify coherent mathematical learning goals of different types and grain sizes appropriate for a particular instructional activity and to understand how the details of the instructional activity are intended to move students toward those goals.

The above cross examinations assisted us to conceptualize the kind of knowledge that we tried to measure through CEMA, influencing our thinking of the four dimensions and helping us refine each of them. For example, Sleep's elaboration in the bulleted list above helped us refine Dimensions 1 and 2; the specification of PCK given by Shulman enabled us to detail Dimensions 3 and 4. Among the four, however, Sleep's conceptualization of teacher knowledge was related most to our notion of *Curriculum Embedded Mathematics Knowledge.* We drew on her specifications of KCC to set the boundary of the knowledge we tried to conceptualize and assess. Since her focus was on mathematical purposing for effective instructional design and teaching, Sleep paid attention to not only knowledge and reasoning but also conceptions and dispositions needed for this teacher task. We narrowed down our focus to the kind of knowledge and reasoning that were required to read and understand the mathematics represented in the curriculum materials.

Examples of the dimensions. Even though these dimensions specified distinct aspects of *Curriculum Embedded Mathematics Knowledge*, the process of developing items revealed that these dimensions were interrelated in many cases. For example, it was often the case that Dimension 1 (knowing the mathematical ideas embedded in a task or identifying a mathematical point of a task) was relevant to all items, even those designed specifically to address other dimensions. The following example excerpt and questions illustrate the four dimensions and how they are interrelated[1]. A student page from Scott Foresman *Mathematics* describes two solution methods for a story problem: "Trisha and her brother, Kyle, collect and sell baseball cards. Kyle has 6 cards to sell. Trisha has 3 cards to sell. If they sell the cards for 8¢ each, how much money will they get all together?" One solution involves finding the total number of cards and multiplying that by 8¢; the other involves finding how much money each person gets and then combining the two. Using this excerpt and the dimensions, we generated four questions to ask. The first question, *What fundamental mathematical idea provides the basis for why the two solution methods produce the same answer?*, posed with four multiple choice responses,

[1] Examples of actual excerpts and items will be presented at the conference. For the protection of the items, in this paper we explain them without providing exact items verbatim.

addresses Dimension 1 (mathematical ideas and points), as it asks teachers to identify what fundamental mathematical idea relates the two solution methods (i.e., distributive property). Another question asks teachers to identify which representation among the four given illustrates the relationship between the two solution strategies. This question is an example of Dimension 4 (connections across representations), as teachers have to relate the problem context and solution methods with an appropriate representation. To see the connections between the problem context and representations, and to answer the question correctly, however, teachers should recognize the relationship between the two solutions, which means that this question also requires Dimension 1. A third question also aims at Dimensions 1 and 4 by using various story problems. The last question associated with this excerpt asks teachers to write an equation in a generalized form that shows the relationship between the two solutions, which addresses Dimensions 1, 2, and 4. The first and the last are based on the same reasons explained above. Dimension 2 (surrounding knowledge) is also required in this question as it asks teachers to represent the mathematical idea in the excerpt (i.e., distributive property) in a symbolic form, which can be applied to any numeric values. A symbolic representation of the distributive property will be introduced to students much later, but teachers are expected not only to know the property in concrete situations, but also to be aware of its abstract form and how it is related to the concrete level.

The excerpt explained above does not include a question that addresses Dimension 3 (problem complexity). As an example of this dimension, one question from another excerpt is illustrated. This question asks teachers to order three division problems from easiest to most difficult, given that students solve them using base-ten blocks and the method pictured in the excerpt (i.e. partitive method of division - finding group size with the given total and the number of groups). Teachers have to assess complexity and difficulty of these division problems as the problems vary in terms of the relationship between the dividend and the divisor, the number of times regrouping is required, and the presence of a remainder.

Empirical check. In order to increase our understanding of what was required to read and use curriculum materials effectively and to gather empirical data on the four dimensions, we conducted interviews with 7 teachers in grades 3-5 in western Michigan and Philadelphia. During the interview, each teacher was asked to think aloud while reading to prepare one lesson from each of two different programs. First, we asked them to talk through the planning of a lesson in the curriculum program they were using. Then, we asked them to talk through a lesson using a novel or unfamiliar program. We used both a familiar and unfamiliar program because we suspected that different kinds of knowledge might be activated in these two instances. The results of the interviews also confirmed that the four dimensions were appropriate for our purpose of the CEMA development.

Excerpt and Item Development

As we refined the four dimensions, we continued to identify excerpts and develop items. With the dimensions in mind, for the purpose of psychometric analysis of the assessment, the project team tried to create items for each dimension. This challenged the project team, however, because the four dimensions were closely related to one another, which made it hard to create items addressing one distinct dimension. After several trials and discussions, we revised our aim to develop items that addressed mathematical ideas described in the dimensions, regardless of the number of dimensions incorporated in each of the items.

The fact that each item was excerpted and adapted from curriculum materials created a particular challenge for item development related to the amount of context necessary to make the item understandable and answerable. The curriculum-embedded nature of the items was critical to the knowledge we hoped to assess; however the challenge was to avoid developing items that were excessively long or complex to comprehend. Following the suggestion of the psychometrician on the team, we adopted a model used by reading comprehension tests that provide a written excerpt followed by questions that assessed different kinds of knowledge related to comprehension of the excerpt. This format appealed to us because it allowed us to begin with a significant curriculum excerpt and then ask multiple questions about the excerpt. Using various parts of the curriculum materials (e.g., student pages, teacher professional development notes, mathematical tasks, representations, etc.), we selected sufficient amounts of content and context for teachers to think about to answer the questions posed.

Eventually, we identified 8 excerpts from the five programs and developed 5-12 items per excerpt to assess teachers' understanding of the important mathematical ideas embedded in each excerpt. Before piloting the items with teachers, we checked informally with classroom teachers and graduate students for clarity and gathered their initial responses. This pre-pilot was conducted in various formats including 3 discussions-in-groups, 3 interviews, and 7 survey-and-interviews. Based on the results of this pre-pilot, we added or reduced the details in the excerpts and revised associated items.

Pilots

The version refined through pre-pilot was used for the first pilot, which involved 26 participating teachers responding to the items on paper (3 excerpts and associated items per teacher). This produced 9 or 10 teacher responses per excerpt. These responses were combined by item to see how many teachers gave a correct answer, what common misconceptions surfaced in teacher responses, and how diverse the responses were. Such examination helped develop multiple-choices of items to get ready for the second pilot. In some cases, we deleted or substantially revised items, especially when the items were too easy, or led to confusion or too much diversity in teacher responses. Even though most of the items were given multiple-choices, a certain number of items remained as open-ended or short-answer.

The second pilot was conducted with additional 27 teachers. These teachers completed the items for three excerpts and then participated in a cognitive interview during which they were asked to explain the intent behind their choices for each item. The purpose of the interviews was to check for reliability between the selected responses and the teachers' intended responses. Based on the results of this pilot, the assessment was revised one more time for external review.

Expert Review

The assessment refined through previous stages was sent to the advisory board members and the external evaluation team for their review. This external review panel was composed of experts in curriculum development, item development, mathematics, and mathematics teaching and teacher education. We provided each reviewer with the eight excerpts and associated items, and a brief description of the process we used to develop the CEMA. Each excerpt-item set was presented in annotated form, which included background information on the source of the excerpt (e.g., grade level, curriculum) and explanatory rationales for each item and its answer(s). We asked the reviewers to use their expertise to examine the assessment in terms of: (1) appropriateness of the mathematical point of item, (2) clarity of item, (3) appropriateness of answers and distracters (multiple choices), and (4) appropriateness of stem (the way each item is posed). Each reviewer provided comments, concerns, and suggestions in the electronic files of excerpt-item sets, which facilitated the final revision before field-test. The greatest challenge during this revision was the coordination of perspectives from experts in various areas. Another challenge we faced was using the revised items for field-test that did not go through pilots. This challenge was overcome in some way when the online CEMA was tested with classroom teachers in the next phase.

Development of online CEMA

The excerpts and items were combined in two groups (group 1 and group 2) for the field testing of the CEMA online. Each group had four excerpts and associated items. None of the excerpts in each group came from the same curriculum. Each excerpt was placed on the left side of the page and associated items on the right. This helped respondents take a look at the excerpt as they responded to the items. This was an effective reminder to the respondents that the items were based on the excerpt and they had to examine the excerpt before answering to the items. The online CEMA was reviewed numerous times to correct any error or mistake that could happen during the transition from paper version to online.

Once the online CEMA was ready for field testing (see Figure 4 for a screenshot[2]), thirteen teachers and former teachers were invited to respond to the online CEMA and provide feedback in terms of changes needed or clarity of language. Their feedback was also used to check to see if the graphics and system worked well (e.g., loading of pages including images, and recording of database). They were also asked to indicate how long it took for them to complete each excerpt-item set and four entire sets in each group, in order to determine estimated time to complete the assessment. All the feedback was used to refine the online CEMA. Table 1 summarizes excerpts and questions refined and used in the field test. Most of the questions require multiple-choice responses, although a few require open-ended or yes/no responses.

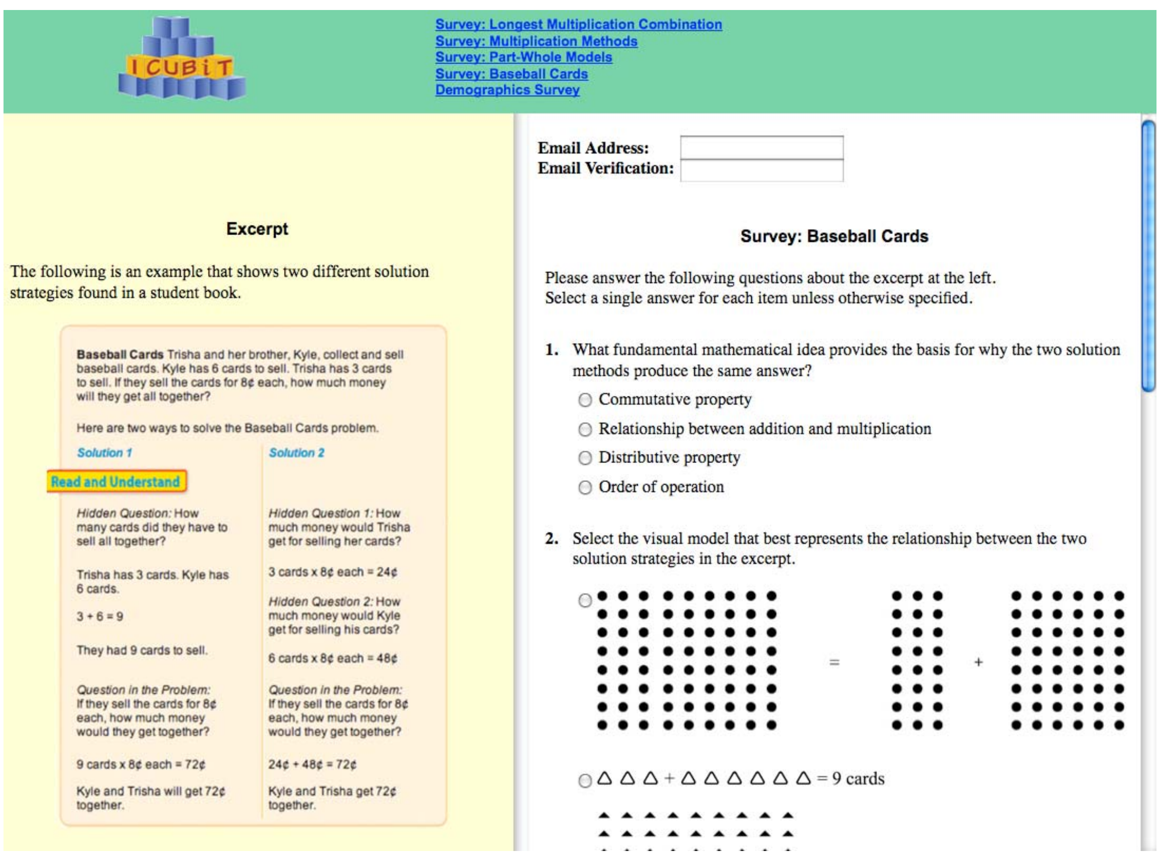

Figure 4. Screenshot of a page from online CEMA

[2] We share this only with the discussants.

Field Test and Item Analysis

We contacted math educators in various locations of the country to recruit teachers for the field testing of the assessment. We also contacted via e-mail individual teachers in various schools from a number of states. Finally, 77 teachers provided their responses to each group. These responses are currently being analyzed using Item Response Theory[3].

[3] The methods and results of the item analysis will be added to this paper when the analysis is completed.

Table 1.

Summary of excerpts and questions

Excerpt	Central mathematical idea	Curriculum	Grade	Group	Number of questions	Number of questions in dimensions*			
						Dim 1	Dim 2	Dim 3	Dim 4
Place value division	The partitive interpretation of long division algorithm using base-ten blocks	Scott Foresman *Mathematics*	4	1	6	5	0	1	5
Baseball cards	Distributive property of multiplication over addition	Scott Foresman *Mathematics*	4	2	4	4	1	0	3
Fact triangles	Commutativity and inverse operations illustrated in fact triangles	*Everyday Mathematics*	3-5	1	4	3	1	1	1
Multiplication methods	Multi-digit multiplication in partial product method, modified repeated addition method, and traditional algorithm	*Everyday Mathematics*	3-5	2	6	6	0	1	3
Longest multiplication	Prime factorization and multiples of numbers	*Investigations*	5	1	6	4	3	0	2
Array model	Multiplying by multiples of 10 and its representations	*Investigations*	4	2	4	4	2	0	3
Square numbers	Patterns associated with square numbers	*Math Trailblazers*	5	1	6	5	3	2	4
Part whole model	Additive and/or multiplicative number relationships represented by part whole models	Singapore Mathematics	4	2	5	4	0	2	4
Total					41	35	10	7	25

* Note that one question can address more than one dimension.

Issues and Challenges

Conceptualizing and assessing *Curriculum Embedded Mathematics Knowledge* was a complex and challenging task. Complexities of the work came from both the difficulty of conceptualizing the kind of teacher knowledge and the challenges of developing an assessment that measures the kind of knowledge conceptualized. There was no existing framework to guide us in this work. In our work, we found the CEMA itself served to develop a conceptual framework. To inform related future work, we discuss specific issues and challenges that we faced while we were conceptualizing and developing the CEMA.

Complexities of the work can be explained by the fact that we had to revisit the framework and the assessment a number of times to refine them. The process needed more rounds of refinement than we originally thought. We had to define the kind of knowledge we intended to assess as well as develop a measure, each process informing and influenced by the other. Especially, the field did not have a clear understanding of the kind of knowledge that Ball, et al. referred to KCC, which was most related to CEMK that we conceptualized. Even though Sleep's elaboration of KCC guided our conceptualization of CEMK, we found that it was still broad, especially when we attempted to measure it. Thus, our work on conceptualizing and assessing CEMK needed repeated check and refinement through somewhat cyclic procedures. More teacher involvement in various phases was needed in this complex process, to gather additional data to confirm our conceptualization and procedures before moving to the next step. For example, teacher interviews with a familiar curriculum and a novel one helped us refine and confirm the four dimensions of CEMK from an empirical sense. In addition, teacher feedback during pre-pilot and during the online CEMA testing enabled us to refine the assessment for pilots and field test, respectively.

Creating items based on common grounds was a challenge as well as a learning opportunity for us. The CEMA should not interfere mathematical integrity in any way, particularly because this assessment measures teacher knowledge of mathematics. When the assessment was externally reviewed, mathematicians' comments on a couple of items pointed out this issue. We went by teachers' responses to create multiple choices and mathematicians were uncomfortable in the way we used mathematical terms. For example, one of the excerpts in the assessment illustrated fact triangles, each of which can show either addition and subtraction facts or multiplication and division facts at the same time. One question asks, "*What does the fact triangle illustrate by introducing multiplication and division at the same time*?" We created four multiple choices based on teacher responses from the pilots. One of the incorrect choices was "*It shows the equivalency of multiplication and division*." The rationale for this choice was making this incorrect one look more appealing by using the mathematical term, equivalence. A mathematician commented that even though she was not sure what it means for two arithmetic operations to be equivalent, multiplication and division, in fact, are equivalent in the sense that $a \div b = c$ if and only if $a = b \times c$ ($b \neq 0$). In the end, we created a new choice to replace this one, in which mathematicians found no problem.

Creating items based on common grounds also means that the correct answers of the items must not vary depending on interpretation. During the pilots, we found that teachers saw things differently and used their own interpretations and reasoning to support their answers. This led us to revise items and multiple choices in a way to eliminate potential interpretations of the items. We clearly pointed out or underlined things that should be used (to give a particular direction to view the questions) to answer to the given questions. For example, on a question asking to rank the order of difficulty in three word problems (Dimension 3), many teachers determined the order based on types of computation (i.e., addition, subtraction, etc.) while the intent was complexity of the part-whole models representing those word problems. Therefore, the question was revised to make this point clear.

In fact, it was very hard to create items addressing Dimension 2 or 3. As shown in Table 1, the number of items in these dimensions was about a third of the number of items in the other dimensions. The difficulty in creating items for Dimensions 2 and 3 was in part because various interpretations were possible. For Dimension 2, when the item is about prerequisites to certain concepts or ideas, it is hard to determine why a particular one is absolutely needed; it is hard to argue why others do not work. The development of mathematical concepts is not linear; various earlier concepts are related to current and future learning. Likewise, in Dimension 3 often it is hard to determine why one is definitely more difficult than others. Various factors contribute to problem difficulty; each problem can include a combination of different factors. Because of these reasons, we had to eliminate many of the items we created for Dimension 2 or 3. This does not mean that the two dimensions are loosely defined. Rather, it illustrates the difficulty of creating items in these dimensions in part because of their nature and in part because of room for interpretations as many interrelated factors are embedded in them.

Coordinating important issues related was another challenge. Throughout the process, we had to maintain mathematical precision, to address pedagogical importance, and to meet measurement criteria. Incorporating all these aspects into the assessment was a difficult task. For example, an expression commonly used by classroom teachers may not convey a precise meaning from a mathematical point of view, whereas identifying appropriate multiple-choice responses for the sake of measurement may lead to incorrect mathematics. The earlier example of equivalency illustrates this difficulty to some degree. We lost mathematical precision while we were trying to fulfill the other two aims. To get the final set of the items for the field test, we revised them numerous times to coordinate all important aspects appropriately.

A final issue we continue to grapple with involves determining the scope of the assessment in terms of the mathematics content. What is the level of mathematics the CEMA (or a similar assessment) should address? How much mathematics beyond grades 3-5 should be included in the CEMA? More specifically, should the CEMA include generalized mathematical expressions even though teachers do not teach them in their grade levels? This was a debate among us when we generated questions and multiple choices. A more fundamental question is, what is the mathematics that elementary teachers need to know? Elementary mathematics has seldom been unpacked. From our perspective, identifying what it takes to understand the mathematics in the elementary curriculum materials is one important task to answer to the questions raised above.

Closing Comments

Teacher knowledge is an important factor to influence how teachers read and use curriculum materials to plan a lesson and make instructional decisions. Through item development, this study conceptualized a specific kind of teacher knowledge - knowledge of the mathematics embedded in tasks and activities in curriculum materials. Currently, little is understood about this kind of knowledge. By elaborating and assessing it we hope to facilitate future research on teacher-curriculum interaction and effective curriculum use. We also expect that it will contribute to teacher education and curriculum design in general.

References

Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group*. Edmonton, AB: CMESG/GCEDM.

Brown, M. W. (2002). *Teaching by design: Understanding the interactions between teacher practice and the design of curricular innovation*. Unpublished doctoral dissertation. Northwestern University, Evanston, Illinois.

Ball, D. L., Lubienski, S.T., & D. S. Mewborn. (2003). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-456). New York: Macmillan.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389-407.

Carter, K. (1990). Teachers' Knowledge and Learning to Teach. In W.R Houston, M. Haberman, & J. P. Silkula (Eds.), *Handbook of research on teacher education*. New York: Macmillan.

Clandinin, D. J., & Connelly, F. M. (1992). Teacher as curriculum maker. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 363-401). New York: Macmillan.

Grossman, P. (1991). Mapping the terrain: Knowledge growth in teaching. In H. C. Waxman & H. J. Walberg (Eds.), *Effective teaching: Current research* (pp. 203-215). Berkeley, CA: McCutchan.

Hill, H., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education, 35*, 330-351.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing Teachers' Mathematical Knowledge: What knowledge matters and what evidence counts? In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning*. New York: Macmillan.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum.

Remillard, J. T. (2000). Can curriculum materials support teachers' learning? *Elementary School Journal, 100*, 331-350.

Remillard, J. (2005). Examining key concepts of research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

Remillard, J.T., Lloyd, G.M., & Herbel-Eisenmann, B.A. (Eds.) (2008). *Mathematics Teachers at Work: Connecting Curriculum Materials and Classroom Instruction.* New York: Routledge.

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

Sleep, L. (2009). *Teaching to the mathematical point: knowing and using mathematics in teaching*. Unpublished dissertation in the University of Michigan.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319-369). New York: Macmillan.

Tirosh, D., & Graeber, A. O. (1989). Preservice elementary teachers' explicit beliefs about multiplication and division. *Educational Studies in Mathematics, 20*, 79-96.

Tirosh, D., & Graeber, A. O. (1990). Evoking cognitive conflict to explore preservice teachers' thinking about division. *Journal for Research in Mathematics Education, 21*(2), 98-108.

Wilson, S., Shulman, L., & Richert, A. (l987). "150 different ways" of knowing: Representations of knowledge in teaching. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 104-124). London: Cassell.