Developing Measures of Content Knowledge for Teaching Reading

Geoffrey Phelps
Stephen Schilling
University of Michigan

The measures development work is being conducted as part of the Study of Instructional Improvement. The research group includes, Geoffrey Phelps, Deborah Loewenberg Ball, Heather C. Hill, and Stephen Schilling. This work is supported by NSF grant REC-9979873, by a subcontract to CPRE on Department of Education (DOE), Office of Educational Research and Improvement (OERI) award #R308A960003, by University of California Office of the President award #8047PCE186, and by the Atlantic Philanthropies. The opinions expressed herein are those of the authors and not of the funding agencies. The authors acknowledge Deborah Loewenberg Ball, Joshua Glazer, Heather Hill, and two anonymous reviewers for helpful comments on earlier drafts of this paper.

Correspondence concerning this article should be addressed to Geoffrey Phelps, Study of Instructional Improvement, 610 South University Ave, School of Education Building #3112, Ann Arbor, Michigan 48109; e-mail gphelps@umich.edu.
Abstract

In this article we present results from a project to develop survey measures of the content knowledge teachers need to teach elementary reading. In areas such as mathematics and science, there has been great interest in the special ways teachers need to know the subject to teach it to others — often referred to as pedagogical content knowledge. However, little is known about what teachers need to know about reading to teach it effectively. We begin the article by discussing what might constitute content knowledge for teaching in the area of reading and by describing the items we wrote. Next, factor and scaling results are presented from a pilot of 261 multiple-choice items with 1,542 elementary teachers. We found that content knowledge for teaching reading includes multiple dimensions, defined both by topic and by how teachers use knowledge in teaching practice. Items within these constructs form reliable scales.
This is a time marked by unprecedented interest in teacher quality. Universities and professional development programs are looking for ways to assess teacher learning. Policymakers argue for tests of teacher competency. Educational researchers are studying teacher knowledge and its effects on instruction and student achievement. Consequently, there is a significant need to develop high-quality measures of instruction and teacher knowledge.

Teacher content knowledge – how teachers need to know a subject to teach it to others – is one area that has received special attention. Researchers have often characterized teacher content knowledge by using proxy measures, such as counts of the number of college courses taken, or by administering mathematics or other subject matter tests. However, these approaches tend to consider only common knowledge, not knowledge specific to teaching. To date there are few tools available for measuring the specialized types of content knowledge needed for teaching a particular subject.

Our work focuses on content knowledge for teaching in the area of elementary reading. Limited attention has been devoted to exploring what teachers need to know about reading to teach students to read and even less to developing measures of this knowledge. We begin by considering what might constitute content knowledge for teaching in the area of reading. In the second section of the paper we present results from a multi-year effort to develop survey measures of this knowledge.

Content Knowledge for Teaching

In areas such as mathematics and science, developing teacher content knowledge has been a major concern because it is generally accepted that teachers who know these subjects are better able to teach them. However, studies seeking to establish connections between teachers’ content knowledge and student achievement are far from conclusive (Ball, Lubienski, & Mewborn, 2001; Begle, 1979). One possible explanation is that teachers need to know content in ways that differ from what is typically taught and learned in university courses.

In the 1980’s, Lee Shulman and his colleagues popularized the concept of “pedagogical content knowledge” and introduced a new way of thinking about the nature and role of the content knowledge needed for teaching (Shulman, 1986, 1987; Wilson, Shulman, & Richert, 1987). Pedagogical content knowledge includes, “the most useful forms of representation …, the most powerful analogies, illustrations, examples, explanations, and demonstrations—in a word, the most useful ways of representing and formulating the subject that make it comprehensible to others…. Pedagogical content knowledge also includes an understanding of what makes the learning of specific topics easy or difficult” (Shulman, 1986, p.7).

Pedagogical content knowledge provided the intellectual impetus for new studies investigating how teachers need to know the subject they teach (Ball, 1988, 1991; Ball et al., 2001; Gess-Newsome & Lederman, 1995; Grossman, 1990, 1991;
Leinhardt & Smith, 1985; Magnusson, Krajcik, & Borko, 1999; Wilson & Wineburg, 1988; Wineburg & Wilson, 1991). These studies all foreground the importance of content in teaching; the discipline, and what there is to be taught and learned about a subject, provides a conceptual foundation for what teachers need to know (Bruner, 1960; Schwab, 1978). However the notion of pedagogical content knowledge goes further, identifying specialized ways that teachers need to know a subject to teach it to others. It provides a basis for differentiating common knowledge of content from the specialized knowledge of content that is the unique province of the classroom teacher.

For example, teachers of mathematics need a special type of understanding of mathematical content itself. “A powerful characteristic of mathematics is its capacity to compress information into abstract and highly usable forms. . . . Mathematicians rely on this compression in their work. However, teachers work with mathematics as it is being learned, which requires a kind of decompression, or ‘unpacking,’ of ideas” (Ball & Bass, 2003). This “unpacked” knowledge may, in part, provide the foundation for knowing how to represent the subject to students or how to understand the mathematical features of student work. Most adults, for example, know that one can “invert and multiply” to get the correct answer to $1 \frac{3}{4}$ divided by $\frac{1}{2}$. However, mathematics teachers must know why such rules work and how to represent the mathematics to facilitate student understanding. Is a student mathematically correct in saying that this problem can be illustrated by splitting $1 \frac{3}{4}$ pies evenly between two families, or in saying that this can be illustrated by calculating how much money you would have if you doubled $1.75$? If not correct, then what is a good story problem that illustrates $1 \frac{3}{4}$ divided by $\frac{1}{2}$? Good mathematics teachers know how to address such questions and how to unpack and represent fractions in ways that are useful in teaching the subject.

These examples from mathematics, illustrate a recent direction in the study of teacher content knowledge that shifts the focus from examining teachers to examining teaching. This approach seeks to identify what teachers need to do and know in the course of teaching practice (Cohen & Ball, 1999; Lampert, 1998). Investigations taking this approach have brought to the fore new ideas about the types of content problems that arise in the course of teaching and the content-specific reasoning involved in solving such problems (Ball & Bass, 2000; Lampert, 2001). What might be the analog of this sort of content knowledge in reading? What do reading teachers need to know about language, text, and reading process, and in what ways, in order to help others learn to read?

**Content Knowledge for Teaching Reading**

In contrast to mathematics and science, the study of teachers’ content knowledge has not been a major area of inquiry in reading. There are several possible explanations. First, reading is not a discipline. There is no single group of scholars

---

1 This example has been widely discussed in the mathematics education community (Ball, 1990).
who identify what there is to be known about the subject. Indeed, it is not even clear what might count as “content” in reading. Second, while many elementary teachers lack substantial understanding of mathematics or science, fewer concerns exist about teachers’ content knowledge in the area of reading, since most teachers are competent readers. A third reason stems from reading research itself. Concern with teacher knowledge and preparation has largely centered on teaching methods, knowledge of curriculum, or the psychology of reading. Attention to content knowledge has been overshadowed by other dominant concerns and hampered by the difficulty of specifying what might constitute content knowledge in reading.

However, interest is growing in the professional knowledge that it takes to teach reading including, for example, attention to teaching methods, student learning, and curriculum (International Reading Association, 2000; Learning First Alliance/American Federation of Teachers (AFT), 1998; National Reading Panel, 2000; Pearson, 1996; Pearson, 2001; Roller, 2001; Snow, 1998). We, however, draw a distinction between studies that seek to identify what reading teachers need to know in general and studies that seek to probe the content knowledge entailed by engaging in the practice of teaching reading. Often studies stop short of investigating how teacher content knowledge impacts the capacity to teach effectively. We focus directly on what teachers know, or need to know, about the content of reading— including knowledge of text, language, reading process, and how this knowledge is used in practice.

To date, research that focuses directly on content in the area of reading has largely examined the knowledge of language and text needed to teach children to read or decode words (Brady & Moats, 1997; McCutchen, Abbott et al., 2002; McCutchen & Berninger, 1999; McCutchen, Harry, & Cox, 2002; Moats, 1994, 2000; Moats & Lyon, 1996; National Board for Professional Teaching Standards (NBPTS), 2001; Wong-Fillmore & Snow, 2002). Initial attempts to study teachers’ linguistic knowledge suggest that teachers lack understanding of specific features of language such as inflected verbs, derivational suffixes, phonemes, schwa sounds, consonant blends, morphemes, and the spelling patterns used to represent sounds (Moats, 1994). Research in this area has been spurred by mounting evidence that beginning readers benefit from explicit opportunities to learn about the language and text elements that make up words (Adams, 1990; Ehri, 1991; National Reading Panel, 2000; Share & Stanovich, 1995; Snow, 1998). If students need such instruction, so the argument goes, then teachers must possess such explicit knowledge of language and text (Learning First Alliance/American Federation of Teachers (AFT), 1998; Wong-Fillmore & Snow, 2002).

Specialized linguistic knowledge is one example of the professional knowledge needed to teach reading. Just as teachers of mathematics must have knowledge of place value to teach students to add, teachers of reading need knowledge of letters and sounds to teach students to decode words. To further explore what might constitute content knowledge for teaching reading, we turn next to an example from a classroom.

A View from the Classroom

Ms. Reynolds, an experienced first grade teacher, is working in a district that has recently revamped its early reading program. She understands, and mostly agrees with the new program. In her view it is helpful and important to teach students to recognize phonemes (i.e., the smallest units of speech in a word). She has long believed that the ability to hear and manipulate sounds in words and then to notice how these are related to letters and spelling patterns are crucial skills in learning to read. Ms. Reynolds has, however, been amazed at how difficult she herself finds many of the activities in the daily lessons. She has to work hard in many cases to notice the very distinctions among letters and sounds that she is now teaching.

Increasingly, she notices distinctions she never recognized before. For example, a few weeks ago she noticed that the letter $u$ makes more than just the long and short $u$ sound. In glancing over the class list of long $u$ words (e.g., flute, tube, cute, tune) she realized that only the $u$ in cute “says the name of $u$,” as she had always told students. The $u$ in the other words is different, more like the sound at the beginning of the word “ooze.” A few weeks before this, she noticed that the class list of words with the $t$ sound (e.g., star, hit, little, top, train, string) contained a variety of different $t$ sounds, not just the single sound she traditionally taught for the letter $t$. She realized that the $t$ in “little” can sound more like the $d$ in “middle.” In “train” the $t$ can sound more like the $ch$ in “child.” Both are different than the $t$ heard in “top” and “hit.” While Ms. Reynolds doesn’t think that teaching such subtle distinctions often makes sense, her growing sensitivity to these issues increasingly influences the specific words she chooses to focus on in instruction, her appreciation of the complexities students face in learning to read these words, and her ability to understand the errors students make while reading.

Ms. Reynolds has just completed a dictated assessment to share with other first grade teachers on her literacy team. One purpose of this exercise is to assess students’ ability to hear and represent phonemes in words. Before going to her meeting, she quickly reviews the words a student named Ron misspelled. The list includes the following words: “lidl” for little; “wit” for white; “chrane” for train; “gif” for gift; “likt” for liked; “runin” for running; “won” for one; “wl” for will; “rele” for real. She is trying to figure out if these misspellings provide evidence that Ron can accurately identify phonemes. This is particularly challenging since she needs to simultaneously make judgments about what sounds Ron is attempting to represent with his unconventional spelling and if these representations demonstrate an ability to hear each phoneme in a word.

This example highlights two points. First, there are many subtleties in the structure of words that are unfamiliar to most adults who do not teach reading. In fact, many adults may find that the ability to notice and work with phonemes is obscured by the very fluency and familiarity they bring to reading and spelling. Teachers, however, require both an “unpacked” understanding of reading and a pedagogically-useful knowledge of the content that can equip them to understand student work.
These points also apply in other areas of early reading instruction. All elementary teachers must figure out what to do when students misread words. In a class of beginning readers, students misread words many times a day. How should a teacher respond? Should she tell the student the word, point out some feature of the word, ask the student to sound out the word, compare the word to another, ask the student to consider context, or something completely different? Although these sound like pedagogical choices, it is less obvious that choosing effectively for a particular word depends, in large part, on the word itself, the type of error, and the surrounding text. The capacity to make good teaching decisions or moves rests in part on the teachers' knowledge of the subtleties of word and text structure.

These examples focus on aspects of word analysis — that is, word reading or decoding. Similar issues exist in the area of comprehension. Many educators argue that students need explicit instruction in strategies for comprehending the meaning of text (Duffy et al., 1987; Duke & Pearson, 2002; Palincsar & Brown, 1986; Pressley, 2000; Pressley et al., 1992). For example, students need to learn how to activate and use their prior knowledge to make sense of new information. One approach for teaching this is to model processes such as generating predictions. Teachers are often directed to do this by “thinking aloud,” essentially sharing their own thought process as they prepare to read a text.

On the surface this seems straightforward. But is it? Is a prediction for a narrative text similar to one for informational text? How might a teacher identify complications in a text and what might it take to design a prediction that helps students notice these features of text and make sense of them while reading? What is actually involved in modeling the process of prediction? This requires more than skill in prediction; it requires explicit awareness of the process and the ability to make it visible to students. This is neither obvious nor straightforward. Few adult readers explicitly frame predictions or consider the kinds of questions sketched above. Furthermore, only reading teachers need to see a text from the perspective of a beginning reader to identify difficulties the text might present for students.

The above discussion, coupled with emerging ideas from research on early reading instruction, provide guidance for conceptualizing content knowledge for teaching in the case of reading. Even competent adult readers require additional knowledge about text, language, and reading process in order to teach reading. This is knowledge that goes beyond being able to read well. Competent adult reading provides a striking analogy to the “compression” that characterizes how mathematicians understand mathematics (Ball & Bass, 2000). Proficient reading requires automaticity and tacit knowledge. Excessive consciousness in the form of explicit knowledge can interfere with smooth functioning. What fluent readers “know” about spelling, words, vocabulary, and text structure, for example, is seldom consciously considered. This is true both for reading words and for comprehending text. If asked, few adult readers could “unpack,” what they are doing. Even fewer could identify the knowledge of text, language, and reading process from which they draw as they read.
A Framework for Investigating Content Knowledge for Teaching Reading

What are the areas of content knowledge teachers might need in order to teach reading? And how would they need to know and use this knowledge? In this paper we examine two areas. One involves knowledge rooted in the structure of language and text. Included here is knowledge that has traditionally been developed by linguists as they study areas such as phonology, morphology, and semantics. Also included are areas of study in English literature that focus on genre, style, and other literary conventions. The second area involves knowledge rooted in reading processes—that is how readers go about reading words and comprehending text. This includes explicit understanding of what readers do, the reading strategies used in decoding and comprehending text, and the interrelated knowledge of how these processes differ according to reading ability, text type, personal background, and more.

Our analysis thus far suggests that content knowledge for reading consists of multiple domains of knowledge. Along one dimension there is knowledge of the “content” of reading. It is likely that this is not just one general area of knowledge but rather is made up of multiple topics. For example, we expect that knowledge of phonemes, as it applies to understanding the relationships between letters and sounds, differs from knowledge of prediction, as it applies to comprehending text.

Along a second dimension are distinctions in the ways teachers know these topics. For example, knowledge of phonemes and the relationships between phonemes and written language may not be sufficient knowledge for making sense of students’ decoding or spelling errors. Just knowing about the importance of prior knowledge and having the ability to frame predictions while reading may not be sufficient for knowing how to represent a useful prediction for a particular text or student. Teachers need knowledge that is usable in the varied contexts, interactions, and practices that define the work of teaching reading.

Developing Measures of Content Knowledge for Teaching Reading

In 2001, the Study of Instructional Improvement initiated a project to develop a large number of content knowledge for teaching items that built on and expanded nearly two years of prior development (Rowan, Schilling, Ball, & Miller, 2001). The primary goal was to create a new type of measure that could be used to study teacher content knowledge and the effects of this knowledge on instruction and student learning. These measures needed to provide a basis for studying teacher knowledge

---

2 The Study of Instructional Improvement seeks to understand the impact of improvement programs on language arts and mathematics instruction and student performance in high poverty elementary schools. Detailed information on the process of item development, considerations that drove item selection, and sample release items can be found on the Study of Instructional Improvement web site at http://www.sii.soe.umich.edu.

across the range of teaching practices common in the elementary grades. Rather than identifying and focusing on a narrowly defined area of the curriculum, or a particular approach to teaching, we set out to generate a general measure of content knowledge for teaching elementary reading.

Over a dozen writers, including leading scholars, teacher educators, and experienced classroom teachers, participated in the project. The items present respondents with situations encountered in the work or practice of teaching reading (e.g., evaluating tasks or curriculum materials, choosing appropriate questions to ask in instruction, responding to a student question or problem, or explaining student progress to a colleague or parent) (Ball & Bass, 2003). This places the focus on identifying the knowledge teachers use in their work. If successful, items directly measure the content knowledge most likely to be associated with the capacity to teach reading effectively. Furthermore, by starting with practice, the development effort is grounded in examples of what teachers do in teaching reading. This, in turn, provides a context for discovering the kinds of knowledge entailed by these practices. Rather than starting with curriculum topics or content organizations from the disciplines, we seek to build both theory and measures based on specific examples of the knowledge that teachers use in their practice.

Our research was guided by three preliminary distinctions in content knowledge for teaching reading: knowledge of content (KC), knowledge of students and content (KSC), and knowledge of teaching and content (KTC). The primary difference between items in each of these categories is in how content knowledge is related to the work of teaching.

Knowledge of teaching and content (KTC) items require respondents to use knowledge of reading to develop or choose teaching actions or moves. This may require using knowledge of content to determine what to say to a student struggling over a difficult word or what type of task to give a student to help her comprehend a challenging passage. Items that tap knowledge of students and content (KSC) require respondents to use knowledge of reading to decipher and interpret students’ products or work. For example, this may entail recognizing typical student errors or approaches to reading. Items in this category do not require making teaching decisions. Knowledge of content (KC) items require respondents to use knowledge of reading in the context of teaching situations. While these items may require a specialized knowledge of content beyond what most adult readers typically understand, they do not require knowledge of students or teaching.

For example, recall Ms. Reynolds’ work on teaching phonemes. She is wrestling with her own knowledge of phonemes as she attempts to identify subtle distinctions in the sounds that make up words. This is just one example of the...
knowledge of reading content (KC) that teachers draw on in the practice of teaching reading. However, Ms. Reynolds needs to do more than just identify phonemes in words; she also needs to make sense of how a student named Ron understands and represents phonemes in his writing. To do this, Ms. Reynolds must draw on an amalgam of her knowledge of phonemes and typical patterns in student work, and then use this knowledge to interpret Ron’s writing (KSC). Ms. Reynolds may also need to teach Ron, or other students in her class, to identify and manipulate phonemes. To do this she will have to decide what to work on with students, including selecting the most useful sounds, words, and ways of representing these (KTC).

In addition to distinctions in the types of content knowledge for teaching noted above, we also sought to measure content knowledge in two major topic areas, comprehension (CMP) and word analysis (WA). These two global distinctions were selected because together they represent a wide range of elementary content and each roughly maps onto a major focal area in the elementary curriculum. For the purpose of guiding item development, we include within comprehension, morphology, vocabulary, comprehension strategies and questions, genre, fluency, and other topics related to comprehending the meaning of words and text. Word analysis includes, phonemic awareness, letter sound relationships, word frequency, and other topics related to the reading and decoding of words and their print and sound elements. For examples of items representing the distinctions discussed above see the Appendix.

Taken together, these two major topic areas and the three types of content knowledge form a two by three matrix defining six domains of content knowledge for teaching reading. These domains are broadly conceived and encompass a wide range of possible reading content. While neither these domains nor the items we wrote fully represent the complex subject of reading, these categories do provide a useful basis for investigating if there are major distinctions in teachers’ content knowledge.

**Method**

The item development project discussed above was designed both to advance our understanding of content knowledge for teaching reading and to generate a large pool of items for a pilot project. In summer 2001, 261 items were piloted in the California Professional Development Institutes. These institutes were part of a large-scale, publicly funded effort in California to increase the knowledge of elementary reading teachers. Because of the large number of items, three separate forms were used, each including roughly equal numbers of items in the domains we sought to measure. Forms were administered to a total of 1,542 teachers participating in 23 separate week-long, summer institutes. Results from this pilot were analyzed using factor analysis, classical test theory and item response theory to identify the best items for further analysis and inclusion in a potential scale of teacher knowledge.
In spring 2002, 599 teachers, the majority of whom also participated in the summer 2001 California Professional Development Institutes, completed a follow-up survey. Data from this survey provides the basis of the analysis discussed in this section. Items on the follow-up survey were selected from the previously piloted items. Selection was driven by the twin considerations of psychometric properties of items and equal representation of items across the domains. The survey includes a total of 77 items: ten situations or scenarios including 42 separate items in the area of comprehension and six situations including 35 separate items in the area of word analysis. We were successful in writing items for only five of the six specified domains. Knowledge of teaching and content items in the area of word analysis are not included in the survey. Brief descriptions are provided in Table 1.

### TABLE 1

**Content Knowledge for Teaching Reading Situations**

<table>
<thead>
<tr>
<th>Knowledge of Content</th>
<th>Knowledge of Students and Content</th>
<th>Knowledge of Teaching and Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comprehension</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of:</td>
<td>Interpreting student:</td>
<td>Teaching students to:</td>
</tr>
<tr>
<td>1. Antonyms (5)</td>
<td>4. Reading to assess comprehension (6)</td>
<td>7. Use the structure of a word (e.g., prefixes, roots) to understand word meaning (5)</td>
</tr>
<tr>
<td>2. Comprehension questions (6)</td>
<td>5. Reading error rate to determine appropriate text difficulty (1)</td>
<td>8. Increase reading fluency (4)</td>
</tr>
<tr>
<td>3. Prefixes (5)</td>
<td>6. Ideas about genre (1)</td>
<td>9. Correct word substitutions (4)</td>
</tr>
<tr>
<td>4. Reading to assess comprehension (6)</td>
<td>5. Reading error rate to determine appropriate text difficulty (1)</td>
<td>9. Correct word substitutions (4)</td>
</tr>
<tr>
<td>5. Reading error rate to determine appropriate text difficulty (1)</td>
<td>6. Ideas about genre (1)</td>
<td>10. Determine meaning of unknown words (5)</td>
</tr>
</tbody>
</table>

| **Word Analysis**    |                                  |                                  |
| Knowledge of:        | Interpreting student:            |                                  |
| 11. Phonemes (7)     | 14. Spelling to assess phoneme knowledge (6) |                                  |
| 12. Spelling regularity (6) | 15. Spelling to assess difficulties with spelling patterns (7) |                                  |
| 13. Word frequency (6) | 16. Reading to assess why students misread particular words (3) |                                  |

**Note:** The number of items within a situation is listed in parenthesis.

---

4 Any question requiring a response is counted as one item. In many cases a situation or scenario includes more than one item. For example, in the Appendix, situation #1 includes five separate items and situation #2 includes six items. However, situation #3 includes just one item.
In the psychometric analysis that follows, we seek to answer two larger questions:

- What dimensions effectively characterize content knowledge for teaching reading?
- Is it possible to develop reliable measures of these dimensions?

There are multiple structures that could characterize content knowledge for teaching reading. For example, each situation in Table 1 might represent a distinct domain requiring a distinct measure; alternatively, distinctions among the various situations may be irrelevant, with all items forming a single factor. These different structures have implications for how teachers might hold or use knowledge of reading. Evidence that items form a single factor would suggest that content knowledge for teaching reading is defined only by common knowledge—that is, what is typically known by competent adult readers. On the other hand evidence of multidimensionality suggests that there are distinct domains of knowledge used in teaching reading.

We expect to find that content knowledge for teaching reading has some degree of multidimensionality. One possibility is that it varies only by knowledge type, and that there are no measurable distinctions between topics such as word analysis and comprehension. In this case, content knowledge for teaching reading would include knowledge of content, knowledge of students and content, knowledge of teaching and content, or some other organization of knowledge type. A second possibility is that there will be no distinction along the dimension of knowledge type, with items organized only according to the topics of word analysis and comprehension, or perhaps even finer grained topics. Still another possibility is that content knowledge for teaching reading may have distinctions in both topic and type of knowledge.

Results

To test these possibilities, we conducted exploratory factor analysis (Thissen & Wainer, 2001), using ORDFAC, a specialized factor program that allows for analysis of ordinal data (Schilling, 2002). In order to control for testlet effects, all items within a situation were scored correct or incorrect and then summed to provide a single ordinal score for each situation or testlet. For in-depth discussion of testlet coding and other technical issues see (Schilling & Phelps, 2003). Table 2 presents chi-square tests of fit of successive models along with Akaike information criterion (AIC). The content knowledge for reading items are clearly fit by three factors, as the three factor model exhibits the lowest AIC, a significant chi-square compared to the two factor model, and a non-significant chi-square compared to the four factor model.
TABLE 2
Comparison of Exploratory Factor Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>-10606</td>
<td></td>
<td></td>
<td></td>
<td>21398</td>
</tr>
<tr>
<td>2 Factor</td>
<td>-10579</td>
<td>54.66</td>
<td>15</td>
<td>0.0000</td>
<td>21373</td>
</tr>
<tr>
<td>3 Factor</td>
<td>-10560</td>
<td>37.18</td>
<td>14</td>
<td>0.0007</td>
<td>21364</td>
</tr>
<tr>
<td>4 Factor</td>
<td>-10555</td>
<td>9.10</td>
<td>13</td>
<td>0.7653</td>
<td>21381</td>
</tr>
</tbody>
</table>

Factor loadings for the 16 situations are given in Table 3. Examination of the factor loadings identifies three clearly defined factors:

- Comprehension / Knowledge of Content (KC)
- Comprehension / Knowledge of Teaching and Content (KTC)
- Word analysis / Knowledge of Content (KC)
<table>
<thead>
<tr>
<th>Situation</th>
<th>Hypothesized Domain</th>
<th>Comprehension Knowledge of Content (CMP/ KC)</th>
<th>Factor Knowledge of Teaching and Content (CMP/ KTC)</th>
<th>Word Analysis Knowledge of Content (WA/ KC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Antonyms</td>
<td>CMP/KC</td>
<td>0.550</td>
<td>-0.011</td>
<td>-0.034</td>
</tr>
<tr>
<td>2. Comprehension questions</td>
<td>CMP/KC</td>
<td>0.430</td>
<td>0.115</td>
<td>0.113</td>
</tr>
<tr>
<td>3. Prefixes</td>
<td>CMP/KC</td>
<td>0.494</td>
<td>-0.017</td>
<td>0.065</td>
</tr>
<tr>
<td>Interpreting student:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Reading to assess comprehension</td>
<td>CMP/KSC a</td>
<td>0.348</td>
<td>0.266</td>
<td>0.065</td>
</tr>
<tr>
<td>5. Reading error rate to determine appropriate text difficulty</td>
<td>CMP/KSC</td>
<td>0.450</td>
<td>0.005</td>
<td>0.150</td>
</tr>
<tr>
<td>6. Ideas about genre</td>
<td>CMP/KSC</td>
<td>0.418</td>
<td>0.214</td>
<td>0.051</td>
</tr>
<tr>
<td>Teaching students to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Use the structure of a word to understand word meaning</td>
<td>CMP/KTC</td>
<td>0.444</td>
<td>0.373</td>
<td>-0.200</td>
</tr>
<tr>
<td>8. Increase reading fluency</td>
<td>CMP/KTC</td>
<td>0.034</td>
<td>0.481</td>
<td>-0.046</td>
</tr>
<tr>
<td>9. Correct word substitutions</td>
<td>CMP/KTC</td>
<td>0.095</td>
<td>0.529</td>
<td>0.022</td>
</tr>
<tr>
<td>10. Determine meaning of unknown words</td>
<td>CMP/KTC</td>
<td>-0.044</td>
<td>0.419</td>
<td>0.059</td>
</tr>
<tr>
<td>Knowledge of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Phonemes</td>
<td>WA/KC</td>
<td>-0.064</td>
<td>0.076</td>
<td>0.613</td>
</tr>
<tr>
<td>12. Spelling regularity</td>
<td>WA/KC</td>
<td>0.267</td>
<td>-0.149</td>
<td>0.449</td>
</tr>
<tr>
<td>13. Word frequency</td>
<td>WA/KC</td>
<td>0.145</td>
<td>-0.061</td>
<td>0.481</td>
</tr>
<tr>
<td>Interpreting student:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Spelling to assess phoneme knowledge</td>
<td>WA/KSC b</td>
<td>-0.091</td>
<td>0.288</td>
<td>0.442</td>
</tr>
<tr>
<td>15. Spelling to assess difficulties with spelling patterns</td>
<td>WA/KSC</td>
<td>0.119</td>
<td>-0.016</td>
<td>0.576</td>
</tr>
<tr>
<td>16. Reading to assess why students misread particular words</td>
<td>WA/KSC</td>
<td>-0.015</td>
<td>0.228</td>
<td>0.392</td>
</tr>
</tbody>
</table>

Note: Situations loading at greater than 0.2 are considered as defining factors and are indicated in bold type. Secondary factor loadings are presented in italics.

a Comprehension / Knowledge of Students and Content
b Word Analysis / Knowledge of Students and Content
Without exception, all situations designed to measure knowledge of content (KC) in comprehension and in word analysis loaded on the respective factor. Only situation #12 had a partial loading on another factor.

Three of the four comprehension situations designed to measure knowledge of teaching and content (KTC) loaded on this factor. Only situation #7 had a primary loading on the knowledge of content (KC) factor. Note, however, that this situation had a moderate secondary loading on the KTC factor, suggesting that it also measures the same underlying trait as the other three KTC situations.

The hypothesized domain of knowledge of students and content (KSC) is indistinguishable from the knowledge of content (KC) and knowledge of teaching and content (KTC) factors. All KSC situations in the area of comprehension loaded on the comprehension knowledge of content (KC) factor and those in the area of word analysis loaded on the word analysis knowledge of content (KC) factor. There are, however, indications that situations written to require respondents to use KSC differ from the situations that do not ask about students. Note that four out of the six KSC situations (#4, 6, 14, and 16) had secondary loadings on the comprehension knowledge of teaching and content (KTC) factor. It appears, that while the KSC situations are primarily measuring knowledge of content, they also measure an underlying trait in common with using knowledge of content in making teaching decisions. This underlying trait is even robust across the topic areas of word analysis and comprehension, with KSC items in both word analysis and comprehension loading on the comprehension knowledge of teaching and content (KTC) factor.

To further investigate the factor structure underlying the items, we fit a series of confirmatory factor models using ORDFAC: 1) a one factor model; 2) a three factor, simple structure model, where items load on one and only one of the three factors identified in the exploratory analysis; and, 3) a five factor, simple structure model, where items load on one and only one factor corresponding to each of the hypothesized domains identified in Table 1. A comparison of these models and the exploratory model using the AIC criterion are presented in Table 4. The three factor, simple structure model fits best by the AIC criterion, providing additional evidence that we have identified three dimensions that could function as scales.

Table 4:

<table>
<thead>
<tr>
<th>Models</th>
<th>Log-likelihood</th>
<th>DF</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>-10606</td>
<td>93</td>
<td>21398</td>
</tr>
<tr>
<td>3 Factor Simple</td>
<td>-10586</td>
<td>96</td>
<td>21364</td>
</tr>
<tr>
<td>5 Factor Simple</td>
<td>-10610</td>
<td>103</td>
<td>21426</td>
</tr>
<tr>
<td>3 Factor Exploratory</td>
<td>-10560</td>
<td>122</td>
<td>21364</td>
</tr>
</tbody>
</table>

5 Situations are grouped as follows: 1-6, 7-10, and 11-16.
To investigate individual item characteristics and the feasibility of creating measures for each of the three factors, items were further analyzed using two parameter item response theory models (Hambleton, Swaminathan, & Rogers, 1991). BILOG (Mislevy & Bock, 1997) was used to analyze item properties and build scales. Individual items were placed on their respective scale based on their parent situation or testlet\(^6\). No single item misfit on any of the three scales. Classical test theory reliabilities (coefficient alpha) and IRT summary statistics are presented in Table 5.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Situations</th>
<th>Items</th>
<th>Coefficient Alpha</th>
<th>IRT Reliability</th>
<th>Test Information Curve Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension/ Knowledge of Content</td>
<td>6</td>
<td>24</td>
<td>0.754</td>
<td>.774</td>
<td>- 2.1</td>
</tr>
<tr>
<td>Comprehension/ Knowledge Teaching and Content</td>
<td>4</td>
<td>18</td>
<td>0.670</td>
<td>.725</td>
<td>0.8</td>
</tr>
<tr>
<td>Word Analysis/ Knowledge of Content</td>
<td>6</td>
<td>35</td>
<td>0.822</td>
<td>.846</td>
<td>-1.8</td>
</tr>
</tbody>
</table>

For each of the three factors, it was possible to create reliable scales. Each scale had IRT estimated reliabilities above 0.7, with coefficient alphas ranging from 0.67, for the 18 item, comprehension/knowledge of teaching and content scale, to 0.82, for the 35 item, word analysis/knowledge of content scale.

Item response theory also provides test information curves to indicate where tests are most effectively targeted. It is noteworthy that in addition to overall reliability, the scales differed in this respect. Figure 1 presents the test information curves for the two comprehension scales. The knowledge of content (KC) scale is effective at the low end of the population, with a peaked information curve at -2.1 standard deviations below the mean. However, this scale is markedly less effective at the high end of the population. In contrast, the knowledge of teaching and content (KTC) scale has a flatter profile and a maximum information curve at 0.8 standard deviations above the mean. This indicates that the KTC scale has good properties throughout a wide range in the population and is well suited for identifying more knowledgeable participants. This has potentially important implications for practice, since it suggests that measures of content knowledge for teaching may need to include KTC items in order to identify those teachers who are more knowledgeable.

---

\(^6\) CMP/KC includes situations 1-6. CMP/KSC includes situations 7-10, and WA/KC includes situations 11-16. See Table 1.
FIGURE 1. Test Information Curves for Comprehension Knowledge of Content and Knowledge of Teaching and Content
Discussion

The factor structure indicates that there are unique domains that make up content knowledge for teaching in the area of reading. There is a clear topical distinction between comprehension and word analysis. Indeed, these two domains are remarkably distinct, with only one of the six situations loading partially on the other factor. Content knowledge for reading is not defined by a general area of knowledge or general reading ability, but instead is made up of multiple domains. This suggests that teachers need to develop knowledge across at least a number of content knowledge dimensions to teach reading.

The factor results further indicate, at least in the area of comprehension, that there are differences in teachers’ understanding of the content of reading specific to the context of the work they do. There is a distinct area measured by the knowledge of content (KC) situations and a second distinct area measured by the knowledge of teaching and content (KTC) situations. It seems plausible that if we had managed to write analogous KTC items in the area of word analysis, these items would define a fourth factor. This finding is important since it strongly suggests that content knowledge for teaching has not only domains related to topical knowledge, but also domains that are more directly rooted in the work of teaching, what could be called pedagogical content knowledge for teaching reading.

One surprise is that the knowledge of content (KC) and knowledge of teaching and content (KTC) domains are distinct. We had expected that answering KTC items would also entail use of knowledge of content more generally. But three of the four situations that define the KTC factor in comprehension load only on this factor. This suggests, that the knowledge of comprehension used in teaching is distinct from more common knowledge of reading.

We were unable to measure domains for knowledge of students and content (KSC) in either comprehension or word analysis. One possible explanation is that content knowledge for teaching does not have distinctions that are as finely drawn as initially expected. Instead of three types of content knowledge, there are only two — one defined by knowledge of content and the second by knowledge of teaching and content. KSC items can be answered by drawing on both types of knowledge. Another possible explanation is that there is a separate KSC domain, but either the particular items were poorly suited for measuring this domain or this knowledge was indistinguishable for the population of respondents that took the survey.

We also set out to investigate whether the items could be used to generate reliable measures for each of the three factors. The answer is clearly yes. But these measures have limitations. The knowledge of content (KC) measures in both word analysis and comprehension are relatively easy. These measures are best suited for differentiating among teachers with less knowledge of reading. These scales could be improved by including more difficult items. The knowledge of teaching and content (KTC) scale in comprehension, on the other hand, is relatively more difficult. This scale is best suited for differentiating the knowledgeable teachers from those with less
knowledge. This scale could be improved by including easier items that measure the knowledge of content teachers use to teach comprehension.

These findings are encouraging. There is evidence for content knowledge in the area of reading. The results support the argument that teachers need to develop understanding of linguistic features of language and text, while providing evidence suggesting that analogous research is warranted in the area of comprehension. The factor results also support the argument that teaching reading like teaching mathematics and science, requires specialized knowledge of content.

At the same time the potential of this work has not been fully realized. The item development team was unable to write usable word analysis items in the knowledge of teaching and content domain. To answer the question of whether knowledge of teaching and content is a single domain of knowledge or whether it includes distinct areas of knowledge in word analysis and comprehension, it is necessary to develop items that have the potential of measuring each of these domains. Furthermore, in order to better measure the domains that have been identified, there is a need for items that cover the full spectrum of ability.

There are also many facets of content knowledge for teaching not addressed in our current work. A broader based approach to item development would serve to generate a greater range of items leading to more comprehensive measures. This would also provide a basis for exploring the topical structure of content knowledge for teaching reading within the relatively broad distinctions of word analysis and comprehension. In short, there is a need to develop a second generation of items that address issues raised in our analysis and that move the project forward to include a wider range of content.

In addition to developing and piloting more items with the goal of improving the quality of the measures, it is important to study the validity of the measures. One approach, currently underway as part of this item development work, is to interview teachers to better understand the types of knowledge used in answering the items (Phelps, 2003). This will provide a useful basis for understanding the knowledge measured by these items, why some items and scales are difficult and others easy, and in general how sensitive the items are to a range of teacher knowledge. These validation efforts should be followed by studies of classroom teaching to explore how differences in teacher knowledge are associated with differences in teaching practice.

Equally important is to begin to use these measures to examine teacher knowledge. Much can be learned by investigating how this knowledge varies among teachers. Do more experienced teachers know more? How do non-teachers and prospective teachers compare? Is there evidence that the measures are sensitive to the differences one would expect across these populations? Are differences in teacher knowledge associated with student achievement? Answers to such questions are substantively important in their own right and also will illuminate the types of knowledge measured by these scales.

There are compelling reasons to devote the resources needed to accomplish these goals. An important long-term objective is to develop tools suitable for studying
content knowledge for teaching, its growth, and its contribution to both improvement in instruction and student achievement. Addressing these problems can move the education field forward in the capacity to understand and promote teacher quality.
References


APPENDIX

Sample Content Knowledge For Teaching Reading Items

1. Comprehension / Knowledge of Content

The Partridge and the Fowler

A Fowler caught a Partridge and was about to kill it. The Partridge earnestly begged him to spare his life, saying, "Pray, master, permit me to live and I will entice many Partridges to you in recompense for your mercy to me." The Fowler replied, "I shall now with less scruple take your life, because you are willing to save it at the cost of betraying your friends and relations."

To assess his students' understanding of The Partridge and the Fowler, Mr. Hamada asks them to work in small groups to select a moral for this fable. He provides a list of possible morals. Which choices capture the meaning of this fable? (Mark YES, NO, or I'M NOT SURE for each choice.)

<table>
<thead>
<tr>
<th>Choice</th>
<th>Yes</th>
<th>No</th>
<th>I'm not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Birds of a feather flock together.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>b) One can not escape one’s own evil deeds.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>c) It is better to take the life of the wicked than the benevolent.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>d) The gods help those that help themselves.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>e) The hero is brave in deeds as well as words.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

7 Items copyright 2001. These items not for use or reproduction without written consent of the Study of Instructional Improvement.
Ms. Smith is teaching a unit on the American Civil War. She has many students in her fifth-grade classroom who are recent immigrants to the United States. Few of her students are familiar with United States culture, even fewer with U.S. history. Most of her students, however, have been in the school long enough to learn both oral and written English, enough at least to decode the words in their reading and content area texts. Ms. Smith has found that many of her students are excellent at memorizing information, but have difficulty using this information to interpret events. Which of the following teaching strategies would be likely to help Ms. Smith’s students build background knowledge they can use to learn from a text about the conflicts of the Civil War? (Mark YES, NO, or I’M NOT SURE for each choice.)

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>I’m not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Elicit their current level of knowledge about the Civil War in the United States.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>b) Read the class an “accessible” storybook about the U.S. Civil War, such as Ted Lewin’s <em>Red Legs: A Drummer Boy of the Civil War</em>.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>c) Divide the class into small groups to play Stratego, a board game involving battle strategies.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>d) Divide the class into “reading regiments” and ask students to design battle flags for their group.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>e) Require each student to memorize the dates of the major battles fought in the Civil War.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>f) Elicit students’ current knowledge about wars, why people fight them, and what their usual consequences are.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
When Mrs. Schwartz’s children proofread their own writing, they sometimes do not notice when they leave off the ending “-ed.” They are especially inattentive when the “-ed” sounds like /t/. Mrs. Schwartz wants to give them targeted, short practice listening for “-ed” when it is pronounced as /t/. Which of the following sets of words should she select? (Mark ONE answer.)

a) wanted, sorted, banded

b) picked, sipped, pitched

c) fringed, dodged, hummed

d) attached, angled, invented
4. Word Analysis / Knowledge of Students and Content

Ms. Reynolds dictated the following story to her class. She plans to use this to assess her students’ ability to hear and represent phonemes (i.e., the smallest segments of speech used to build a word).

I got a little white train.
   It was my best gift.
   I liked running it fast.
   One day I will go on a real train.

She looked at her students' papers. Ron’s paper looked like this:

I got a lidl wit chrane.
   It was my bst gif.
   I likt rung it fast.
   Won day I wl go on a rele chrane.

Which of the following words from Ron’s writing provide evidence that Ron can identify each phoneme heard when saying the word? (Mark YES, NO, or I'M NOT SURE for each word.)

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>I'm not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) “wit” for white</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>b) “chrane” for train</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>c) “bst” for best</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>d) “gif” for gift</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>e) “likt” for liked</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>f) “runin” for running</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>g) “wl” for will</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>h) “rele” for real</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>