



***APPENDIX B:***

***Measuring Teachers' Pedagogical Content Knowledge in Surveys:  
Detailed Results for the Domain of Reading/Language Arts\****

**Brian Rowan  
Steven G. Schilling  
Deborah L. Ball  
Robert Miller**

**With**

**Sally Atkins-Burnett  
Eric Camburn  
Delena Harrison  
Geoff Phelps**

**October, 2001**

\* Work on this paper was supported by grants from the Educational Statistics Services Institute of the American Institutes for Research, the Atlantic Philanthropies –North America, the Office of Educational Research and Improvement of the U.S. Department of Education, and the National Science Foundation (Interagency Educational Research Initiative). The opinions expressed here are those of the authors and are not specifically endorsed by the sponsors.

**Appendix B:**  
**Measuring Teachers' Pedagogical Content Knowledge in Surveys:**  
**Detailed Results for the Curricular Domain of Reading/Language Arts**

This appendix presents a detailed analysis of the measures of pedagogical content knowledge developed in the area of reading/language arts.

**Item Pool**

As Table 1 shows, we began work in the curricular domain of reading/language arts with 22 different scenarios. Six of these scenarios were designed to measure the facet of pedagogical content knowledge that we called “content knowledge,” and within these 6 scenarios, we had available a total of 26 separate items. For the facet labeled “knowledge of students’ thinking,” we developed 16 different scenarios, with a total of 38 items. As the table shows, the distribution of scenarios and items makes it impossible for us to measure both facets of pedagogical content knowledge across all “fine-grained” curricular domains selected for study.

**Table 2: Number of Items Assessing Teachers' Pedagogical Content Knowledge in Reading/Language Arts**

	Facet of Pedagogical Content Knowledge	
	<u>Content Knowledge</u>	<u>Knowledge of Students' Thinking</u>
<b><u>Word Analysis</u></b>		
• Letter-sound relationships	13 items (3 scenarios)	
• Phonemes	8 items (1 scenario)	
• Word recognition/sight words		12 items (4 scenarios)
• Phonetic cues		6 items (4 scenarios)
• Context/Picture/Syntactical Cues		16 items (4 scenarios)
<b><u>Reading Comprehension</u></b>		
• Monitoring for meaning		4 items (4 scenarios)
<b><u>Writing</u></b>		
• Editing process	5 items (2 scenarios)	

**Results**

In the following pages, we begin our discussion with scales constructed at the most fine-grained level of analysis, where we are measuring a single facet of teachers’ pedagogical content knowledge in a single, “fine-grained” area of the mathematics curriculum. We then present scales at larger grain sizes, which in the larger domain of reading/language arts, involved developing scales to measure a *single* facet of teachers’ pedagogical content knowledge across all of the fine-grained curricular areas where items measuring that facet were developed.

Given the format of the questionnaire items used in the study, it is not possible to show the scenarios and items used to form particular scales within the body of the tables. Instead, in each table, we simply list a number for each item (e.g., B22a) and provide a very

brief description of item content (in both the text and the table). Please note that the item numbers listed in the tables refer to the questionnaire form (A or B), as well as the scenario/items from which the scale was built. Thus, the referent “B22a” refers to scenario 22, item a, from form B. Readers interested in examining the exact wording and format of the items included in scales are referred to the questionnaires attached to this report.

The relevant results for reading/language arts are presented in Tables 2 – 11. Each of the tables shows all of the items considered for inclusion into a scale, and each table sorts these items according to whether they were kept in or deleted from the scale. Items kept in a scale are listed in order of difficulty (as estimated by the Rasch model), with the hardest items at the top and easier items listed in descending order of difficulty. The reader will note that these estimated difficulties do not necessarily correspond to actual p-values, which are the percentage of respondents in the sample answering that item correctly. Also note that the tables include information on items that we deleted from the final scales. These items are listed in reverse order of deletion, with those deleted at later points in the analysis listed before those deleted at earlier points. Each table also shows the item-to-scale biserial correlations for all items. For deleted items, the biserial correlation listed is the one estimated at the stage just prior to deletion. For retained items, the biserial correlation is the one estimated for the final scale.

### **Content Knowledge Scales**

#### **Letter/Sound Relationships**

We begin with a discussion of the fine-grained curricular areas where we attempted to build measures of teachers’ content knowledge. The first of these domains is letter/sound relationships. As Table 2 (next page) shows, most of the items in the Letter/Sound Relationships scale come from a single scenario (A2), where we presented teachers with a list of 11 letter or letter sound combinations and asked them whether they could think of a word making the long sound of “A” containing those letter sound combinations. Two additional items that appear in Table 2 come from different scenarios, but include only a single item from these scenarios. Scenario B3 asked teachers whether they could identify the “A” sound in alkaloid, scenario A3 asked teachers to identify the “X” sound in Xanax. The final scale consisted of 8 items with biserial correlation’s ranging from 0.162 to 0.816. Three of the deleted items had high p-values (A2E, A2F, A3), which indicates they were “too easy” to provide separation in the fit statistics. The other two deleted items (A2A and A2G) are items where it is difficult to generate examples of words where “ae” and “ea” have the long A sound. The reliability of the final scale was 0.697.

**Table 2: Letter Sound Relationships – Teachers’ Content Knowledge**

<i>Reliability = 0.697</i>			
<u>Items</u>	<i>% Correct</i>	<i>Item-to-Scale Biserial</i>	Item Content
<i>Kept</i>			
B3	0.964	0.162	• “a” sound in alkaloid
A2B	0.889	0.362	• long “A” sound: ai
A2C	0.542	0.559	• long “A” sound: ao
A2D	0.767	0.816	• long “A” sound: au
A2H	0.571	0.488	• long “A” sound: ei
A2I	0.852	0.787	• long “A” sound: ia
A2J	0.919	0.594	• long “A” sound: oa
A2K	0.918	0.776	• long “A” sound: ua
<i>Deleted</i>			
A2A	0.548	-0.204	• long “A” sound: ae
A2E	0.985	0.061	• long “A” sound: ay
A2F	0.859	0.030	• long “A” sound: a
A2G	0.317	0.041	• long “A” sound: ea
A3	0.910	0.107	• beginning sound in Xanax

**Phonemes**

We also attempted to measure teachers’ content knowledge in the area of phonemes. As Table 3 shows, all of the items in the Phonemes scale come from a single scenario (B4), which instructed respondents to provide the number of phonemes for each word presented in the scenario, where the words were chosen from a primary grade vocabulary list. The final scale consists of seven (of the eight) items from this scenario and had high internal consistency, with biserial correlation’s for kept items ranging from 0.911 to 0.999. The

**Table 3: Phonemes – Teachers’ Content Knowledge**

<i>Reliability = 0.999</i>			
<u>Items</u>	<i>% Correct</i>	<i>Item-to-Scale Biserial</i>	Item Content
<i>Kept</i>			
B4A	0.725	0.999	• phoneme: after
B4B	0.642	0.944	• phoneme: battle
B4C	0.755	0.999	• phoneme: chime
B4D	0.808	0.999	• phoneme: die
B4F	0.755	0.999	• phoneme: ship
B4G	0.755	0.999	• phoneme: shoe
B4F	0.712	0.911	• phoneme: terse
<i>Deleted</i>			
B4E	0.423	0.533	• phoneme: exit

only item deleted from the B4 scenario in this scale was item was B4E, which asks about the word “exit.” The word “exit” consists of four letters, but the letter “x” is composed of two letter sounds making the total number of phonemes in the word five. If the correct answer for B4E was changed to 4 phonemes, then B4E would have a positive biserial correlation of 0.533 and the reliability for the final scale would be 0.999.

### **Word Attack**

Table 4 presents a scale that combines items from the Letter/Sound Relationships scale and the Phonemes scale into a single scale that can be viewed as measuring teachers’ broader knowledge in the domain of word attack skills. As Table 2 shows, to form such a measure, we simply joined items from the previous two scales into the scale that is presented in Table 4. The 21 items put into this scaling analysis yielded a final 12-item scale with a reliability of 0.911. The biserial correlations are quite high ranging from 0.596 to 0.999. The measure order breaks, however, along subscale lines with the phonemes items at

<b>Table 4: Word Attack – Teachers’ Content Knowledge</b>			
<b>Reliability = 0.911</b>			
<b>Items</b>	<b>% Correct</b>	<b>Item-to-Scale Biserial</b>	<b>Item Content</b>
<b><i>Kept</i></b>			
<b>B4A</b>	<b>0.725</b>	<b>0.999</b>	• phoneme: after
<b>B4B</b>	<b>0.642</b>	<b>0.938</b>	• phoneme: battle
<b>B4C</b>	<b>0.755</b>	<b>0.999</b>	• phoneme: chime
<b>B4D</b>	<b>0.808</b>	<b>0.999</b>	• phoneme: die
<b>B4F</b>	<b>0.755</b>	<b>0.999</b>	• phoneme: ship
<b>B4G</b>	<b>0.755</b>	<b>0.999</b>	• phoneme: shoe
<b>B4H</b>	<b>0.712</b>	<b>0.905</b>	• phoneme: terse
<b>A2C</b>	<b>0.542</b>	<b>0.596</b>	• long “A” sound: ao
<b>A2D</b>	<b>0.767</b>	<b>0.609</b>	• long “A” sound: au
<b>A2I</b>	<b>0.852</b>	<b>0.910</b>	• long “A” sound: ia
<b>A2J</b>	<b>0.919</b>	<b>0.715</b>	• long “A” sound: oa
<b>A2K</b>	<b>0.918</b>	<b>0.811</b>	• long “A” sound: ua
<b><i>Deleted</i></b>			
<b>B3</b>	<b>0.964</b>	<b>0.427</b>	• “a” sound in alkaloid
<b>B4E</b>	<b>0.423</b>	<b>0.522</b>	• phoneme: exit
<b>A2A</b>	<b>0.548</b>	<b>-0.278</b>	• long “A” sound: ae
<b>A2B</b>	<b>0.889</b>	<b>0.212</b>	• long “A” sound: ai
<b>A2E</b>	<b>0.985</b>	<b>-0.067</b>	• long “A” sound: ay
<b>A2F</b>	<b>0.859</b>	<b>0.129</b>	• long “A” sound: a
<b>A2G</b>	<b>0.317</b>	<b>-0.067</b>	• long “A” sound: ea
<b>A2H</b>	<b>0.571</b>	<b>0.346</b>	• long “A” sound: ei
<b>A3</b>	<b>0.910</b>	<b>0.177</b>	• beginning sound in Xanax

the top of the range and letter/sound relationship items grouped at the end of the order. The seven items from the final Phonemes scale (see Table 3) are kept in this scale, while five of the eight items from the final Letter/Sound Relationships scale (see Table 2) are kept in this combined measure (items A2B, A2H, and B3 are deleted).

***Editing***

To this point, we have been measuring teachers’ content knowledge in the curricular domains of word analysis. However, we also developed one measure of teachers’ content knowledge in the area of writing. This measure focused on teachers’ knowledge of editing. Here, we presented respondents with 5 items in 2 scenarios that asked them to choose the best sentence from a selection four to five sentences, of which only one or two were grammatically acceptable. The results shown in Table 5 (below) show that the scale had a very low reliability, largely because the items were very easy for teachers to answer correctly. The percent correct for each item ranges from 0.712 to 0.985, for example. Given these high p-values, it appears that the items we developed were not particularly challenging for teachers, resulting in a scale with very little item separation. As Table 5 shows, the Editing scale ended up having a reliability of only 0.106.

<b><i>Table 5: Editing – Teachers’ Content Knowledge</i></b>			
<b><i>Reliability = 0.106</i></b>			
<b><i>Items</i></b>	<b><i>% Correct</i></b>	<b><i>Item-to-Scale Biserial</i></b>	<b><i>Item Content</i></b>
<b><i>Kept</i></b>			
<b>B5A</b>	<b>0.966</b>	<b>-0.081</b>	• <b>sentence editing</b>
<b>B5B</b>	<b>0.724</b>	<b>0.227</b>	• <b>sentence editing</b>
<b>B5C</b>	<b>0.914</b>	<b>0.009</b>	• <b>sentence editing</b>
<b>A5A</b>	<b>0.985</b>	<b>-0.260</b>	• <b>sentence editing</b>
<b>A5B</b>	<b>0.712</b>	<b>-0.030</b>	• <b>sentence editing</b>

***Reading/Language Arts – Overall Measure of Content Knowledge***

The final content knowledge measure in reading/language arts combines items from all of the previous scales, that is, the Phonemes, Letter/Sound Relationships, and Sentence Editing scales, augmented by other items for which there were not sufficient items within the curricular topic to create subscales. These additional items include teaching scenarios in which teachers divided the words “aphasia” and “syllabication” into syllables (A4, B2) and items in which teachers’ were asked to identify the steps for both the SQ3R and KWL reading strategies. The results are reported in Table 6 (next page). We started the analysis with 30 items that produced a final scale with 21 items and a reliability of 0.870. The biserial correlation’s ranged from 0.204 to 0.999. Once again, the measure order shows that the items from the Phonemes and Letter/Sound Relationships subscales (see Tables 18 & 19)

are kept and essentially remain nested within their respective subscale groupings. However, two sentence editing items (A5A, B5C), the two reading strategy items (SQ3R, K), and one syllable division item (A4) also appear in the final scale.

<b>Table 6: Teachers' Content Knowledge – All Items</b>			
<b>Reliability = 0.870</b>			
<b>Items</b>	<b>% Correct</b>	<b>Item-to-Scale Biserial</b>	<b>Item Content</b>
<b>Kept</b>			
B3	0.966	0.424	• “a” sound in alkaloid
B4A	0.725	0.999	• phoneme: after
B4B	0.642	0.989	• phoneme: battle
B4C	0.755	0.935	• phoneme: chime
B4D	0.808	0.999	• phoneme: die
B4E	0.423	0.531	• phoneme: exit
B4F	0.755	0.999	• phoneme: ship
B4G	0.755	0.999	• phoneme: shoe
B4H	0.712	0.815	• phoneme: terse
B5C	0.914	0.352	• sentence editing
A2B	0.889	0.204	• long “A” sound: ai
A2C	0.542	0.639	• long “A” sound: ao
A2D	0.767	0.768	• long “A” sound: au
A2H	0.571	0.487	• long “A” sound: ei
A2I	0.852	0.752	• long “A” sound: ia
A2J	0.919	0.596	• long “A” sound: oa
A2K	0.918	0.615	• long “A” sound: ua
A4	0.627	0.240	• divide “aphasia” into syllables
A5A	0.985	0.580	• sentence editing
SQ3R	0.629	0.343	• identify SQ3R steps
KWL	0.780	0.288	• identify KWL steps
<b>Deleted</b>			
B2	0.224	0.062	• divide “syllabication” into syllables
B5A	0.966	0.021	• sentence editing
B5B	0.724	0.134	• sentence editing
A2A	0.548	-0.357	• long “A” sound: ae
A2E	0.985	0.056	• long “A” sound: ay
A2F	0.859	0.184	• long “A” sound: a
A2G	0.317	0.043	• long “A” sound: ea
A3	0.910	0.179	• beginning sound in Xanax
A5B	0.712	0.136	• sentence editing

## **Knowledge of Students' Thinking Scales**

In this section, we turn to scales designed to measure teachers' *knowledge of students' thinking* in various fine-grained areas of the reading/language arts curriculum. As Table 1 shows, scales were constructed in the areas of word recognition/sight words, use of phonetic cues, use of context, picture, and syntactical cues, and monitors for meaning. In addition, we created an overall measure of teachers' knowledge of students' thinking from items in all of these domains.

### **Word Recognition/Sight Words**

Table 7 shows the scale for measuring teachers' knowledge of students' thinking in the domain of word recognition/sight words. This scale was constructed from 12 different items appearing in 4 different scenarios. The scenarios presented teachers with cases of students reading passages from a grade-appropriate text and then asked teachers to determine if the depicted student: a) was exchanging visually similar sight words, b) was reading only words in his/her sight vocabulary, or c) was showing mastery of common sight words. Of the twelve items available for analysis, half were kept in the final scale (Table 7). The biserial correlation's range from 0.148 to 0.402, and the reliability for the scale is 0.486.

<b>Table 7: Word Recognition/Sight Words – Knowledge of Students' Thinking</b>			
<b>Reliability = 0.486</b>			
<b>Items</b>	<b>% Correct</b>	<b>Item-to-Scale Biserial</b>	<b>Item Content</b>
<b><i>Kept</i></b>			
<b>A7F</b>	<b>0.231</b>	<b>0.205</b>	• diagnostic: exchanges visually- similar sight words
<b>A8F</b>	<b>0.200</b>	<b>0.402</b>	• diagnostic: exchanges visually- similar sight words
<b>A8I</b>	<b>0.515</b>	<b>0.148</b>	• diagnostic: only reads words in sight vocabulary
<b>B7I</b>	<b>0.439</b>	<b>0.323</b>	• diagnostic: only reads words in sight vocabulary
<b>B8B</b>	<b>0.754</b>	<b>0.317</b>	• diagnostic: mastery of common sight words
<b>B8I</b>	<b>0.456</b>	<b>0.327</b>	• diagnostic: only reads words in sight vocabulary
<b><i>Deleted</i></b>			
<b>A7B</b>	<b>0.545</b>	<b>0.055</b>	• diagnostic: mastery of common sight words
<b>A7I</b>	<b>0.585</b>	<b>-0.277</b>	• diagnostic: only reads words in sight vocabulary
<b>A8B</b>	<b>0.561</b>	<b>-0.489</b>	• diagnostic: mastery of common sight words
<b>B7B</b>	<b>0.464</b>	<b>0.105</b>	• diagnostic: mastery of common sight words
<b>B7F</b>	<b>0.246</b>	<b>-0.150</b>	• diagnostic: exchanges visually- similar sight words
<b>B8F</b>	<b>0.596</b>	<b>-0.284</b>	• diagnostic: exchanges visually- similar sight words

### *Use of Phonetic Cues*

A scale measuring teachers' knowledge of student thinking in the domain of students' use of phonetic cues was constructed from an additional 6 items from the same four scenarios just mentioned. In response to the passages, teachers were asked if students were: (a) proficient with consonant blends and/or b) relied too heavily on phonetic details. Table 8 shows that only three items were kept in this scale, and that the overall reliability for the scale was 0.374. This low reliability results in part from too few items in the scale, but it is also the case that the biserial correlation's for all six items in the analysis were low as well (ranging from -0.047 to 0.433). Looking at the p-values for these items, it appears that were better able to determine whether a student relies too heavily on phonetic details in comparison to determining whether a student was proficient with consonant blends.

<i>Table 8: Use of Phonetic Cues – Knowledge of Students' Thinking</i>			
<i>Reliability = 0.374</i>			
<u>Items</u>	<u>% Correct</u>	<u>Item-to-Scale Biserial</u>	<u>Item Content</u>
<i>Kept</i>			
<b>B7A</b>	<b>0.246</b>	<b>0.373</b>	• <b>diagnostic: proficient with consonant blends</b>
<b>B7G</b>	<b>0.105</b>	<b>0.433</b>	• <b>diagnostic: relies heavily on phonetic details</b>
<b>B8G</b>	<b>0.789</b>	<b>0.121</b>	• <b>diagnostic: relies heavily on phonetic details</b>
<i>Deleted</i>			
<b>A7A</b>	<b>0.104</b>	<b>0.412</b>	• <b>diagnostic: proficient with consonant blends</b>
<b>A7G</b>	<b>0.789</b>	<b>-0.047</b>	• <b>diagnostic: relies heavily on phonetic details</b>
<b>A8G</b>	<b>0.697</b>	<b>-0.156</b>	• <b>diagnostic: relies heavily on phonetic details</b>

### *Use of Context, Picture, and Syntactical Cues*

We now move our analysis to teachers' knowledge of reading comprehension skills. In this area, the first fine-grained domain considered is teachers' knowledge of students' use of context, picture, and syntactical cues. Here, teachers were asked to determine whether the script of a student's attempt at reading various passages suggested that the student: a) uses knowledge of syntax, b) monitors for meaning, c) recognizes important content words, and/or d) relies too heavily on pictures and other context clues. The final scale is presented in Table 9 (next page). Of the original 16 items, 11 are retained, resulting in a scale with a reliability of 0.724. The biserial correlation's ranged from 0.358 to 0.615. All four items about the student "monitoring for meaning" and the four items regarding whether the student "uses knowledge of syntax" were kept in the final scale. Only two of the four items in the area of "recognizes important content words" are kept and only one "relies too heavily on pictures and other context clues" item remains.

**Table 9: Use of Context, Picture, and Syntactical Cues – Knowledge of Students’ Thinking**

<i>Reliability = 0.724</i>			
<u>Items</u>	<i>% Correct</i>	<i>Item-to-Scale Biserial</i>	<b>Item Content</b>
<i>Kept</i>			
A7C	0.754	0.453	• diagnostic: knowledge of syntax
A7D	0.672	0.385	• diagnostic: monitors for meaning
A8C	0.547	0.407	• diagnostic: knowledge of syntax
A8D	0.788	0.615	• diagnostic: monitors for meaning
A8E	0.892	0.588	• diagnostic: recognizes important content words
A8H	0.667	0.473	• diagnostic: relies heavily on pictures and other context clues
B7C	0.807	0.505	• diagnostic: knowledge of syntax
B7D	0.702	0.415	• diagnostic: monitors for meaning
B8C	0.589	0.358	• diagnostic: knowledge of syntax
B8D	0.789	0.555	• diagnostic: monitors for meaning
B8E	0.839	0.637	• diagnostic: recognizes important content words
<i>Deleted</i>			
A7E	0.697	-0.108	• diagnostic: recognizes important content words
A7H	0.646	0.176	• diagnostic: relies heavily on pictures and other context clues
B7E	0.891	-0.053	• diagnostic: recognizes important content words
B7H	0.684	0.154	• diagnostic: relies heavily on pictures and other context clues
B8H	0.75	0.043	• diagnostic: relies heavily on pictures and other context clues

**Monitors for Meaning**

Table 10 presents results for a scale designed to measure teachers’ knowledge of students’ ability to monitor for meaning during reading. Here, four items were used. Table 10 shows that the scale reliability is 0.433. The biserial correlations range from 0.193 to 0.433.

**Table 10: Monitors for Meaning – Knowledge of Students’ Thinking**

<b>Reliability = 0.433</b>			
<b>Items</b>	<b>% Correct</b>	<b>Item-to-Scale Biserial</b>	<b>Item Content</b>
<b><i>Kept</i></b>			
<b>A7D</b>	<b>0.672</b>	<b>0.333</b>	• diagnostic: monitors for meaning
<b>A8D</b>	<b>0.781</b>	<b>0.433</b>	• diagnostic: monitors for meaning
<b>B7D</b>	<b>0.702</b>	<b>0.193</b>	• diagnostic: monitors for meaning
<b>B8D</b>	<b>0.789</b>	<b>0.158</b>	• diagnostic: monitors for meaning

**Overall Measure of Teachers’ Knowledge of Students’ Thinking**

Our final scale combines items from the A7, A8, B7, and B8 series discussed above, along with additional items that could not be used in separate scales due to an insufficient number of items representing a specific curricular topic. These additional items include three items (AB6A, AB6B, AB6C) asking teachers to identify which sets of skills are more difficult for elementary students: a) identifying the vowel sound in “came” or ending consonant in “came”, b) identifying story details or identifying the main idea, and c) identifying each phoneme in “bat” or identifying the syllables in “umbrella.” It also includes items from the A14 and B15 series posing scenarios in which teachers are asked to recommend possible remediation strategies for students struggling with specific comprehension difficulties. Yet another teaching situation appears in the A11 series, where teachers are asked to assess the reading comprehension abilities of a particular student.

Table 11 (next page) shows that of the 47 items input for this analysis, 28 items were kept in the final scale (with 19 items deleted). Mostly due to the sheer number of items the reliability for this scale is 0.798. The biserial correlations range from 0.184 to 0.915. The four “monitoring for meaning” items and the four “uses knowledge of syntax” items from the A7, A8, B7, and B8 series remain in the scale. However, all of the other items from these series were deleted. Also, Table 11 shows that the final scale contains items AB6A, AB6C, A11B, A11D, and A11E.

**Table 11: All Items - Knowledge of Students' Thinking**

<b>Reliability = 0.798</b>			
<b>Items</b>	<b>% Correct</b>	<b>Item-to-Scale Biserial</b>	<b>Item Content</b>
<b>Kept</b>			
<b>AB6A</b>	<b>0.677</b>	<b>0.188</b>	• task difficulty: vowel sound/ending consonant
<b>AB6C</b>	<b>0.781</b>	<b>0.406</b>	• task difficulty: identify phonemes/ identify syllables
<b>A14A</b>	<b>0.968</b>	<b>0.588</b>	• diagnostic: comprehension skills
<b>A7C</b>	<b>0.754</b>	<b>0.184</b>	• diagnostic: knowledge of syntax
<b>A7D</b>	<b>0.672</b>	<b>0.262</b>	• diagnostic: monitors for meaning
<b>A7G</b>	<b>0.844</b>	<b>0.224</b>	• diagnostic: relies heavily on phonetic details
<b>A7H</b>	<b>0.646</b>	<b>0.342</b>	• diagnostic: relies heavily on pictures and other context clues
<b>A8C</b>	<b>0.547</b>	<b>0.508</b>	• diagnostic: knowledge of syntax
<b>A8D</b>	<b>0.788</b>	<b>0.636</b>	• diagnostic: monitors for meaning
<b>A8E</b>	<b>0.892</b>	<b>0.915</b>	• diagnostic: recognizes important content words
<b>A8F</b>	<b>0.200</b>	<b>0.338</b>	• diagnostic: exchanges visually-similar sight words
<b>A8G</b>	<b>0.833</b>	<b>0.339</b>	• diagnostic: relies heavily on phonetic details
<b>A8H</b>	<b>0.667</b>	<b>0.254</b>	• diagnostic: relies heavily on pictures and other context clues
<b>A8I</b>	<b>0.515</b>	<b>0.351</b>	• diagnostic: only reads words in sight vocabulary
<b>B7A</b>	<b>0.246</b>	<b>0.453</b>	• diagnostic: proficient with consonant blends
<b>B7C</b>	<b>0.807</b>	<b>0.366</b>	• diagnostic: uses knowledge of syntax
<b>B7G</b>	<b>0.105</b>	<b>0.317</b>	• diagnostic: relies heavily on phonetic details
<b>B7H</b>	<b>0.684</b>	<b>0.323</b>	• diagnostic: relies heavily on pictures and other context cues
<b>B7I</b>	<b>0.439</b>	<b>0.383</b>	• diagnostic: only reads words in sight vocabulary
<b>B8B</b>	<b>0.754</b>	<b>0.424</b>	• diagnostic: mastery of common sight words
<b>B8C</b>	<b>0.589</b>	<b>0.540</b>	• diagnostic: uses knowledge of syntax
<b>B8D</b>	<b>0.789</b>	<b>0.779</b>	• diagnostic: monitors for meaning
<b>B8E</b>	<b>0.839</b>	<b>0.842</b>	• diagnostic: recognizes important content words
<b>B8G</b>	<b>0.789</b>	<b>0.496</b>	• diagnostic: relies heavily on phonetic details
<b>B8I</b>	<b>0.456</b>	<b>0.605</b>	• diagnostic: only reads words in sight vocabulary
<b>A11B</b>	<b>0.844</b>	<b>0.560</b>	• student has trouble moving beyond her background experience
<b>A11D</b>	<b>0.583</b>	<b>0.301</b>	• student ignores critical facts
<b>A11E</b>	<b>0.674</b>	<b>0.264</b>	• student misses overall point of the text

Table continues on next page with deleted items

**Table 11 (cont.): All Items - Knowledge of Students' Thinking**

<b>Deleted</b>			
<b>AB6B</b>	<b>0.677</b>	<b>-0.085</b>	• task difficulty: story details/ identify main idea
<b>A14B</b>	<b>0.956</b>	<b>-0.218</b>	• diagnostic: comprehension skills
<b>A15A</b>	<b>0.948</b>	<b>0.052</b>	• diagnostic: comprehension skills
<b>A15B</b>	<b>0.776</b>	<b>0.141</b>	• diagnostic: comprehension skills
<b>A7A</b>	<b>0.697</b>	<b>0.148</b>	• diagnostic: proficient with consonant blends
<b>A7B</b>	<b>0.545</b>	<b>-0.001</b>	• diagnostic: mastery of common sight words
<b>A7E</b>	<b>0.697</b>	<b>-0.117</b>	• diagnostic: recognizes important content words
<b>A7F</b>	<b>0.231</b>	<b>-0.314</b>	• diagnostic: exchanges visually-similar sight words
<b>A7I</b>	<b>0.585</b>	<b>0.136</b>	• diagnostic: only reads words in sight vocabulary
<b>A8B</b>	<b>0.561</b>	<b>-0.044</b>	• diagnostic: mastery of common sight words
<b>B7B</b>	<b>0.464</b>	<b>0.012</b>	• diagnostic: mastery of common sight words
<b>B7D</b>	<b>0.702</b>	<b>0.170</b>	• diagnostic: monitors for meaning
<b>B7E</b>	<b>0.891</b>	<b>-0.019</b>	• diagnostic: recognizes important content words
<b>B7F</b>	<b>0.246</b>	<b>-0.143</b>	• diagnostic: exchanges visually- similar sight words
<b>B8F</b>	<b>0.596</b>	<b>-0.306</b>	• diagnostic: exchanges visually- similar sight words
<b>B8H</b>	<b>0.750</b>	<b>-0.011</b>	• diagnostic: relies heavily on pictures and other context cues
<b>A11A</b>	<b>0.702</b>	<b>0.049</b>	• student gets lost in details
<b>A11C</b>	<b>0.646</b>	<b>-0.012</b>	• student did not understand question
<b>A11F</b>	<b>0.750</b>	<b>0.138</b>	• student did not read the text