

## **Abstract Title Page**

**Title:**

Relating knowledge about reading to teaching practice: An exploratory validity study of a teacher knowledge assessment.

**Author(s):**

Geoffrey Phelps, PhD, University of Michigan; David Johnson, PhD, University of Michigan; Joanne Carlisle, PhD, University of Michigan.

## Abstract Body

### Background/context:

Teachers are central to the improvement of reading instruction. In the landmark report published by the National Commission for Teaching and America's Future (1996), the authors made the point succinctly, stating that "what teachers know and can do is the most important influence on what students learn" (pg. 6). Current interest in teacher quality has led to a wide range of policies that emphasize the importance of teacher education and professional development. For example, federal policies such as the No Child Left Behind Act of 2001, devote substantial resources to teachers' professional development, viewing teacher improvement as a primary lever for improving the quality of reading instruction and ultimately student achievement.

In recent years, there have been a number of reports seeking to clarify what elementary reading teachers need to know and learn in order to teach effectively (Anders, Hoffman, & Duffy, 2000; Hoffman & Pearson, 2000; International Reading Association, 2000; Moats, 1999; Snow, Griffin, & Burns, 2005). According to these reports, teachers need a foundational understanding of the structure of the English language, knowledge about how students develop as readers, and knowledge about effective methods for teaching reading. A common theme across these reports is the recognition that teaching reading is complex work that demands deep professional knowledge. Further, this is knowledge not likely to be held by virtue of being a literate, well-educated adult. According to Snow, Griffin, and Burns (2005) the knowledge about reading that teachers need is not just content knowledge, divorced from the job of teaching reading, but rather knowledge about reading that is embedded in the practice of teaching—what they refer to as "usable knowledge" (p. 11). Without formal opportunities to develop this knowledge through teacher education or professional development, reading teachers are likely to be poorly prepared to teach young children to read.

Interest in teacher knowledge in the area of reading has been accompanied by interest in development and use of research measures—those that are suitable for studying teachers' knowledge, its role in teaching practice, and its development. One indicator of this interest is the frequent requirement from research funders, such as the Institute of Education Sciences, that studies focused on teacher development and learning make use of validated measures of teacher knowledge to assess program effects. While these recommendations are sensible, the reality is that the development of high quality measures of teacher knowledge has lagged behind interest. There is scant evidence that extant measures of teacher knowledge are valid indicators of teacher quality (Phelps & Carlisle, 2008). In particular, there is limited evidence that differences in teacher knowledge are associated with meaningful differences in outcomes such as student learning (see for example, Cirino, Pollard-Durodola, Foorman, Carlson & Francis, 2007; Carlisle, Correnti, Phelps & Zeng, in press). The general lack of information on the validity of teacher knowledge measures leaves researchers on unsure footing. Even the best-designed studies are not likely to succeed if scores on measures are poorly related to constructs of interest.

**Purpose/objective/research question/focus of study:**

The research reported in this paper is focused directly on assessing the validity of the *Teaching Knowledge about Reading and Reading Practices* (TKRRP) assessment. Following the recommendations of the Standards for Educational and Psychological Testing (APA/AERA, 1999), we see validation as a process of constructing an argument that builds evidence for valid inferences that can be drawn regarding the proposed uses of assessment scores. In this study, we investigate a central claim about the knowledge we are seeking to measure with the TKRRP to inform our understanding of the inferences that can be validity drawn from the test scores – that is, *teacher’s scores on the Teacher Knowledge of Reading and Reading Practices (TKRRP) are associated with the quality of the teacher’s reading instruction.*

In stating this claim, we assume that the TKRRP is a measure of teacher quality. That is, we assume that teachers who score higher on the TKRRP will be more effective teachers than those who have lower TKRRP scores. In this study we investigate this claim by collecting, scoring, and analyzing videotapes of teachers’ reading blocks to determine the relationship between the quality of their reading instruction and their knowledge.

**Setting:**

The study was carried out in the context of the Michigan Reading First evaluation. All first through third grade teachers (n=1,635) in schools participating in Reading First were administered a questionnaire containing the *Teaching Knowledge about Reading and Reading Practices* in the spring of 2008. This census of Reading First teachers provided a context for examining the psychometric properties of the measure of teachers’ knowledge and for examining the relationship of the teacher knowledge score to other traditional teacher quality indicators such as degree attainment, certification status, professional preparation, and teaching experience (Carlisle, Johnson, Phelps & Rowan, 2008). See Appendix B for assessment questions and basic information on the psychometric properties of the TKRRP.

**Population/Participants/Subjects:**

In the Spring of 2008, second and third grade teachers in Reading First schools in Michigan were invited to participate in a video tape validation study. These schools were selected because they were generally representative of RF schools in Michigan and because they were geographically close enough to the research center to allow for efficient collection of videotapes of instruction. Participating teachers (n=31) were videotaped for entire literacy blocks. Teachers were contacted within the week for 30-minute debriefing interviews conducted over the telephone.

**Intervention/Program/Practice:**

While this study was carried out in the context of Michigan Reading First, the particular intervention is not directly applicable to this proposal.

## **Research Design:**

The validation study employs two different approaches to investigate the validity of the TKRRP. (See Appendix for TKRRP questions and basic psychometric characteristics). First we make use of the data from a systematic coding of the quality of teachers' lessons using the Video Assessment of Teaching (VAT-R). We use a two level HLM with lessons nested within classrooms to predict variation in lesson quality. In a parallel analysis using the same sample of 31 teachers, we conducted a qualitative case study comparing high knowledge teachers and low knowledge teachers. In this analysis we use constant comparative techniques to investigate characteristics of the reading instruction of teachers with high levels of knowledge compared to teachers with low levels of knowledge.

## **Data Collection and Analysis:**

Participants in the study were 31 second and third grade Michigan Reading First teachers who volunteered to be videotaped and interviewed. The videotaping visit was scheduled in cooperation with the teachers. While teachers were instructed to not change their regular daily course of instruction, it is likely that the videotaped lessons for many teachers represent a best effort at providing high quality reading instruction within the context of Reading First and thus a good example of strong application of teacher knowledge to teaching practice. The Video Assessment of Teaching-Reading (VAT-R) was developed to provide a reliable coding system for characterizing the quality of reading lessons. The VAT-R coding is applied at the lesson level, with lessons defined as a shift in the subject matter purpose of the lesson. Most teachers taught between 4 and 7 lessons in their literacy block. Lessons were coded for the subject matter content and related teaching approaches used by the teacher during the lesson. In addition to these descriptive categories, a set of 13 evaluative criteria, with a single overarching evaluation category, were developed to characterize the quality of the reading instruction (Figure 1). This analysis makes use of the quality ratings. Three coders were trained to use the VAT-R system and coding commenced when inter rater reliability exceeded .80.

We will use a two level HLM to investigate the relation of teacher knowledge and lesson quality. At level 1 we will enter lesson level variables (e.g., lesson length) and at level 2 classroom level variables (e.g., teacher attainment, certification, experience, % classroom free and reduced lunch, % limited English language proficiency). When coding is complete we expect to have approximately 200 lessons nested within 31 teachers. The analysis will focus on the contribution of teacher knowledge to variation in lesson quality. As part of this analysis, we will also investigate the structure of the VAT-R measure of lesson quality, exploring different strategies for constructing composite measures and for exploring whether teacher quality is more strongly associated with particular indicators of lesson quality. The hierarchical structure of the data will allow us to describe the extent to which lesson quality varies within and among teachers.

As described above, the 31 teachers participating in the study are also participating in the Michigan Reading First evaluation. The larger evaluation study includes all 1,635 first through third grade teachers in Michigan teaching in Reading First schools. We have used the population distribution of Michigan Reading First 1-3 teacher knowledge scores to identify a high knowledge strata (top 12.5%) and a low knowledge strata (bottom 12.5%) for the population.

The sample of 31 teachers participating in this study has 7 teachers in the low knowledge strata and 5 in the high knowledge strata. To keep our analysis manageable we selected the 4 highest scoring teachers and four lowest scoring teachers from the 31 teachers participating in the video validation study. Analyses of the video were conducted in stages. First, video files for the high knowledge teachers were watched and provisional categories common across the teachers and related to the quality of instruction were identified. Second, the same procedure was carried out for the low knowledge teachers. Third, categories were compared across high and low knowledge cases with special attention given to categories represented in one instance but not the other. This stage was particularly important as it provides a way of assessing whether there are features of practice that are absent for high knowledge teachers, but observed for low knowledge teachers and vice versa. Fourth, the coding schemes for the high and low cases were refined and used to analyze all video files to identify any characteristics of teaching practice that appeared to be related to high scores on the knowledge test and vice versa. Finally, high and low knowledge teachers who did not fit the profile of their knowledge were examined as a way to generate hypotheses about the factors that might hinder or support the use of knowledge in practice.

### **Findings/Results:**

The coding of instruction using the VAT-R is ongoing with completion expected by the end of December 2008. Preliminary analysis including 18 of the 31 teachers and employing a simple OLS regression were promising. In this simple analysis, we coded all data at the lesson level and included measures for teachers' knowledge, years teaching experience, master's degree, % school free and reduced hot lunch, and % school limited English proficiency. Because master's degree and years experience were highly correlated and exhibited multicollinearity, these variables were entered separately into the regression model. The coefficients were similar across models and in the final model reported in Table 2 master's degree was excluded. In this model, teachers' knowledge is significantly associated with the quality of reading lessons ( $\beta=.264, p < .009$ ). This provides preliminary evidence that scores on the TKRRP are associated with instructional quality.

The qualitative case analysis is still underway. Preliminary results suggest a number of interesting findings. Unsurprisingly, many of the same features of quality that were prominent in the VAT-R coding were also noted in the qualitative analysis. However, there were also noticeable differences in instruction that were not represented in the VAT-R. For example, preliminary analyses suggest that the instruction of the higher knowledge teachers differed from the lower knowledge teachers. Higher knowledge teachers appeared to provide more elaborate and accurate descriptions of ideas and linked classroom work more directly to literacy concepts. High knowledge teachers asked more demanding questions and probed student responses more deeply. High knowledge teachers tended to use language more precisely to explain concepts and elaborate on classroom discussions. While these findings are preliminary, they do suggest that there were noticeable differences in the reading instruction observed for the high and low knowledge cases.

Taken together these two analyses provide useful insights into the two different approaches to characterizing instructional quality. The VAT-R was informed by a literature review of lesson features that are thought to be associated with lesson quality. Categories were further refined by piloting the instrument on videos of instruction to identify domains that had been overlooked and

ones that could be coded reliably. This process was largely theory driven and constrained by the need to obtain acceptable levels of inter-rater reliability. The second approach was largely inductive, working from cases defined by teachers' knowledge. This approach allowed for attention to features of teachers instruction that would demand a very high level of inference from coders and thus would be unsuitable for many measurement instruments. It also allowed for identifying features of instruction that might be related to teacher's knowledge scores but not prominent in the literature.

### **Conclusions:**

Program evaluators, teacher educators, and others studying teacher development and instructional improvement will find this session useful for a number of reasons. One purpose of the session is to introduce and discuss a measure of teacher knowledge that can be used in large-scale survey studies. Researchers interested in tools suitable for measuring teacher knowledge will have an opportunity to learn about one of the few instruments that have been developed for these purposes and validated with a large sample of teachers. However, the results from this session are substantively interesting in their own right. There is remarkably limited evidence that differences in teacher knowledge are associated with meaningful outcomes. These analyses, while they should rightly be seen as preliminary and exploratory given the limitations imposed by the size of the sample and the selection of participants, do indicate that variation in teacher knowledge as measured by the TKRRP is associated with higher quality instruction.

## Appendixes

### Appendix A. References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anders, P. L., Hoffman, J. V., & Duffy, G. G. (2000). Teaching teachers to teach reading: Paradigm shifts, persistent problems, and challenges. In M. I. Kamil, J. Mosenthal, D. Pearson & R. Barr (Eds.), *Handbook of Reading Research: Volume III* (pp. 719-742). Mahwah, NJ: Erlbaum.
- Carlisle, J., Correnti, R., Phelps, G., & Zeng, J. (in press). Investigating Teachers' Knowledge of Language Structure and its Relation to Students' Reading Achievement. *Reading and Writing: An Interdisciplinary Journal*, Accepted for publication.
- Carlisle, J., Johnson, D., Phelps, G., & Rowan, B. (2008). *Teacher knowledge of Reading and Reading Practices*. Ann Arbor: University of Michigan.
- Cirino, P. T., Pollard-Durodola, S. D., Foorman, B. R., Carlson, C. D., & Francis, D. J. (2007). Teacher characteristics, classroom instruction, and student literacy and language outcomes in bilingual kindergartners. *Elementary School Journal*, 107(4), 341-364.
- Hoffman, J. V., & Pearson, P. D. (2000). Reading teacher education in the next millennium: What your grandmother's teacher didn't know that your granddaughter's teacher should. *Reading Research Quarterly*, 35, 28-44.
- International Reading Association. (2000). Excellent reading teachers: A position statement of the International Reading Association. *The Reading Teacher*, 54(2), 235-240.
- Moats, L. (1999). *Teaching reading IS rocket science: What expert teachers of reading should know and be able to do*. (No. 39-0372). Washington, DC: American Federation of Teachers.
- National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future*. Washington, DC: <http://www.nctaf.org>.
- Phelps, G., & Carlisle, J. (2008). The measurement of teachers' knowledge about reading: A research synthesis. *Manuscript submitted for publication*.
- Snow, C. E., Griffin, P., & Burns, M. S. (Eds.). (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*. San Francisco, CA: Jossey-Bass.

**Appendix B. Tables and Figures**  
*Not included in page count.*

Table 1.

*Predictors of lesson quality (lesson n=117)*

Variables	<i>B</i>	<i>SE B</i>	$\beta$
Teacher Knowledge	.28**	.106	2.67**
Professional Development	-.18**	.066	-2.66**
Years Teaching	.09**	.028	3.59**
% Limited English Proficiency	.56	.601	.91
% Free and Reduced Lunch	-1.66**	.516	-.32**

$R^2 = .389$ .

\*\*  $p < .01$ .



Figure 1.

*Video Assessment of Teaching-Reading Quality Codes*

1. The text used was a good choice given the purpose of the lesson.
2. The text was a good choice, given the students' reading skill and understanding of relevant concepts, ideas, background knowledge.
3. The teacher made good use of the text during the lesson.
4. The teacher prepared the students for the literacy activity or reading event.
5. The teacher selected appropriate instructional methods and/or activities to foster students' learning (given the purpose of the lesson).
6. The teacher kept children's attention focused on the lesson and managed their behavior effectively.
7. The teacher demonstrated sensitivity to cultures and languages of the students.
8. The teacher's lesson was cognitively challenging for the students.
9. The teacher used effective methods to foster students' interest in the lesson and a positive attitude toward the content and activities of the lesson
10. The teacher was sensitive to the needs of individual students.
11. The lesson was coherent. The parts of the lesson worked together effectively.
12. The pace of the teachers' explanations and activities was appropriate.
13. The teacher's explanations of literacy concepts were accurate and/or complete.
14. Overall, the teacher's design, instructional method, use of materials, and management of the lesson suggest that he/she holds relevant knowledge about reading and reading instruction and used this knowledge effectively in teaching this lesson (here a 6 would be expert/excellent and a 1 would be ineffective instruction and support for learning to read).

*Note: All codes are rated on a six point scale using the anchors "strongly agree" and "strongly disagree. There is a "not observed" option for every code.*

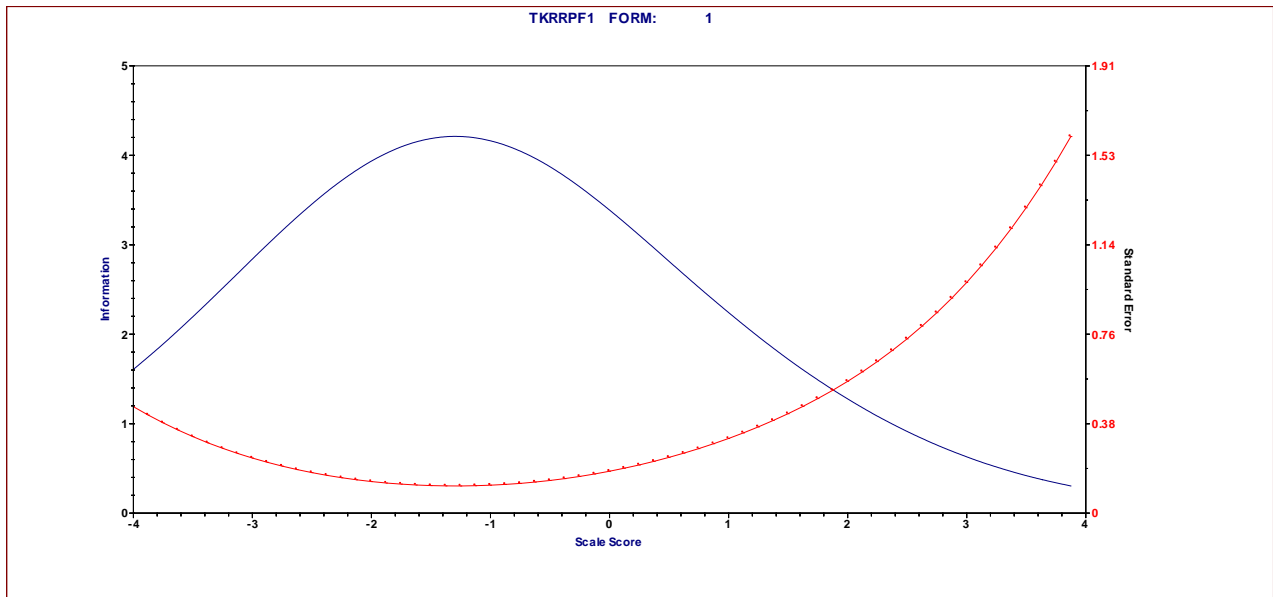
## Appendix C (Measures)

### Teacher Knowledge about Reading and Reading Practices (TKRRP)

#### *TKRRP Scale Results*

Questions	Items	Coefficient Alpha	One Parameter IRT	
			1PL IRT Reliability	Test Information Curve Maximum
13	21	.756	.762	-1.25

#### *One Parameter Test Information Curve*



## Part 2: Knowledge about Reading and Reading Practices

Mark the best response to each question.

31. Mr. Burnett noticed that some of his second graders are having difficulty reading common irregular words. To address this problem, Mr. Burnett created sets of words for students to practice. Which set is most suitable for this purpose? (Mark (X) one)
- a. when, until, which, after
  - b. sweet, sugar, milk, banana
  - c. because, does, again, their
  - d. light, house, my, they
32. In her kindergarten class, Ms. Frank uses several different tasks to help her students identify sounds in words. Which directions indicate the use of a blending task? (Mark (X) one)
- a. "Put the sounds together to say the word. /t/ /a/ /p/."
  - b. "Tell me the first sound of 'tap'."
  - c. "Say 'tap'. Now say it again but don't say /t/."
  - d. "Say each sound in 'tap'."
33. Mr. Rink asked an aide to present each of the following words orally to a group of children and to have the children tell the aide how many phonemes (speech sounds) are in each word. Help create an answer key that Mr. Rink's aide could use by marking (X) the number of phonemes contained in each word.

	1	2	3	4	5
a. freight	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. ship	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c. nation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

34. A parent asks you what to do to help Juan, her second-grade son, become a more fluent reader. Which of the following the recommendation is most likely to help Juan develop reading fluency? (Mark (X) one)
- a. Have Juan read each book several times.
  - b. Have him listen to books on tape.
  - c. Have him read on his own for 20 minutes every evening.
  - d. Read books to him every day.
35. A new third-grade teacher is having trouble picking books that are at the right reading level for his students. He asks you how he can help a student figure out whether a book is too hard. You suggest that he tell the student (Mark (X) one)
- a. to pick books on topics he/she knows something about.
  - b. to avoid books with small print and few pictures or illustrations.
  - c. not to pick books with more than five hard words on a page.
  - d. not to select books written by unfamiliar authors.
36. During reading, analysis of word structure would be a useful strategy for understanding which of the following words? (Mark (X) one)
- a. discriminate
  - b. inalterable
  - c. perspective
  - d. institution

37. Mr. Danks, a kindergarten teacher, has students learn to recite nursery rhymes (such as Little Miss Muffet) and to sing songs (such as Twinkle, Twinkle Little Star). In what way are these activities most likely to support children's early reading development? Through fostering their (Mark (X) one)

- a. understanding of story structure.
- b. enjoyment of literature.
- c. development of vocabulary.
- d. development of phonological awareness.

38. The following are common words that children are usually taught to read in grades one through three. Some are phonetically regular (i.e., they conform to frequently-taught phonic rules in English), whereas others are phonetically irregular (i.e., they are exceptions to phonic rules). Please mark (X) whether each of the following words is phonetically regular or irregular.

	Regular	Irregular
a. snowy	<input type="checkbox"/>	<input type="checkbox"/>
b. was	<input type="checkbox"/>	<input type="checkbox"/>
c. chunk	<input type="checkbox"/>	<input type="checkbox"/>
d. done	<input type="checkbox"/>	<input type="checkbox"/>
e. give	<input type="checkbox"/>	<input type="checkbox"/>
f. peach	<input type="checkbox"/>	<input type="checkbox"/>

39. Mr. Lewis' class has been learning spelling rules for adding "ing" to base words. He is looking for groups of words that illustrate the various rules to give his students a complex challenge. Which of the following groups of words would be best for this purpose? (Mark (X) one)

- a. hopping, running, sending, getting
- b. hoping, buying, caring, baking
- c. seeing, letting, liking, carrying
- d. All of the word sets are useful for this purpose.

40. Mr. Hamilton, a first-grade teacher, notices that Rafael spends much of his free time writing. He notes that Rafael misspells many words but that his misspellings suggest knowledge of some letter-sound relations. For instance, he spelled *zipper* as *zipr* and *elephant* as *elifint*. To promote Rafael's spelling development, which would be the best step for Mr. Hamilton to take? (Mark (X) one)

- a. Engage Rafael in activities that promote phonological awareness.
- b. Teach him standard spelling patterns before he spends more time writing.
- c. Teach him standard spelling patterns within the context of his compositions.
- d. Encourage him to continue to write a lot.



41 . Ms. Rico dictated the following story to her class:

**I have a black and white dog.  
Her name is Skipper.  
One day she went to my school.  
She liked playing with the kids.**

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

**I have a blk and wit bog.  
Hra name is skpr.  
Wone bay she wat to mui skul.  
She likt playg wethe the kibs.**

Which of the following words in Jesse's writing provide evidence that Jesse can identify the correct number of speech sounds in words? (Mark "Yes" or "No" for each word.)

- |                       |  |  |
|-----------------------|---|---|
| a. "blk" for black    | <input type="checkbox"/>  | <input type="checkbox"/>  |
| b. "wit" for white    | <input type="checkbox"/>  | <input type="checkbox"/>  |
| c. "skpr" for skipper | <input type="checkbox"/>  | <input type="checkbox"/>  |

42. Ms. Stanley, a kindergarten teacher, is preparing activities to teach phonological awareness in a developmentally appropriate sequence. Which of the following should she teach first? (Mark (X) one)

- a. Matching word sounds and letters.
- b. Identifying words that rhyme.
- c. Identifying vowels that say their own name.
- d. Counting the number of speech sounds in words.

43. A first-grade teacher is preparing a read-aloud lesson for her class. She is thinking about selecting four or five words from the story to discuss with the students. Which category of words below, if selected by the teacher, will most affect whether students will understand the story? (Mark (X) one)

- a. names of characters
- b. the words that are hardest to pronounce
- c. words that students will encounter in other texts
- d. specialized words in the story