Boise State University

ScholarWorks

Literacy, Language, and Culture Faculty Publications and Presentations

Department of Literacy, Language, and Culture

1-24-2023

The Teachers' Perceptions and Knowledge of Reading Assessment Survey: A Validation Study

Rachel Beachy
Plano Independent School District

Daibao Guo
Boise State University

Katherine Landau Wright
Boise State University

Erin M. McTigue
University of Stavangar

Publication Information

Beachy, Rachel; Guo, Daibao; Wright, Katherine Landau; and McTigue, Erin M. (2023). "The Teachers' Perceptions and Knowledge of Reading Assessment Survey: A Validation Study". *Reading & Writing Quarterly, 39(6), 559-581*. https://doi.org/10.1080/10573569.2022.2156954

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

The Teachers' Perceptions and Knowledge of Reading Assessment Survey: A Validation Study

Rachel Beachy
Plano Independent School District

Daibao Guo
Boise State University

Katherine Landau Wright
Boise State University

Erin M. McTigue
University of Stavanger

Abstract

Despite many calls, there is little research addressing teachers' knowledge of reading assessments and how they utilize assessment for reading instruction. Therefore, the current research developed and validated a reliable measure of teachers' perceptions and knowledge of reading assessments, called the *Perceptions and Knowledge of Assessment in Literacy Survey* (PKALS). Through the analysis of two separate administrations, we provide evidence for validity of the PKALS and examine the associations between teachers' characteristics and their performance on the PKALS. This research also found that teacher experience was correlated with higher knowledge; however, teachers' certification and graduate degree status were not. The PKALS can support future researchers and teacher preparation programs to identify gaps in teacher knowledge, allowing for interventions to promote student reading success.

Keywords: validation, reading assessment, data literacy, teacher

There is a strong and unsurprising relationship between teachers' understanding of the art of teaching in their discipline and students' learning achievement (e.g., Carlisle et al., 2009; Carlisle et al., 2011; Kelcey, 2011). It is also well-accepted that teachers should be adaptive and implement their content knowledge in a "pedagogically powerful" manner (Shulman, 1987, p. 15), adapting to the various abilities and backgrounds of the students. Logically then, as teachers grow in content knowledge, this would extend to how their students would grow in learning. Yet, although previous research spotlights teacher knowledge as it relates to specific content, the research often fails to illustrate *how* teachers' knowledge is connected to instructional decision-making process (Cunningham et al., 2004; Moats & Foorman, 2003; Phelps & Schilling, 2004).

One important aspect of a teacher's job that requires careful and intentional decision making is the use of assessment data to inform literacy instruction. Being able to effectively use assessment data requires teachers to be fluent in both the content related literacy instruction and the practice of interpreting assessment results. To date, there are existing measures of teachers' knowledge of literacy concepts including early language and literacy development (e.g., Neuman & Cunningham, 2009), teacher knowledge regarding language structures (e.g., Cunningham et al., 2009; Hindman & Wasik, 2011; Moats, 1994), and teacher beliefs about language and literacy instruction (e.g., Hindman & Waski, 2008). Furthermore, numerous researchers have explored the connections between teacher knowledge and classroom practices (e.g., Piasta et al, 2020; Schachter et al., 2016; Spear-Swerling & Zibulsky, 2014) While much of this research may begin to tap into assessment knowledge, there is a dearth of studies investigating teachers' preparedness to use assessment data effectively within their classrooms (Mandinach, 2012; Means et al., 2010; Wayman & Stringfield, 2006). Especially, there is a lack of research to provide us with comprehensive insights connecting K-12 teachers' language assessment literacy with their knowledge of literacy and language (Coombe, et al., 2020).

Although many have called for including assessment and data literacy in teacher preparation and professional development, there is relatively little research available addressing specific teacher skills related to the ability to utilize assessment for reading instruction (Carlisle et al., 2009). Educators not only need the literacy content knowledge but also the flexibility, knowledge, and adaptation necessary to adjust instruction based on students' needs. While multiple tools to assess teacher knowledge of literacy are readily available (e.g., Bos et al. 2001; Moats 1994) these do not assess data literacy or the connection between assessment and instruction. Therefore, the current research aimed to develop and validate a reliable measure of teachers' perceptions and knowledge of reading assessment called the

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Perceptions and Knowledge of Assessment in Literacy Survey (PKALS). Such a formal assessment would support future researchers, literacy coaches, and teacher preparation programs to identify the gaps in teacher knowledge, so that we can develop teachers' content knowledge in reading instruction and literacy assessment, allowing for appropriate intervention to support PreK-12 student reading success.

Several perspectives supported the development of the PKALS. In the following section, we first define the constructs of data literacy and data-driven instruction. Then we present a model of data-driven instruction and discuss how teacher preparation has played a role in the development of educators' data literacy.

Data Literacy and Data-Driven Instruction

In the United States, No Child Left Behind Act (NCLB) (2002) redefined the role of standardized testing by linking school performance to funding thus elevating the role of standardized tests (Gertsi-Pepin & Woodside-Jiron, 2005). Subsequent school reform efforts, including the establishment of Common Core State Standards (National Governors Association, 2010) have continued to elevate the role of reading assessment (and implicitly teachers' data literacy) by emphasizing high stakes assessment for students and higher quality preparation for literacy educators (Afflerbach, 2010; Blank, 2013; Jacobs et al., 2012). Most recently, the focus on the "science of reading," (SOR) a term lacking consensus in the field of reading research (Goodwin & Jimenez, 2020) has, nonetheless, maintained focus on the importance of reading assessment as well as impacted policies for teacher preparation in this area (Mosley Wetzel et al. 2020).

Taken together (NCLB, Common Core State Standards and SOR), the past twenty years have demanded for teachers' effective interpretation of student data. Instructional decision making has become the keystone to connect increased assessments and data to actual student literacy achievement. Ideally, the cycle of administering, evaluating, and interpreting assessments allows educators to differentiate literacy instruction based on student competencies and needs. Teachers' instructional priorities can then be guided by students' difficulties (Connor et al., 2009; Moats, 2014). Yet, this linkage only occurs if the data is used effectively.

This focus on assessment has led to the widespread use of terms (and concerns) such as test-centric instruction (Davis & Wilson, 2015) and assessment saturated classrooms (Kontovourki, 2012). As such, an overarching question is this: Being presented with all this assessment data, do teachers have the knowledge and skills to evaluate, integrate, and interpret assessment data in a literacy classroom? At this point, we do not fully have that answer, partially due to a lack of tools to efficiently measure teachers' data literacy.

Before proceeding further, it is necessary to define the terms *data literacy* and *data-driven instruction* – the key constructs of this research. Hamilton et al. (2009) defined data literacy as an educator's ability to ask and answer questions about the collection, analysis, evaluation, and interpretation of data. Data literacy has been found to be a learned-through-practice instructional technique, as opposed to being learned through formal education (Jacobs et al., 2012). It includes teachers' knowledge and skills that reinforce effective use of data for instruction by examining various data sources (Mandinach & Gummer, 2013a). For the purpose of the current research, teachers' data literacy is defined as how they work with data from reading assessments. We define *data-driven instruction* (see Hamilton et al., 2009; Mandinach, 2012) as systematic collection, analysis, evaluation, and interpretation of a variety of data (e.g., demographic, observational, achievement) to inform educational practice. In this study, we understand that teachers' data literacy guides and supports the implementation of data-driven instruction within their classrooms.

Although research on *data literacy* and *data-driven instruction* has received increasing attention in the field of literacy within the past few decades (Mandinach & Gummer, 2013b), there are many obstacles regarding data literacy in teacher education programs, including a lack of research on how teachers acquire data literacy. However, before these problems can be addressed, we need to understand teachers' current perception and knowledge of literacy assessments and their needs related to data literacy. Thus, it is essential to develop a reliable and validated instrument to explore teachers' perceptions, content knowledge, understanding of reading assessments, as well as their data literacy skills.

The Cycle of Data-Driven Instruction

Effectively using data to guide instructional decisions begins with evaluation, which allows teachers to gain a deeper insight into the learning needs of students (Hamilton et al., 2009). Teachers must know how to look at data through an evaluative lens that supports the analysis of student assessment results, which in turn guides instructional decisions.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

The frameworks for assessment and diagnosis, established by Cooper et al. (2006) and Hamilton et al. (2009), provided a foundation for the cycle developed for this research. This cycle of developing instruction based on students' needs, revealed by assessment data and the interpretation of data to drive that instruction, establishes the foundation for continuous reading improvement (Mokhtari et al., 2007). In this cycle, the evaluation of data is a necessary bridge between the administration of an assessment and the interpretation of assessment data for instructional decision making. Thus, understanding teachers' data literacy is essential for promoting data-driven instruction.

The Role of Teacher Preparation

Teacher education is imperative to the development of a teacher workforce (Athanases et al., 2013) and should include cycles of inquiry and investigation of student learning (i.e., data-driven instruction, Darling-Hammond, 1989, 2002; Lieberman & Miller, 2008; Lieberman & Wood, 2003). Coursework focused on the attainment and comprehension of these cycles supports data-driven schools and gives novice teachers the tools to provide appropriately differentiated instruction. Research suggests that teacher preparation programs must educate future teachers to use data from an assortment of assessments, including quantitative tests, student attendance, demographics, and student engagement, to support the development and adjustment of instruction (Greenberg & Walsh, 2012). However, little has been documented on the preparation of pre-service teachers in the acquisition of data literacy (Mandinach & Gummer, 2013a; Mann & Simon, 2010).

Similarly, with enduring importance placed on students' observable growth, in-service teachers are also expected to track student development. Consequently, they should be able to "analyze and reflect on their practice, to assess the effects of their teaching, and to refine and improve their instruction" (Darling-Hammond, 2008, p. 93). Unfortunately, the body of research on the process of assessing, collecting, and deciphering data remains minimal. Even as state and local governments place a higher emphasis on data-driven instruction (e.g., National Governors Association, 2010; No Child Left Behind Act, 2001; Texas Education Agency, 2015), limited research has been conducted to investigate the preparation of teachers' data literacy (Brookhart, 2011; Popham, 2009).

To conclude, the necessity for teachers' data literacy skills, in conjunction with limited research in this area, calls for researchers to examine teacher perceptions and knowledge related to literacy assessments. As such, the purpose of this research is to develop and validate a reliable instrument to collect information on teachers' perceptions, content knowledge, understanding of reading assessments and data literacy instruction. We developed and analyzed a survey to measure teachers' perceptions, content knowledge, and understanding of reading assessments in PreK-12 grades through two independent data collections.

Methods

To measure teachers' perceptions and knowledge of reading assessment and the interpretation of data for instruction, we developed the *Perceptions and Knowledge of Assessment in Literacy Survey* (PKALS). The dimensions of knowledge covered by the PKALS focused on the cognitive aspects of reading and their assessment, namely the five essential components of reading instruction: phonemic awareness, phonics, vocabulary, fluency, and comprehension (National Reading Panel, 2000). Below we describe the survey development, participant recruitment, and data analysis.

Survey Construction

To develop the survey, we reviewed prior related studies, which allowed us to identify items from previous research that best supported our research interests. Items were selected and adapted (if needed) from the survey of teachers' perceptions and knowledge of early reading instruction (Bos et al., 2001), the survey of teachers' knowledge of reading concepts (Moats & Foorman, 2003), and Salinger et al.'s survey of early reading instructional knowledge (2010). However, only the Bos et al. survey was used intact and represented our section for teacher knowledge of basic literacy concepts.

Furthermore, content knowledge and assessment knowledge items were adapted from state certification tests and preparation manuals, including the Michigan Test for Teacher Certification (2015), New York State Teacher Certification Exam Preparation Guide (2015), the TExES Preparation Manual (2015, Generalist EC-6) and the TExES English Language Arts and Reading (4-8), the Massachusetts Test for Educator Licensure Preparation Materials for

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Reading Specialist, (2015), and the Florida Teacher Certification Guide for Reading K-12 (2015). Additionally, items were research developed from textbooks on reading assessment (e.g., McKenna & Stahl, 2009) and common reading assessment (e.g., DIBELS).

We then consulted an expert panel of seven literacy professors from six institutions for feedback on the content and wording as well as accuracy of the information. To work to counterbalance regional biases, the faculty were from state universities across multiple regions: New York, Virginia, Maryland, Alabama, Illinois, and Nevada. Second, we administered the survey to 17 elementary reading teachers in a small urban school district in Texas, representing potential users of the survey. After the participants piloted the survey, the first author got individual feedback regarding the alignment of the survey items to their current use of assessment, confusing terminology and wording issues. We then revised the survey based on both sources of feedback. This process yielded a survey with a demographic section and three subscales (60 items). The three subscales were: (a) *Perception* subscale, consisting of 25 Likert items, measuring teacher perceptions of reading assessment and instruction; (b) *Assessment Knowledge* subscale, focused on teacher knowledge of reading assessments and data literacy (15 multiple-choice questions); and (c) a *Knowledge of Phonemic Awareness and Phonics* subscale, measuring teacher knowledge of basic literacy concepts about *phonemic awareness and phonics* (20 multiple-choice questions, taken in whole from Bos et al., 2001).

We worked to include numerous related skills throughout the subscales, including knowledge related to: a) literacy instruction in general (sample item: *Repeated readings of the same text is a way students can improve their fluency. (T/F)*); b) assessment principles (sample item: *If two standardized reading assessments both have a mean of 87, but the standard deviation (SD) of test A is 15 and the SD of test B is 5, we know that...*), c) interpretation of tests (*When given the prompt of, "What rhymes with pat?", the student responds with "Pen." This student lacks an understanding of rhyming. (T/F))* and d) translation of assessment results to instruction (*To support a highly proficient reader who reads several grade levels above other students in his class, the teacher should use yearly summative reading assessments to support and select challenging and engaging literature on that student's independent and instructional levels*.). These skill sets interact with each other for if a teacher does not have correct interpretation of assessments, then they cannot translate the results to instruction. Furthermore, we also included questions related to early reading decoding skills (phonemic awareness, sight words) and later developing skills (fluency, comprehension) so that this could be useful for a wide range of teachers, whereas many existing instruments focus only on early literacy skills.

First Administration

To begin survey validation and refinement, we recruited a national sample of teacher-participants who completed the survey through Qualtrics. We recruited potential participants through direct email requests, word of mouth, and snowball sampling including professional colleagues and personal connections. The recruitment email invited current and recent PreK-12 grade reading, literacy, and/or language arts educators to participate. While not all participants completed the entire survey, 178 participants completed the demographic and the *Perceptions* subscale, 152 participants completed the *Assessment Knowledge* subscale, and 144 participants completed the *Knowledge of Phonemic Awareness and Phonics* subscale. To maximize our sample size, we analyzed each subscale independently.

The demographic section, with 22 items identifying demographic and professional information, was analyzed descriptively. We recoded items in the *Assessment Knowledge* and *Knowledge of Phonemic Awareness and Phonics* subscales to indicate right or wrong answers and then analyzed the construct structure of each subscale through Exploratory Factor Analysis (EFA) using SPSS. For each subscale, we then analyzed inter-item reliability using Cronbach's alpha and alpha if deleted analyses. Finally, we conducted an analysis based on item response theory with both multiple-choice sections to further evaluate the validity of this instrumentation and guide item revisions and deletions for our second survey administration.

One of the primary purposes of conducting an EFA is to inform score validity (Thompson, 2004). As the items on each of the subscale were meant to measure different constructs, there was no theoretical rationale for conducting an EFA on all 60 items as one scale. Therefore, we focused on using an EFA to explore the nature of the constructs within each subscale (Thompson, 2004), and further examined the subscales for validity evidence via reliability and item response theory analyses.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Survey Validation and Refinement

Following refinement based upon the first administration, the PKALS was composed of 82 items with a demographic section and three subscales. Compared to similar measures (e.g., Saligner et al., 2010, $\alpha = 0.790$), the *Perception* subscale produced high overall reliability within three of the four subfactors. The low reliability of subfactor four (perceived knowledge of text interactions and exposure) may be related to the small number of items within this subfactor, but it is more likely due to a lack of clarity in the underlying construct. Through the process of item deletion, revision, future versions of the *Perception* subscale are intended to have more clearly defined and reliable subfactors. Although the *Assessment Knowledge* subscale produced an overall low reliability score, it is important to note that this subscale was developed anew. Analysis from the EFA and the IRT approach gave us information to improve the scale before the next survey. Finally, the *Knowledge of Phonemic Awareness and Phonics* subscale produced a relatively high overall and subfactor reliability. Compared to the Bos et al. (2001) survey from which it was taken, this survey produced a higher level of reliability. Through the process of item deletion, we reduced the number of items through IRT analysis (i.e., items 45, 50) and EFA (i.e., items 38, 46, and 48). With the deletion of the five items, the *Knowledge of Phonemic Awareness and Phonics* subscale was shortened from 20 to 15 multiple choice questions for the second administration, but it still maintained a high reliability score.

Second Survey Administration

The second data collection occurred during a two-week period. We administered the survey to PreK-12 grade teachers and collected the data in the Qualtrics online survey system. Campus-specific emails were sent to 16 schools in a small urban school district in south-central Texas.

The demographic section of the survey was analyzed through pure observation and categorization. We used a Confirmatory Factor Analysis (CFA), guided by the results of the EFA in the first administration, to analyze the three subscales. Additionally, the *Assessment Knowledge* and *Knowledge of Phonemic Awareness and Phonics* subscales were also analyzed using IRT, just as they were in the first administration.

Finally, we performed one-way Analysis of Variance (ANOVA) to determine any statistically significant relationships between participants' information from the demographic section of the survey (e.g., certification, years of teaching experience, years of teaching experiences in reading/language arts) and demonstrated knowledge of reading assessment and language.

Results

First Administration Results

Of the 178 participants who started the survey, 144 participants completed the entire survey. However, we included a participant in each analysis if they completed the target section. The demographic data for the participants in the first administration achieved the goal of a nationwide sample. As shown in Table 1, although participants from three states, Texas, Nevada, and Virginia, comprised the majority of the sample, 13 states were represented in total. Additionally, the majority of the participants were white (80.9%) females (96.1%). While not a diverse sample, it is similar to the national statistics describing the teaching force, in which 81.9% are white, and 76.3% are female (U.S. Department of Education, 2012). The sample contained much variation in years of teaching, but over half of the participants had been teaching for eight or more years (58.4%). The years of experience of these educators and the fact that the majority reported holding a master's degree characterize this sample as a particularly well-informed group of teachers.

[*Table 1* about here]

Perceptions Subscale Results – First Administration

The *Perception* subscale consisted of 25 Likert response items addressing teachers' perceived knowledge of reading instruction and assessment. Upon the first analyses of these items, the overall Cronbach's alpha was 0.793. However, the EFA determined eight possible subfactors, the item relationships were unclear, and only two subfactors had an α of 0.70 or higher. Therefore, through a combination of EFAs, Cronbach's alpha if deleted results, and theoretical rationale, we dropped five items (i.e., items 4, 7, 18,19, 24), increasing Cronbach's alpha to 0.833. The final model that best fit the data contained four subfactors. Table 2 shows how the items were grouped by subfactors.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

[Table 2 about here]

Assessment Knowledge Subscale Results – First Administration

The *Assessment Knowledge* subscale contained 15 multiple-choice items that assessed the participants' general reading assessment knowledge and data literacy. We used Excel to conduct an Item Response Theory (IRT) analysis to determine item difficulty and index discrimination. Then we used SPSS to test for reliability and conduct an EFA. We determined appropriate items to remove or revise through a combination of IRT, reliability and EFA results in a recursive manner.

The overall item difficulty of this subscale was 0.68, which was in the ideal range (Thorndike et al.,1991). The index discrimination was calculated by comparing the number of participants with mostly correct or high responses (top 27%) to the number of participants who answered the same item correctly but had mostly incorrect or low responses (bottom 27%) (Binks-Cantrell et al., 2012). Our analysis showed that ten items had a discrimination score of greater than 0.30, which is considered the reasonably good range (Ebel & Frisbie, 1986), and the overall discrimination index for all reading assessment knowledge and data literacy items was good at 0.36. However, three items had a discrimination index of ≤0.19, which indicated these items were candidates for either revisions or deletion. The IRT results are presented in Table 3.

[Table 3 about here*]*

The initial reliability analysis produced an alpha of 0.483, which is considered a poor value of internal consistency. However, the analysis also identified an improved alpha when deleting item 30. Additionally, based on the IRT analysis, we removed items 23 and 27 due to ease in answering.

Then we ran an EFA on the *Assessment Knowledge subscale*. The EFA yielded two subfactors (alpha scores are 0.487 and 0.419 respectively). However, this analysis revealed improved reliability if we deleted specific items. Therefore, we conducted another EFA, which yielded three subfactors explaining 41% of the variance. However, these subfactors all maintained lower alpha scores (.504, .436, and .289, respectively).

Knowledge of Phonemic Awareness and Phonics Subscale Results – First Administration

The *Knowledge of Phonemic Awareness and Phonics* subscale consisted of 20 multiple-choice questions. According to IRT analysis, because the items in this subscale had four to five answer choices, the ideal range of item difficulty was between 60.0% and 62.5% (Binks-Cantrell et al., 2012). Our findings showed seven items were within ±0.1 the range of difficulty, and the overall item difficulty was .66. The index discrimination analysis found 14 items were greater than .30, and the overall discrimination index for all items was .39, which all fell within the good range of scores. Three items were identified as poor, needing major revision or deletion. The IRT results are present in Table 4.

[Table 4 about here]

A maximum Cronbach's alpha was attained at $\alpha = 0.734$ with the deletion of items 45 and 50. Because Bos et al. (2001) already established these survey items as reliable ($\alpha = 0.600$), we consulted their results when revising this subscale. Then through a combination of index discrimination results and Cronbach's alpha if deleted results and EFA analyses, we deleted five items (i.e., items 45, 50, 38, 46 and 48), which shorten this subscale to 15 items.

<u>Second Administration Results</u>

Upon completion of the first administration, in which initial evidence of reliability and validity of the new survey instrument was developed and the instrument was revised, we conducted the second administration.

Demographic Information and Preparedness Results – Second Administration

The majority of the 125 participants were white (86.4%) and female (93.6%) (see Table 5). Again, while this is not a diverse sample, it does correspond to national educator averages. Nearly 60% of participants reported they took two or more reading assessment courses in college, and 80% of the participants reported they received professional development on reading assessment.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

[Table 5 about here]

The majority of participants assessed students at least once per week (96.0%). As for the time spent assessing students, 73.6% of the participants indicated they spent one to five hours assessing children. The data collected on participants' perceived preparedness showed that a majority felt adequately to well prepared to teach children to read (81.6%) and even adequately to well prepared to support the growth of struggling readers (83.2%). Finally, a large portion of the participants also felt adequately or well prepared to use the two foundational components of teaching reading for reading instruction (75.2%). In total, this group presented as reasonably confident teachers of reading with frequent experience in assessment. Detailed information is presented in Table 6.

[*Table 6 about here*]

Perception Subscale Results – Second Administration

Next, the CFA for this subscale was performed using the SPSS Amos software to examine the model fit indices and determine how well the model fit the data. The four subfactors established during the EFA remained. However, the model that best fit our data was a model without the three alpha if deleted items (identified in the previous reliability analysis) and subfactor 4. We, therefore, removed these three items and subfactor 4 from further analysis. Overall, the data indicated a probability value of 0.00, the χ 2 /df at 1.79 (good), the RMSEA at 0.08 (good), and the CFI at 0.92 (good), indicating a good model fit for the *Perception* subscale.

Assessment Knowledge Subscale Results – Second Administration

The reliability for this portion of the survey was low (Overall Cronbach's alpha = 0.488). The low reliability could partially be attributed to the number of items in this subscale or the participant sample size (Hayes, 2008). However, as these items were developed explicitly for this work and not adapted from other subscales, the current phrasing of items likely contains ambiguity of interpretation for participants. Alpha if deleted analysis demonstrated that deleting item 27 would raise the first subfactor reliability estimate to 0.463. Deleting item 31 would raise the second subfactor reliability estimate to 0.366 for an overall estimate of 0.508.

Following the guidelines defined in the first administration, 5 out of 12 items were found to be within the ideal difficulty range (≥ 0.50 and ≤ 0.70) and had a total difficulty index of 65.71. All 12 items had a discrimination index of ≥ 0.30 and a total of a 0.41 discrimination index value, meaning all items were determined to be good.

We conducted two Confirmatory Factor Analyses, one using all items in this subscale and one excluding the two alpha if deleted items. We consulted Meyers et al. (2013) in the appropriate indices to determine the best model fit. The model that best fit the data excluded those two items, so they were removed from the final subscale. The data indicated a probability value of 0.37, the χ 2 /df at 1.06 (good), the RMSEA at 0.03 (good), and the CFI at 0.96 (good), indicating a good model fit for the *Assessment Knowledge* subscale. In total, despite the low reliability via Cronbach's alpha, the CFA model fit and strong results from the IRT analysis indicate that this subscale provides a functional starting point for measuring teachers' assessment knowledge and data literacy.

Knowledge of Phonemic Awareness and Phonics Subscale Results – Second Administration

The overall reliability of this subscale was 0.733, with the first subfactor (Phonological awareness) yielding an α of 0.558 and the second (Phonics) yielding an α of 0.565. There were no items identified that would increase reliability if deleted.

Following the predetermined guidelines defined in the first administration, it was determined that 7 of the 15 items were within the ideal difficulty range of ≥ 0.50 and ≤ 0.70, with a total difficult index value of 63.88. Twelve of the 15 items were found to be good items based on their discrimination index values of ≥ 0.30, with a total value of 0.46. Thus, the IRT for this subscale further supported the validity of this survey tool.

The analysis was performed in SPSS Amos and only involved a CFA with all items since no alpha if deleted items were identified. The data indicated a probability value of 0.09, the χ 2 /df at 1.19 (good), the RMSEA at 0.05 (good), and the CFI at 0.86 (adequate), indicating an adequate to good model fit for the *Knowledge of Phonemic Awareness and Phonics* subscale.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Relationship Between Teacher Experience and Teacher Knowledge

One-way Analyses of Variance (ANOVA) were performed to describe the relationship between participant characteristics that would likely impact knowledge (i.e., presence of a Master Reading Teacher Certification, total years teaching, total years teaching English/language arts or reading), and demonstrated knowledge of reading assessment. Two subscales of the survey (i.e., *Assessment Knowle*dge and *Knowledge of Phonemic Awareness and Phonics*) contained items with a correct/incorrect response; thus, they specifically assessed participants' knowledge. A combined average correct on these two subscales was calculated for each of the participants. This "mean knowledge score" was used as the dependent variable in each of the ANOVAs. According to Levene's Test, all three analyses met the assumption of homogeneity of variances (p ranged from 0.054 to 0.319).

There were three levels of response for the question, "Do you have a master reading teacher certification (Texas) or another state-level certification specializing in reading instruction?" The three response choices were as follows: (a) Yes, (b) No, and (c) I am currently working on this certification. Table 7 below shows the results of this one-way ANOVA. Findings showed that there was no statistically significant difference between these three groups.

[Table 7 about here]

We asked participants to indicate the total number of years they had been teachers. There were five possible responses to this question: (a) 1 −3 years, (b) 4−7 years, (c) 8−15 years, (d) 16−19 years, and (e) more than 20 years. Table 8 displays the results from this one-way ANOVA. The results of this analysis indicated there were no overall statistically significant differences in this data set. However, despite the lack of statistical significance in the omnibus test, the Tukey-Kramer post hoc analysis revealed there was a difference between the highest and lowest performing groups. Specifically, educators teaching 16−19 years scored higher compared to those teaching for 1−3 years ($p = 0.047$).

[Table 8 about here]

The final ANOVA analysis considered differences between the means of participants' years teaching reading/language arts specifically ("How many years, including this year, have you taught reading/language arts?"). This question also had five possible responses: (a) 1−3 years, (b) 4−7 years, (c) 8−15 years, (d) 16−19 years, and (e) more than 20 years. The results of this one-way ANOVA can be found in Table 9.

[Table 9 about here]

The results displayed a statistically significant difference at the $p < .01$ value. The Tukey-Kramer post hoc analysis disclosed three statistically significant differences at the $p < .05$ and $p \leq .001$ levels between means of participants teaching reading/language art: (a) those teaching for 1−3 years and those teaching 4−7 years ($p = 0.024$), (b) those teaching 1−3 years and those teaching 8−15 years ($p = 0.043$), and (c) those teaching 1−3 years and those teaching 16−19 years ($p = 0.001$). In all three comparisons, the teachers with greater years teaching language arts demonstrated higher knowledge than the teachers who were within their first three years.

Discussion

The primary goal of this study was developing and collecting evidence of validation of the PKALS for the purpose of measuring teachers' perceptions and knowledge of reading assessment and data literacy. This work provides a potentially important tool for both practitioners and researchers because there is a lack of validated instruments in an area that holds increasing importance for education. Within the validation process, we considered scale structure and scale reliability and performed one-way ANOVAs to determine associations between participants' characteristics (e.g., means of participants' certifications, years of experience) and demonstrated performance. In the following section, we first discuss major findings from our first administration. Then we summarize and interpret the main findings in the second administration. Finally, we debrief findings on validation between teacher experience and teacher knowledge.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Major Findings from the First Administration

In this first administration, we worked to collect evidence of reliability (through IRT and internal structure) and validity (through factor structure) across a diverse sample of teachers, as well as find areas of weakness to revise the second administration.

Participant Demographics

A strength of our first administration was the diversity of participants across grade levels, which is important as we are developing an instrument that could be used across settings and within a variety of teacher preparation programs. This sample may have had a self-selection bias as teachers who would choose to complete a long survey on reading assessment were more likely to be highly committed to reading instruction. Accordingly, the majority of participants felt they were well prepared to teach reading (53.2%) and well prepared to teach struggling readers (50.5%). In fact, only 3.2% felt they were not prepared to teach reading. This perception of confidence is congruent to this group of teachers' reported preparation.

Reliability and Validity

The PKALS initially administered was comprised of three subscales, each measuring multiple constructs or subfactors. The *Perception* subscale produced high overall reliability within three of its four subfactors. Compared to similar measures (e.g., Saligner et al., 2010, $\alpha = 0.790$), this survey demonstrated a higher overall level of reliability, and except for subfactor 4, it demonstrated high within factor reliability. The low reliability of subfactor 4 (perceived knowledge of text interactions and exposure) in this subscale may be related to the small number of items within this subfactor, but it is more likely due to a lack of clarity in the underlying construct. Through the process of item deletion, revision, and further development of items, we attempted to improve this subfactor.

In contrast, the *Assessment Knowledge* subscale produced an overall low reliability score. It is important to note that, unlike the teacher perception subscale, this subscale was developed anew. There were also few instruments from which to adapt items. Through analysis from both the IRT approach and the EFA, the subscale was marginally improved post hoc. However, as there are no available measures in this area, this initial work provides promise and direction for future research.

The *Knowledge of Phonemic Awareness and Phonics subscale* produced a relatively high overall and subfactor reliability. Compared to the Bos et al. (2001) survey from which it was derived, our subscale produced a higher level of reliability. Through item deletion, we attempted to reduce the number of items in this subscale while maintaining a high reliability score.

Validation of the PKALS in the Second Administration

We revised the PKALS survey based on the findings from our first administration and administrated the new survey instrument. In this second administration, we recruited a sample of teachers from one district in the southwest United States. We considered both evidence of reliability and validity. Although this is a small-scale study, the wide distribution of participants' total years in the classroom and years teaching reading/language arts allowed an opportunity to analyze the depth of teachers' knowledge given varying years of experience and professional preparation. In the following section, we summarize and interpret the main findings.

Perception Subscale

The analysis demonstrated that the *Perception* subscale had high overall reliability with $\alpha = 0.857$, indicating that the items have a strong relationship with other items in the same subscale. The factor analysis indicated a four-subfactor structure; however, despite earlier revisions, the low alpha level of subfactor 4 ($\alpha = 0.453$) led to the decision to delete subfactor 4 (Perceived Knowledge of Text Interactions and Exposure) in future research. After such changes, the CFA supported the intended three subfactor structure of this subscale. In total, this subscale measures the following: (a) teacher perceptions of phonemic awareness and phonics, (b) instructional practices and student response, and (c) reading strategies and teacher actions. Additionally, this portion of the survey showed a "good" fit across all reported indices. Therefore, this subscale represents teachers' perceptions regarding reading instruction.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

In general, the findings for the *Perception* subscale indicated that teachers' perceived knowledge of phonemic awareness and phonics was that these skills play an important role in the development of early literacy and decoding, as the majority of participants indicated that they strongly agreed ($\mu = 65.0\%$) to the items in subfactor 1. For subfactor 2, teachers' perceptions were split between strongly agree ($\mu = 45.12\%$) and agree ($\mu = 41.76\%$) regarding research-based instructional practices and student responses to those practices (e.g., summarization, learning log reflections) to improve reading in the classroom. The results of the data collected for subfactor 3 indicated participants' perceived knowledge of the given reading strategies and teacher actions (e.g., modeling, explicit instruction, and thinking aloud) were appropriate practices to support reading success as indicated by their strongly agree ($\mu = 68.32\%$) response to these items. Subfactor 4 was not further analyzed.

Assessment Knowledge Subscale

As described earlier, the 12 items in the *Assessment Knowledge* subscale were researcher-developed based on prior research. After small adjustments, according to the reliability and CFA results, this subscale's reliability produced an $\alpha = 0.508$ and a "good" fit for all goodness of fit indices reported. In general, the findings from this sample of 96 educators indicated these teachers had an overall mean reading assessment knowledge and data literacy score of 65.71% for the 12 items on this subscale. Regarding distribution, only one participant got all 12 items correct; however, 60 of the 96 participants (62.5%) got at least 8 of the 12 items correct (66.7%). The data from this subscale shows the majority of participants (62.5%) had some knowledge of reading assessments and data literacy but also had room for growth.

Knowledge of Phonemic Awareness and Phonics Subscale

The 15 items of the *Knowledge of Phonemic Awareness and Phonics* subscale, which measure teachers' *knowledge* of phonemic awareness and phonics, were adapted from Bos et al. (2001). This subscale proved to have high reliability, and the reliability of this sample, $\alpha = 0.733$, was even higher than the reliability from the original administrations by Bos et al. ($\alpha = 0.600$). Factor analysis revealed two stable subfactors, and all goodness of fit indices indicated an adequate to good fit for these items. Therefore, this subscale represents a highly reliable and valid scale for use with teacher research.

For this sample, the findings indicated that 91 participant teachers had a mean score of 63.9% on the items in *Knowledge of Phonemic Awareness and Phonics*. Not a single participant got all the answers correct on these items; however, 48 of the 91 participants (52.8%) answered at least 10 of the 15 items correctly for this subscale. This data indicates that only half of educators who participated in this study demonstrated a relatively accurate knowledge of language structure critical to teaching reading, thus leaving half who did not. Unfortunately, these results mirror Bos et al.'s results (average score of 60% on language structure) published 15 years before this data was collected.

Validation Between Teacher Experience and Teacher Knowledge

Master Reading Teacher Certification

A Master Reading Teacher Certification (MRT) in the state in which the data was collected could be awarded to teachers with three years of teaching experience and the completion of a set of four graduate-level courses in reading education. Therefore, we predicted that teachers with an MRT might outperform those teachers without the additional certification. A one-way ANOVA was performed to determine if there were statistically significant differences between the knowledge scores of three groups as follows: (a) participants with an MRT, (b) without an MRT, or (c) getting an MRT certification.

Surprisingly, teachers with an MRT scored only 66% average on the knowledge questions in the survey, and findings showed no statistically significant differences in knowledge scores between those with or without an MRT certification. We believe two factors can help explain this finding. The first is the low number of teachers with MRT certification prevented meaningful comparisons between groups or generalizations from this sample, as there were only 174 teachers who reported that they had or were pursuing an MRT, and statistical significance is highly related to sample size (Thompson, 2006). More consequentially, we see this as further evidence of the gap between research and practice that is common throughout education. While our results indicate that teachers, even those with advanced education, may not be familiar with the terms and concepts inherent in data literacy, that does not mean they would

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

not benefit from this knowledge. Further research with a more diverse population is necessary to generalize results, but these preliminary findings from our validation work indicate that there is a need for additional training related to data literacy, and tools such as the PKALS to measure teachers' related knowledge base.

Years of Teaching Experience

Beyond additional professional development, we predicted that teachers with more years of teaching experience would perform better on the PKAL. A relationship between years of teaching experience and scores could provide further validation that the PKALS measures practical aspects of reading assessment. In our study, we found the highest mean knowledge score for total years teaching was from the group of participants in the 16−19 years level, at 72%. Interestingly, the trend did not continue to the most experienced teachers, as participants with 20+ years of experience did not have a higher mean score than those with 16−19 years. This may be a result of the time when they were in teacher preparation in the 1980s and early 1990s when much of the instruction was based around whole language and deemphasized both formal assessment and phonics (Alexander & Fox, 2018).

Moreover, we performed an ANOVA to determine if there was a statistical significance of the knowledge scores of the five subgroups as follows: (a) 1−3 years, (b) 4−7 years, (c) 8−15 years, (d) 16−19 years, and (e) more than 20 years. Interestingly, although the result yielded no overall statistically significant difference in this data set, the Tukey-Kramer post hoc analysis revealed there was a difference between the highest and lowest performing groups. Specifically, those educators teaching 16−19 years compared to those teaching 1−3 years ($p = 0.047$). This finding is consistent with previous research, demonstrating that teacher knowledge tends to increase with years of experience (Bos et al., 2001; Goldhaber, 2002; Kraut et al., 2016; Salinger et al., 2010; Stronge et al., 2011).

Years Teaching Reading/Language Arts

Similarly, we predicted that teachers who have been teaching reading/language arts specifically for multiple years would perform better on the PKAL. For years of teaching reading/language arts, we categorized participants in the following five groups based on their responses: (a) 1−3 years, (b) 4–7 years, (c) 8−15 years, (d) 16−19 years, and (e) more than 20 years. The overall ANOVA results yielded a statistically significant difference at the $p < .01$ value. The Tukey-Kramer post hoc analysis disclosed three statistically significant differences between means of participants teaching reading/language art: (a) those teaching for 1−3 years and those teaching 4−7 years ($p = 0.024$); (b) those teaching 1−3 years and those teaching 8−15 years ($p = 0.043$); and (c) those teaching 1−3 years and those teaching 16−19 years ($p = 0.001$). The teachers with greater years teaching language arts demonstrated higher knowledge than the teachers who were within their first three years. Regarding our goal at providing evidence of validation, the trend that teachers with more experience, particularly for language arts and reading, performed better on the PKAL than newer teachers provide additional evidence.

It is important to highlight that we found that participants with 1−3 years teaching experience in reading/language arts had the lowest mean knowledge score with 50%. One may expect that their recent training would prepare them for knowledge on reading assessment. This finding, that novice teachers did not perform better, is concerning because these topics (assessment, phonics, research-based reading skills) are currently prioritized by teacher preparation for literacy instruction. Research indicates there are two likely possibilities for this result of underperformance by new teachers. First, emerging from recent debates regarding Science of Reading (see Goodwin & Jimenez, 2020 for an overview of issues), is concern that preservice teachers are not effectively prepared to teach and assess phonemic awareness, phonics, and other research-based approaches (e.g., Moats, 2020). A second, more nuanced, possibility, is that the information is presented, but the manner in which preservice teachers are instructed results in knowledge that can remain at a superficial level, or inert, as they proceed to the classroom (Englert, et al. 2019). Likely, classroom instruction itself is insufficient and practical experience is also necessary to fully understand the complexities of reading assessment and interpretation, so this is a competency developed by a combination of formal instruction and applied experience.

Limitations

One limitation of this research is that, due to the location and personal connections of the authors, we recruited significantly more teachers from Texas than from other geographic locations. Additionally, while our sample population reflected that of the teacher workforce in the United States, it was primarily composed of white female teachers. Future research should include more diverse populations of teachers to generalize results. Finally, the use of

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

snowball technique, particularly paired with an incentive, reduces researcher control of the sample as was can only rely on self-report. However, as researchers, we able to have high control of the second sample (used in the CFA) due to cooperation and support from a single school district.

Additionally, our final subscales contained inconsistent numbers of items. This was likely a result of the fact that more items existed for topics related to teacher knowledge than data literacy, so there were more items in our initial pool related to this topic. However, all subscales yielded acceptable reliability estimates, indicating that the small number of items in some areas did not have a grave impact on the scale. Finally, in the second survey administration, we conducted CFA with results from 125 participants. The small sample size could be a limitation of our study, although model fit is still good. Therefore, future research can address this issue by increasing the sample size.

Implications & Future Research

Our results provide a call for further research to better understand the needs of educators acquiring data literacy. Most concerning was the low demonstrated knowledge of teachers who were in their first three years of teaching as they had likely completed their reading assessment courses recently but had not mastered the complexity of assessment and interpretation. Unfortunately, while this survey development work opens questions about teacher preparation and professional development for reading assessment, it does not yet provide clear answers. To do so, we would recommend using such a tool, with a larger systematic sampling of teachers and more widely disseminated survey boundaries, would yield better understanding about what educators currently know about reading assessment and how to support their learning through professional development as well as implications for teacher preparation. Furthermore, future work in this area should quantify the types and amount of literacy coursework reported by survey takers. Additionally, as the data indicates that assessment knowledge may develop through a combination of content knowledge and practical application, longitudinal work that follows teachers within their preparation programs and throughout their early career of teaching would provide insight into where teachers are gaining their knowledge.

We envision that this survey could be utilized as a pre- and post- assessment for school administrators or regional service providers in determining teachers' areas of strengths and needs prior to planning professional development in reading assessment. Furthermore, at the university level, this survey could be used as a benchmark to measure preservice teachers' knowledge before, during and after field placements, where they would be gaining practical experience for administering and interpreting reading assessments. Additionally, this survey could provide evidence of growth for a reading assessment course, or, at the program level, for teachers engaged in preparation for advanced certifications in literacy.

As an instructional implication, it is important to address that our findings suggest that teachers are entering the profession without mastery of reading assessment and data literacy, but appear to develop more competency as they proceed in their teaching. We echo the views highlighted by Torotelli and colleagues (2021) that pre-service teachers must not have only discrete knowledge about literacy and assessment but also pedagogical or situated knowledge in order to maintain and apply their knowledge. Therefore, we would encourage teacher preparation programs, whether through the increased use of case studies, individual assessment and tutoring of K-12 students, or more systematically developed within field placements, to help students translate their assessment and literacy skill knowledge to a more applied level. We also encourage instructors of reading assessment courses to assess student knowledge in applied manners, in addition, to discrete knowledge. For example, students could be given a range of assessment information, and challenged to translate this data into a profile of strengths and weaknesses, as well as translate to instructional recommendations.

Conclusion

The findings of our first administration provide the basis for a valid and reliable instrument to capture teachers' perceptions and knowledge regarding reading assessment. The use of both IRT and traditional reliability analyses provided a rigorous manner to analyze the items and adapt the instrument. The instrument, therefore, was refined and employed with a unique sample of participants in the second administration of the new survey instrument.

In summary, findings from the second administration provided evidence for validating this instrument. First, through Confirmatory Factor Analysis, reliability analysis, and the relationship between scores and teacher experience, this study provided evidence that the PKALS can be used to measure teacher perceptions of reading assessment, as well

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

as teachers' content knowledge of reading assessment, language structure and data literacy. These findings also indicate that the teachers in this sample, with more years teaching language arts and reading, demonstrated greater knowledge of reading assessment and reading instruction than their less experienced peers.

References

Afflerbach, P. (2010). *Essential Readings of Assessment.* Newark, DE: International Reading Association.

Alexander, P. A., & Fox, E. (2018). Reading research and practice over the decades: A historical analysis. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 51-80). Routledge.

Athanases, S. Z., Bennett, L. H., & Wahleithner, J. M. (2013). Fostering data literacy through preservice teacher inquiry in English language arts. *The Teacher Educator, 48,* 8−28. https://doi.org/10.1080/08878730.2012.740151

Binks-Cantrell, E., Joshi, R. M., & Washburn, E. K. (2012). Validation of an instrument for assessing teacher knowledge of basic language constructs of literacy. *Annals of Dyslexia, 62*, 153−171. https://10.1007/s11881-012-0070-8

Binks-Cantrell, E., Washburn, E. K., Joshi, R. M., & Hougen, M. (2012). Perter effect in preparation of reading teachers. *Scientific Studies of Reading, 16,* 526−536. https:// doi.org/10.1080/10888438.2011.601434

Blank, R. K. (2013). What research tells us: Common characteristics of professional learning that leads to student achievement. *Journal of Staff Development, 34, 50−53.*

Bos, C., Mather, N., Dickson, S., Podhajski, B., & Chard, D. (2001). Perceptions and knowledge of preservice and inservice educators about early reading instruction. *Annals of Dyslexia, 51*, 97−120. https://doi.org/10.1177/002221940103400508

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30,* 3−12. https:// doi.org/10.1111/j.1745-3992.2010.00195.x

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing: An Interdisciplinary Journal, 22*, 457−486. https://doi.org/10.1007/s11145-009-9165-y

Carlisle, J. F., Kelcey, B., Rowan, B., & Phelps, G. (2011). Teachers' knowledge about early reading: Effects on students' gains in reading achievement. *Journal of Research on Educational Effectiveness, 4*, 289−321. https://doi.org/10.1080/19345747.2010.539297

Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn?. *Language Testing in Asia, 10*, 1-16.

Cooper, J. D., Chard, D. J., & Kiger, N. D. (2006). *The struggling reader: Interventions that work.* Scholastic.

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., Underwood, P., & Morrison, F. J. (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development*, *80*, 77−100. https://doi.org/10.1111/j.1467-8624.2008.01247.x

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia, 54*, 139−167. https://doi.org/10.1007/s11881-004-0007-y

Cunningham, A. E., Zibulsky, J., & Callahan, M. D. (2009). Starting small: Building preschool teacher knowledge that supports early literacy development. *Reading and Writing, 22*, 487-510. https://doi.org/10.1007/s11145-009-9164-z

Darling-Hammond, L. (1989). Accountability for professional practice. *Teachers College Record, 91*, 60−80.

Darling-Hammond, L. (2002). *Learning to teach for social justice*. Teachers College Press.

Darling-Hammond, L. (2008). Teacher learning that supports student learning. *Teaching for Intelligence, 2*, 91−100.

Davis, D. S., & Willson, A. (2015). Practices and commitments of test-centric literacy instruction: Lessons from a testing transition. *Reading Research Quarterly, 50*, 357-379. https://doi.org/10.1002/rrq.103

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of educational measurement.* Prentice Hall.

Englert, C.S., Mariage, T.V., Truckenmiller, A.J., Brehmer, J., Hicks, K., & Chamberlain, C. (2020). Preparing special education preservice teachers to teach phonics to struggling readers: Reducing the gap between expert and novice performance. *Teacher Education and Special Education, 43*, 235–256. https://doi.org/10.1177/0888406419863365

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Gerstl-Pepin, C. I., & Woodside-Jiron, H. (2005). Tensions between the "science" of reading and a "love of learning": One high-poverty school's struggle with NCLB. *Equity & Excellence in Education, 38(3)*, 232-241.

Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teacher about K-12 assessment: A review.* Washington, DC: National Council on Teacher Quality.

Goldhaber, D. (2002). The mystery of good teaching. *Education Next*, *2*, 50−55.

Goodwin, A. P., & Jiménez, R. T. (2020). The science of reading: Supports, critiques, and questions. *Reading Research Quarterly, 55*, S7-S16.

Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Hayes, B. E. (2008). *Measuring customer satisfaction and loyalty: Survey design, use, and statistical analysis methods* (3rd ed.). Quality Press.

Hindman, A. H., & Wasik, B. A. (2008). Head Start teachers' beliefs about language and literacy instruction. *Early Childhood Research Quarterly, 23*, 479-492. doi: 10.1016/j.ecresq.2008.06.002

Hindman, A. H., & Wasik, B. A. (2011). Measuring teachers' knowledge about early language and literacy: Practical implications and considerations. *NHSA Dialog, 14*(4), 351-356. https://doi.org/10.1080/15240754.2011.618647

Jacobs, J., Gregory, A., Hoppey, D., Yendol-Hoppey, D. (2012). Data literacy: Understanding teachers' data use in a context of accountability and response to intervention. *Action in Teacher Education, 31,*41−55. https://doi.org/10.1080/01626620.2009.10463527

Kelcey, B. (2011). Assessing the effects of teachers' reading knowledge on students' achievement using multilevel propensity score stratification. *Educational Evaluation and Policy Analysis, 33*, 458−482. https://doi.org/10.3102/0162373711415262

Kontovourki, S. (2012). Reading leveled books in assessment-saturated classrooms: A close examination of unmarked processes of assessment. *Reading Research Quarterly*, *47*, 153−171. https: doi.org/10.1002/RRQ.014

Kraut, R., Chandler, T., & Hertenstein, K. (2016). The interplay of teacher training, access to resources, years of experience and professional development in tertiary ESL reading teachers' perceived self-efficacy. *Gist: Education and Learning Research Journal*, *12,* 132-151. https://doi.org/10.26817/16925777.248

Lieberman, A., & Miller, L. (2008). *Teachers in professional communities: Improving teaching and learning*. Teachers College Press.

Lieberman, A., & Wood, D. R. (2003). *Inside the national writing project: Connecting network learning and classroom teaching*. Teachers College Press.

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist, 47*, 71-85. https://doi.org/10.1080/00461520.2012.667064

Mandinach, E. B., & Gummer, E. S. (2013a). Building educators' data literacy: Differing perspectives. *Journal of Educational Research & Policy Studies, 13*, 1-5

Mandinach, E.B., & Gummer, E. S. (2013b). A systematic view of implementing data literacy in educator preparation. *Educational Researcher,42,* 30-37. https://doi.org/10.3102/0013189X12459803

Mann, B., & Simon, T. (2010, July). *Teaching teachers to use data*. Presentation at the NCES STATS-DC 2010 Data Conference, Bethesda, MD.

McKenna, M. C., & Dougherty Stahl, K. A. (2009). *Assessment for reading instruction* (3rd ed.). The Guildford Press.

Means, B., Padilla, C., & Gallagher, L. (2010). *Use of education data at the local level: From accountability to instructional improvement*. U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. http://www2.ed.gov/rschstat/eval/tech/use-of-education-data/use-of-education-data.pdf

Meyers, L. S., Gamst, G., & Guarino, A. J. (2013). *Applied multivariate research: Design and interpretation.* Sage Publications.

Moats, L. C. (2014). What teachers don't know and why they aren't learning it: Addressing the need for content and pedagogy in teacher education. *Australian Journal of Learning Difficulties, 19*, 75−91. https://doi.org/10.1080/19404158.2014.941093

Moats, L. C. (2020). Teaching reading is rocket science: What expert teachers of reading should know and be able to do. *American Educator, 44*(2), 4.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Moats, L. C., & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia, 53*, 23−45. https://doi.org/10.1007/s11881-003-0003-7

Mokhtari, K., Rosemary, C. A., & Edwards, P. A. (2007). Making instructional decisions based on data: What, how, and why. *The Reading Teacher, 61,*354−359. https://doi.org/10.1598/RT.61.4.10

Mosley Wetzel, M., Skerrett, A., Maloch, B., Flores, T. T., Infante-Sheridan, M., Murdter-Atkinson, J., ... & Duffy, A. (2020). Resisting positionings of struggle in "science of teaching reading" discourse: Counterstories of teachers and teacher educators in Texas. *Reading Research Quarterly*, 55, 319-330.

National Governors Association (2010). *Common Core State Standards*. National Governors Association Center for Best Practices, Council of Chief State School Officers.

Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532-566. https://doi.org/10.3102/000283120832808

No Child Left Behind Act of 2001, ESEA, 2001, Title 1, Part B, Subpart 1, Section 1202(c)(7)(A)(IV)(2).

National Reading Panel (US), National Institute of Child Health, Human Development (US), National Reading Excellence Initiative, National Institute for Literacy (US), & United States Department of Health. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Phelps, G., & Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *The Elementary School Journal, 105*, 31−48.

Piasta, S. B., Ramirez, P. S., Farley, K. S., Justice, L. M., & Park, S. (2020). Exploring the nature of associations between educators' knowledge and their emergent literacy practices. *Reading and Writing, 33*, 1399-1422. https://doi.org/10.1007/s11145-019-10013-4

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice, 48,* 4−11. https://doi.org/10.1080/00405840802577536

Salinger, T., Mueller, L., Song, M., Jin, Y., Zmach, C., Toplitz, M., Partridge, M., & Bickford, A. (2010). *Study of Teacher Preparation in Early Reading Instruction* (NCEE 2010-4036). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Schachter, R. E., Spear, C., Piasta, S., Justice, L. M., & Logan, J. A. R. (2016). Early childhood educators' knowledge, beliefs, education, experiences, and children's language- and literacy-learning opportunities: What is the connection? *Early Childhood Research Quarterly, 36*, 281-294. doi: 10.1016/j.ecresq.2016.01.008

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1−21.

Spear-Swerling, L. & Zibulsky, J. (2014). Making time for literacy: Teacher knowledge and time allocation in instructional planning. *Reading and Writing, 27*, 1353-1378. https://doi.org/10.1007/s11145-013-9491-y

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of teacher Education*, *62*, 339−355. https://doi.org/10.1177/0022487111404241

Texas Education Agency (2007−2015). *Student assessment division*. Retrieved March 19, 2015, from http://tea.texas.gov/student.assessment/

Thompson, B. (2004). *Exploratory and confirmatory factor analysis.* American Psychological Association.

Thompson, B. (2006). *Foundations of behavioral statistics*. The Guilford Press.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education.* Macmillan.

Tortorelli, L. S., Lupo, S. M., & Wheatley, B. C. (2021). Examining teacher preparation for code-related reading instruction: An integrated literature review. *Reading Research Quarterly, 56*, S317-S337.

Wayman, J. C., & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education, 11,* 549−571. https://doi.org/10.1086/505059

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2012). *Number and percentage distribution of teachers in public and private elementary and secondary schools, by selected teacher characteristics.* [Data file]. Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_209.10.asp

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 1

First Administration: Participant Demographics (n = 178)

	n	Percentage
Gender		
Male	7	3.9%
Female	171	96.1%
Ethnicity		
African American	9	5.1%
Asian/Pacific Islander	4	2.2%
Hispanic	16	9.0%
White	144	80.9%
Other	5	2.8%
Years Teaching		
1−3 years	32	18.0%
4−7 years	42	23.6%
8−15 years	53	29.7%
16−20 years	19	10.7%
20 or more years	32	18.0%
Years Teaching Reading/Language Arts		
1−3 years	47	26.4%
4−7 years	39	21.9%
8−15 years	50	28.1%
16−20 years	18	10.1%
20 or more years	24	13.5%
Current Grade Level Taught		
Early Childhood (PK-Kinder)	30	16.9%
Elementary (1st–4th grade)	93	52.2%
Intermediate or Middle School (5th–8th grade)	39	21.9%
High School (9th–12th grade)	16	9.0%
Master's Degree		
Yes, in Reading/Language Arts	41	23.0%
Yes, in another content area	51	28.7%
In Progress	34	19.1%
No	52	29.2%
Specialized Reading Certification		
Yes	40	22.5%
No	126	70.8%
In Progress	12	6.7%
Current State in Which You Teach		
California	5	2.8%
Georgia	1	0.6%
Illinois	7	3.9%
Iowa	5	2.8%
Louisiana	1	0.6%
Massachusetts	2	1.1%
Mississippi	1	0.6%
Nevada	21	11.8%
Oklahoma	2	1.1%
Tennessee	1	0.6%
Texas	107	60.1%
South Carolina	1	0.6%
Virginia	24	13.4%

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Certification Area (participants could select multiple responses)		
	131	29.6%
Generalist (1st–6th)	72	16.3%
Early Childhood (PK–K)	11	2.5%
Bilingual Educator	63	14.3%
English as a Second Language	30	6.8%
Special Education	35	7.9%
Reading Specialist/Master Reading Teacher	57	12.9%
Content Specific	1	0.2%
Educational Diagnostician	15	3.4%
Educational Administration	27	6.1%
Other		
College Reading Assessment Courses Taken		
0	19	10.7%
1	29	16.3%
2	58	32.6%
3 or more	72	40.4%
Professional Development on Reading Assessment		
Yes	138	77.5%
No	40	22.5%

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 2

First Administration: Likert Items Exploratory Factor Analysis, Cronbach's α, and α if Deleted Items

Item	Subfactor **1** Phonemic Awareness and Phonics	Subfactor **2** Instructional Practices and Student Actions	Subfactor **3** Reading Strategies and Teacher Actions	Subfactor **4** Text Interactions and Exposure
13. Phonics instruction promotes decoding skills.	**.761**	.434	.464	.292
10. K-2 teachers should know how to teach phonics (letter/sound correspondences).	**.729**	.241	.441	.153
16. Phonics instruction is beneficial for children who are struggling to learn to read.	**.607**	.553	.208	.397
14. It is important for teachers to demonstrate to struggling readers how to segment words into phonemes when reading.	**.594**	.414	.302	.167
1. K-2 teachers should know how to teach and assess phonological awareness.	**.553**	.240	.289	.153
21. Effective instruction for word recognition and decoding emphasizes students' development of graphophonemic skills.	.374	**.720**	.393	.183
23. To grow students' understanding of the relationship between written and spoken word, a teacher could read aloud from a big book while pointing to each word as its read.	.422	**.600**	.482	.362
22. Literal comprehension instruction can include retelling the beginning, middle, and end of a story.	.337	**.567**	.307	.393
15. Direct, explicit, instruction in phonemic awareness supports a student's ability to rhyme.	.289	**.519**	.245	.327
25. Having students write in a learning log about what they learned and what they do not understand, during and after reading, supports self-monitoring to improve comprehension.	.222	**.447**	.412	.221
17. The teacher thinking aloud during reading promotes students' active construction of meaning and comprehension.			.346 .309	**.667** .145

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

12. Teacher modeling of skills during guided reading will help foster student's ability to utilize these skills.	.388	.249	**.631**	.172
20. Comprehension can be supported through teaching students explicit strategies to monitor their understanding.	.270	.551	**.564**	.111
11. Picture cues can help children identify words in the early stages of reading.	.371	.108	**.538**	.242
8. Children should read different types of text for different instructional purposes (i.e., decodable texts, genre based children's literature, rhyming texts).	.415	.435	**.463**	.278
9. Repeated readings of the same text is an example of an instructional strategy to improve fluency.	.429	.313	**.451**	.326
5. Time children spend reading or being read contributes directly to reading improvement.	.135	.309	**.397**	.223
3. Controlling text through consistent spelling patterns (The fat car sat on a hat.) is an example of an effective method for children who struggle to learn to identify words.	.329	.416	.167	**.782**
2. Literacy experiences in the home contribute to early reading success.	.103	.147	.220	**.387**
6. Learning to use context clues (syntax and semantics) is more important than learning to use graphophonemic cues (letters and sounds) when learning to read.	-.255	.128	.058	**.355**
Cronbach's α	.774	.721	.728	.424
Overall Cronbach's α			**.833**	

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 3

First Administration: Item Response Theory Results in Assessment Knowledge Subscale

Item	Item Difficulty	Index Discrimination
19. Examining a class reading average on a district benchmark/unit assessment does not...	88.19	0.21
20. An example of a formative reading assessment is…	51.39	0.62
21. If a teacher wants an assessment that gives current data to be used for the adjustment of instructional goals for a student, the teacher should use…	45.83	0.54
22. A student's independent, instructional and frustrational reading levels, can be found through the administration of…	57.64	0.36
23. A running record is an example of…	91.67	0.13
24. To interpret a norm-referenced reading assessment a teacher compares a student's raw score to…	40.97	0.44
25. On a given assessment, percentile rank is the…	63.19	0.49
26. A criterion-referenced reading assessment is useful in determining the…	86.11	0.28
27. A standardized assessment…	93.06	0.13
28. The international Literacy Association advises educators to ignore grade-equivalent scores, because they…	68.75	0.31
29. To best assess a student's knowledge of a specific reading skill, a teacher or specialist should use…	57.64	0.41
30. A student who can identify 7 letter-sound correspondences and can blend and segment CVC words presented orally is most likely ready to…	32.64	0.15
31. If two standardized reading assessments both have a mean of 87, but the standard deviation of Test A is 15 and the SD of Test B is 5, we know that…	77.78	0.49
32. An observational checklist of a student's reading behaviors allows the teacher to…	82.64	0.46
33. Standard deviation is the term used to identify the…	83.33	0.36
Overall Totals	**68.06**	**0.36**

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 4

First Administration: The Knowledge of Phonemic Awareness and Phonics Subscale Item Response Theory Results

Item	Item Difficulty	Index Discrimination
34. Which word contains a short vowel sound?	92.36	0.26
35. A phoneme refers to:	90.28	0.31
36. A pronounceable group of letters containing a vowel sound is a:	71.53	0.41
37. If *tife* were a word, the letter "i" would probably sound like the "i' in:	95.83	0.08
38. A combination of two or three consonants pronounced so that each letter keeps its own identity is called a…	70.83	0.62
39. An example of a voiced and unvoiced consonant pair would be:	40.28	0.41
40. Two combined letters that represent one single speech sound are a:	65.28	0.67
41. How many speech sounds are in the word "eight"?	87.50	0.28
42. How many speech sounds are in the word "box"?	17.36	0.41
43. How many speech sounds are in the word "grass"?	47.22	0.64
44. What type of task would this be? *Say the word "cat." Now say cat without the /c/ sound.*	80.56	0.46
45. What type of task would this be? *I am going to say some sounds that will make one word when you put them together. What does /sh/ /oe/ say?*	89.58	0.56
46. Mark the statement that is false:	50.00	0.56
47. What is the second sound in the word "queen"?	36.11	0.23
48. A reading method that focuses on teaching the application of speech sounds to letters is called:	61.11	0.46
49. A soft c is in the word:	86.81	0.33
50. Identify a pair of words that begin with the same sound:	99.31	0.03
51. All of the following nonsense words have silent letters, except:	47.92	0.67
52. If you say the word, and then reverse the order of the sounds, **ice** would be:	73.61	0.64
53. If you say the word, and then reverse the order of the sounds, **enough** would be:	67.36	0.59
Overall Totals	**66.39**	**0.39**

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 5

Second Administration: Participant Demographics (n = 125)

	n	Percentage
Gender		
Male	8	6.4%
Female	117	93.6%
Ethnicity		
African American	1	0.8%
Asian/Pacific Islander	0	0.0%
Hispanic	15	12.0%
White	108	86.4%
Other	1	0.8%
Years Teaching		
1−3 years	12	9.6%
4−7 years	23	18.4%
8−15 years	50	40.0%
16−19 years	13	10.4%
20 or more years	27	20.6%
Years Teaching Reading/Language Arts		
1−3 years	24	19.2%
4−7 years	21	16.8%
8−15 years	46	36.8%
16−19 years	14	11.2%
20 or more years	20	16.0%
Current Grade Level Taught		
Early Childhood (PK-Kinder)	20	16.0%
Elementary (1st–4th grade)	68	54.4%
Intermediate or Middle School (5th–8th grade)	20	16.0%
High School (9th–12th grade)	17	13.6%
Master's Degree		
Yes, in Reading/Language Arts	7	5.6%
Yes, in another content area	46	36.8%
In Progress	12	9.6%
No	60	48.0%
Specialized Reading Certification		
Yes	17	13.6%
No	105	84.0%
In Progress	3	2.4%

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Certification Area (participants could select multiple responses)		
Generalist (1st – 6th)	98	26.9%
Early Childhood (PK – K)	60	16.6%
Bilingual Educator	13	3.6%
English as a Second Language	78	21.4%
Special Education	26	7.1%
Reading Specialist/Master Reading Teacher	13	3.6%
Content-Specific	44	12.1%
Educational Diagnostician	2	0.5%
Educational Administration	12	3.3%
Other	18	4.9%
College Reading Assessment Courses Taken		
0	31	24.8%
1	20	16.0%
2	34	27.2%
3 or more	40	32.0%
Professional Development on Reading Assessment		
Yes	100	80.0%
No	25	20.0%

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 6

Second Administration: Participants General Reading Assessment Experience and Perceived Instructional Preparedness (n = 125)

	n	Percentage
Frequency of Instructional Decisions Made Based on Students' Data		
Every day	96	76.8%
At least once per week	24	19.2%
Other	5	4.0%
Frequency Students are Assessed		
Every day	39	31.2%
At least once per week	50	40.0%
Every 2−3 weeks (Progress monitor)	31	24.8%
Once per grading period	5	4.0%
Time Spent Assessing based on Frequency Assessed		
1−2 hours	64	51.2%
3−5 hours	28	22.4%
6−10 hours	6	4.8%
10 or more hours	2	1.6%
Other	25	20.0%
Perceived Preparedness to Teach Children to Read		
Not Prepared	6	4.8%
Somewhat Prepared	17	13.6%
Adequately Prepared	40	32.0%
Well Prepared	62	49.6%
Perceived Preparedness to Support the Growth of Struggling Readers		
Not Prepared	2	1.6%
Somewhat Prepared	19	15.2%
Adequately Prepared	50	40.0%
Well Prepared	54	43.2%
Perceived Preparedness to Use Phonological Awareness and Phonics in Teaching Reading		
	11	8.8%
Not Prepared	20	16.0%
Somewhat Prepared	33	26.4%
Adequately Prepared	61	48.8%
Well Prepared		

Table 7

Second Administration: ANOVA Results for Mean Knowledge Score by Master Reading Teacher Certification Status

				Master Reading Teacher Certification Status					
				Yes MRT (*n* = 14)		**No MRT (*n* = 74)**		**Working on MRT (*n* = 3)**	
	df	*F*	*p*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mean Knowledge Score	2	0.482	0.619	0.66	0.17	0.63	0.20	0.73	0.13

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2022.2156954

Table 8

Second Administration: ANOVA Results for Mean Knowledge Score by Total Years Teaching

				Total Years of Teaching									
				1-3 years (n = 10)		**4-7 years (n = 15)**		**8-15 years (n = 32)**		**16-19 years (n = 12)**		**20+ years (n = 22)**	
	df	*F*	*p*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mean Knowledge Score	4	2.083	0.090	0.49	0.23	0.65	0.19	0.64	0.13	0.72	0.25	0.65	0.21

Table 9

Second Administration: ANOVA Results for Mean Knowledge Score by Years Teaching Reading/Language Arts

				Years of Teaching Reading/Language Arts									
				1−3 years (n = 19)		**4−7 years (n = 15)**		**8−15 years (n = 28)**		**16−19 years (n = 12)**		**20+ years (n = 17)**	
	df	*F*	*p**	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mean Knowledge Score	4	4.617	0.002	0.50	0.21	0.69	0.13	0.65	0.14	0.77	0.25	0.64	0.19

* $p < .01$