NCEE 2010-4036

U.S. DEPARTMENT OF EDUCATION

Study of Teacher Preparation in Early Reading Instruction

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

Study of Teacher Preparation in Early Reading Instruction

SEPTEMBER 2010

Terry Salinger
Lorin Mueller
Mengli Song
Ying Jin
Courtney Zmach
Michele Toplitz
American Institutes for Research

Mark Partridge
Adam Bickford
Optimal Solutions Group LLC

NCEE 2010-4036
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
Rebecca Maynard
Commissioner

September 2010

This report was prepared for the Institute of Education Sciences under Contract No. ED-04-CO-0062/0001. The project officer was Tracy Rimdzius in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the reports.

To order copies of this report,

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 22207, Alexandria, VA 22304.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 703-605-6794.
- Order online at www.edpubs.gov.

This report also is available on the IES website at http://ies.ed.gov/ncee.

Upon request, this report is available in alternate formats such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

ACKNOWLEDGMENTS

This study represents a collaborative effort between Optimal Solutions Group LLC (Optimal) and the American Institutes for Research (AIR). We appreciate the willingness of 99 teacher preparation programs to participate in the study and of the 2,237 pre-service teachers who completed the *Pre-service Teacher Preparation Program and Knowledge Survey.* Additionally, we are grateful for the interest in the study shown by 15 reading researchers, 19 reading teacher educators, and 28 recent college graduates, who also agreed to take the *Pre-service Teacher Preparation Program and Knowledge Survey.* Their efforts provided us comparison groups against which to gauge the performance of the pre-service teachers.

Our work was strengthened by our expert technical working group. Members included Joanne Carlisle, University of Michigan; Donald Compton, Vanderbilt University; Linnea Ehri, City University of New York; Geoffrey Phelps, University of Michigan; Ray Reutzel, Utah State University; and Joanna Williams, Teachers College.

Work on this study consisted of many stages, and numerous people have been involved during the development of the measures, recruitment of institutions and students, administration of the *Pre-service Teacher Preparation Program and Knowledge Survey,* and data analysis. At AIR, we wish to acknowledge David Baker, Rebecca Holland-Coviello, Andrea Muse, Derrick Johnson, Jim Kurtessis, Nancy Matheson, Mark Smith, Stephane Baldi, Elizabeth Sparano, Catherine Lopez, Melanie Skemer, Charlene Impey, Amy Holtzman, and Juliet Jason. At Optimal, we recognize Mark Turner, Monica Leal Priddy, Eric Asongwed, Amy Peterson, Andrew Gendreau, Annie Oonjit, Brandon Newbury, Laurel Davis Rasinski, Charlene Gatewood, Alexander Kyerematen, Deanna Bussard, Shelley Marshall, Helga Gross, Anna Habash, Michelle Blair, Allison Buckley, Ja-Nee Jackson, Veronika Peleshchuk, Diana Behrendt, Grace Hong, Kelly Gent, and Christopher Hill.

Finally, we wish to thank Michael Garet, Dan Sherman, Tsze Chan of AIR and Larry Dolan and Toks Fashola of Optimal for technical reviews; and Holly Cutting Baker and Aida Gomez for helping to make the report useful and understandable.

Contents

List of Tables

List of Figures

EXECUTIVE SUMMARY

A component of the *No Child Left Behind* Act (NCLB) (PL 107-110) is its emphasis on the importance of systematic and explicit instruction in early reading using practices that are grounded in scientific research. The Reading First legislation (Title I, Part B, Subpart 1)[1] within NCLB is designed to support state and local education agencies so that they can in turn base their early reading instruction on scientific research and focus on five "essential components" of early reading instruction, as defined by the legislation and informed by the National Reading Panel[2]: (1) phonemic awareness; (2) phonics; (3) vocabulary development; (4) reading fluency, particularly oral reading skills; and (5) reading comprehension strategies.

The *Study of Teacher Preparation in Early Reading Instruction* responds to a congressional mandate in the Reading First legislation for "a measurement of how well students preparing to enter the teaching profession are prepared to teach the essential components of reading instruction" (No Child Left Behind Act, 2001, Section 1205(c)(8))[3]. The study was commissioned by the National Center for Education Evaluation and Regional Assistance at the U.S. Department of Education's Institute of Education Sciences. The study plan included a survey about teacher education programs and an assessment of pre-service teachers' knowledge about the essential components of early reading instruction.

Primary Research Questions

Two primary research questions guided the *Study of Teacher Preparation in Early Reading Instruction*.

- Research Question 1: To what extent does the content of teacher education programs focus on the essential components of early reading instruction?
- Research Question 2: To what extent are graduating pre-service teachers knowledgeable about the essential components of early reading instruction?

Study Design

The study collected data from a sample of 2,237 pre-service teachers attending a nationally representative sample of 99 institutions that prepare teachers for initial certification using the *Pre-Service Teacher Preparation Program and Knowledge Survey*. The 99 institutions were sampled from all institutions of higher education located in the contiguous United States that graduate at least 50 individuals at the bachelor's, post-bachelor's certificate, or master's level from one or more of the following programs:

[1] See Title I, Part B, Subpart 1, Student Reading Skills Improvement Grants – Reading First, especially Section 1208(3). Downloadable from: http://www.ed.gov/policy/elsec/leg/esea02/pg4.html

[2] National Institute of Child Health and Human Development. (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* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office. Downloadable from: http://www.nationalreadingpanel.org/

[3] See Title I, Part B, Subpart 1, Student Reading Skills Improvement Grants – Reading First, especially Section 1205(c)(8). Downloadable from: http://www.ed.gov/policy/elsec/leg/esea02/pg4.html

- General Education
- Elementary Education and Teaching
- Teacher Education, Multiple Levels
- Early Childhood Education and Teaching
- Reading Teacher Education
- Multi/Interdisciplinary Studies-Other

The 2,237 pre-service teachers were sampled from all students in the sampled institutions who would be eligible to graduate with an elementary teaching education certificate in spring or summer 2007. The pre-service teachers also needed to earn their degree and/or complete their teacher preparation program primarily on-site, not through courses taken mostly online.

The *Pre-Service Teacher Preparation Program and Knowledge Survey* consists of two parts, the Program Survey and the Knowledge Assessment. The Program Survey includes questions that gather background characteristics and items that elicit pre-service teachers' self-reports about the emphasis within their coursework and their exposure through field experiences to the essential components of early reading instruction. The Knowledge Assessment consists of multiple-choice questions about the essential components of reading instruction, especially as they are taught in kindergarten to grade 3 classrooms. The items address relevant research in the teaching and learning of each of the five essential components of early reading instruction.

The *Pre-Service Teacher Preparation Program and Knowledge Survey* was administered to participating pre-service teachers at the participating institutions in the spring and summer of 2007, which represented the end of the samples' teacher training programs. The final distribution of items is shown in Table ES-1.

Table ES-1. Distribution of items in operational Pre-Service Teacher Preparation Program and Knowledge Survey

Component		Total items
Program Survey		
Part 1:	Background characteristic items	22
Part 2:	Exposure/emphasis items	35
	Preparedness items	13
Total: Program Survey		*70*
Knowledge Assessment		
Phonemic Awareness		12
Phonics		8
Fluency		12
Vocabulary		12
Comprehension		12
Total: Knowledge Assessment		*56*
Total items to be completed in 2 hours		***126***

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

A total of 2,237 pre-service teachers attending 99 teacher preparation programs took the *Pre-Service Teacher Preparation Program and Knowledge Survey.* The teacher preparation programs were in 24 states and included both public and private institutions; all had programs that trained teachers to teach in kindergarten to grade six classrooms.

The final sample of 2,237 pre-service teachers included 71.7 percent who were 25 years old or younger. Of the sampled pre-service teachers, 72.2 percent were working toward an undergraduate degree, 63.1 percent had an elementary education major or concentration, and 92.7 percent had no prior teaching certification. Ninety percent of the pre-service teachers in this sample reported that they planned to teach in fall 2007.

Data Analysis and Constructs

In discussing the results from both the Program Survey and the Knowledge Assessment, this report refers to variables related to the five essential components of early reading instruction (phonemic awareness, phonics, fluency, vocabulary, and comprehension) as follows:

- All components: information gathered from items on all five essential components of early reading instruction and analyzed as a single factor
- Alphabetics: information from items on phonemic awareness and phonics
- Fluency: information from oral reading fluency items
- Meaning: information from vocabulary and comprehension items

The study team measured variables related to the emphasis of pre-service teacher programs on the essential components through coursework and the exposure to these components through field experiences using a four-point metric ranging from *none* (0) to *considerable* (3). The overall focus of pre-service teacher programs on the essential components was represented by

the average of reports of emphasis in coursework and exposure through field experiences such as practica and student teaching.

The combination of the five essential components into the three-factor model used in this report is consistent with the *Report of the National Reading Panel* (NICHD, 2000). As conceptualized in the NRP Report and operationalized for this study, alphabetics encompasses phonemic awareness and phonics, which include concepts such as predictors of reading acquisition, phoneme manipulation skills, and letter-sound correspondence. Fluency refers to concepts such as oral fluency, repeated reading, and automaticity of word recognition. Meaning encompasses vocabulary and comprehension; items address concepts such as effective instruction of comprehension skill and strategy, types of vocabulary knowledge found to be essential to reading acquisition, and approaches to vocabulary development.

Study Findings

Research Question 1

The first research question concerns content related to early reading instruction presented in pre-service teacher education programs, specifically, pre-service teachers' perceptions of the extent to which their pre-service teacher programs emphasized the essential components of early reading instruction through coursework and provided opportunities through field experiences such as school-based practica or student teaching that would expose them to such instruction either through observation or actual practice teaching.

Data collected with the Program Survey provide national estimates[4] of pre-service teachers' perceptions of the focus of their training programs on the essential components of early reading instruction.

- On average, pre-service teachers rated the overall focus (based on coursework and field experience data combined) of their training programs as being above "little" but below "moderate," or 1.76 on a zero-to-three scale. On average, pre-service teachers also rated coursework emphasis (1.66) and field experience exposure (1.86) as being above "little" but below "moderate" on a zero-to-three scale.

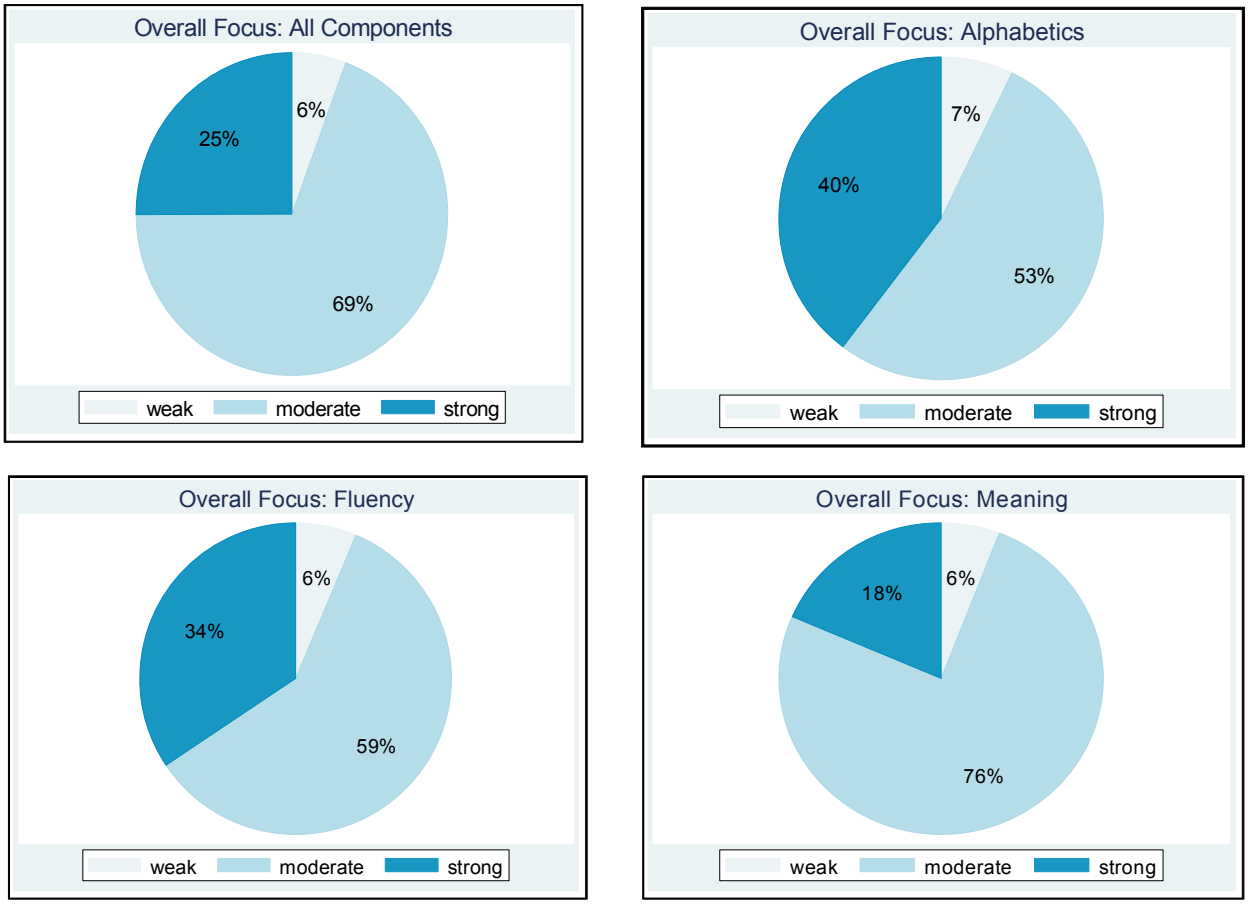
- Sixty-nine percent of pre-service teachers reported a moderate overall programmatic focus (rating greater than 1, but less than or equal to 2 on a zero-to-three scale) on the essential components of early reading instruction, 25 percent reported a strong focus (rating greater than 2 on the scale), and 6 percent reported a weak focus (rating less than or equal to 1 on the scale). See Figure ES-1.

- Pre-service teachers were twice as likely to report a strong focus on the essential components in their field experience than in their coursework (40 percent versus 21 percent across all components) ($p < .001$). See Figures ES-2 and ES-3.

[4] The national estimates provided in the report were derived using data from pre-service teachers at the 99 institutions that agreed to participate in the study.

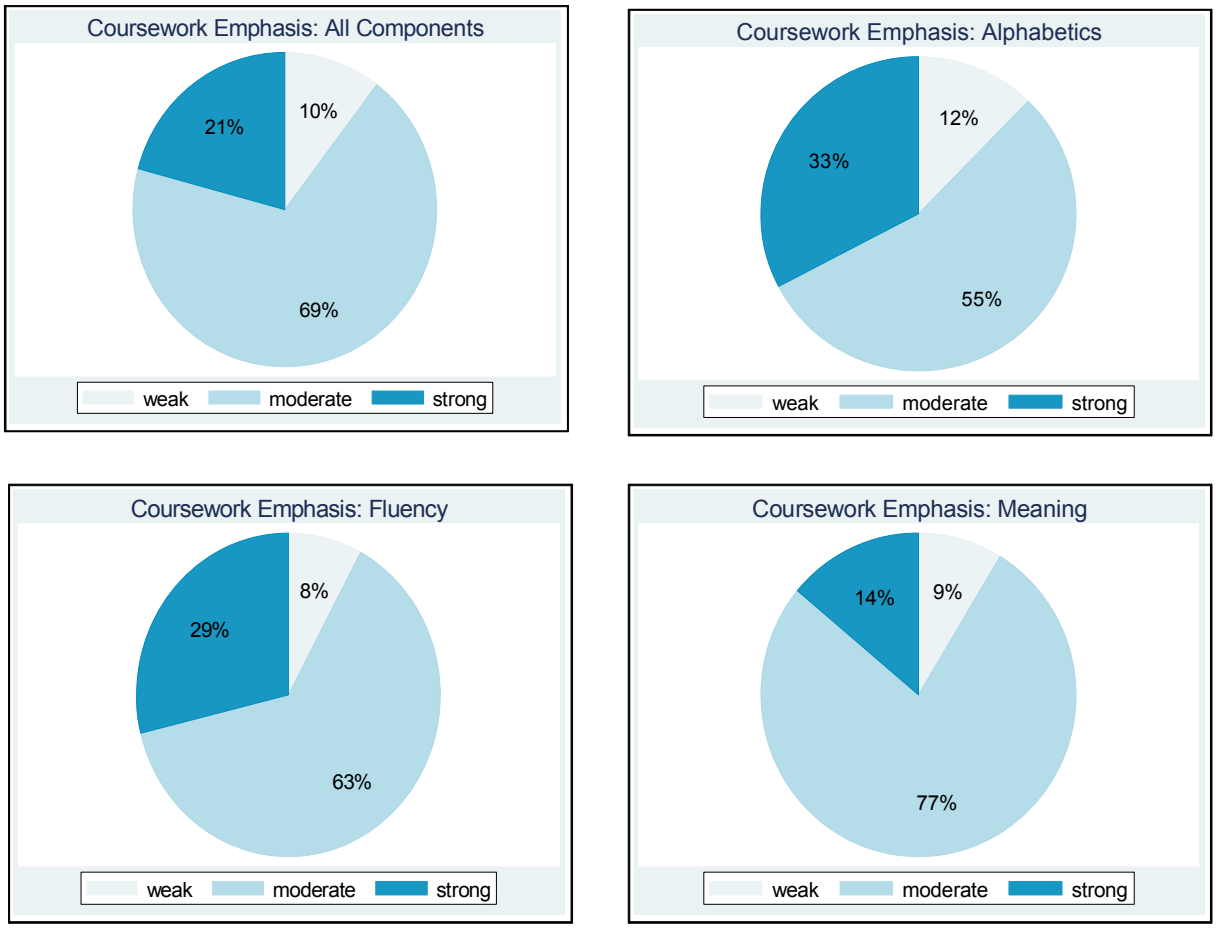
- Pre-service teachers were twice as likely to report a strong overall programmatic focus on alphabetics (40 percent) and fluency (34 percent) than on meaning (18 percent). See Figure ES-1. This was also the case for coursework emphasis (33 percent for alphabetics, 29 percent for fluency, and 14 percent for meaning). See Figure ES-2.

Figure ES-1. Percentage of pre-service teachers reporting weak, moderate, or strong overall program focus on the essential components of early reading instruction, by essential component



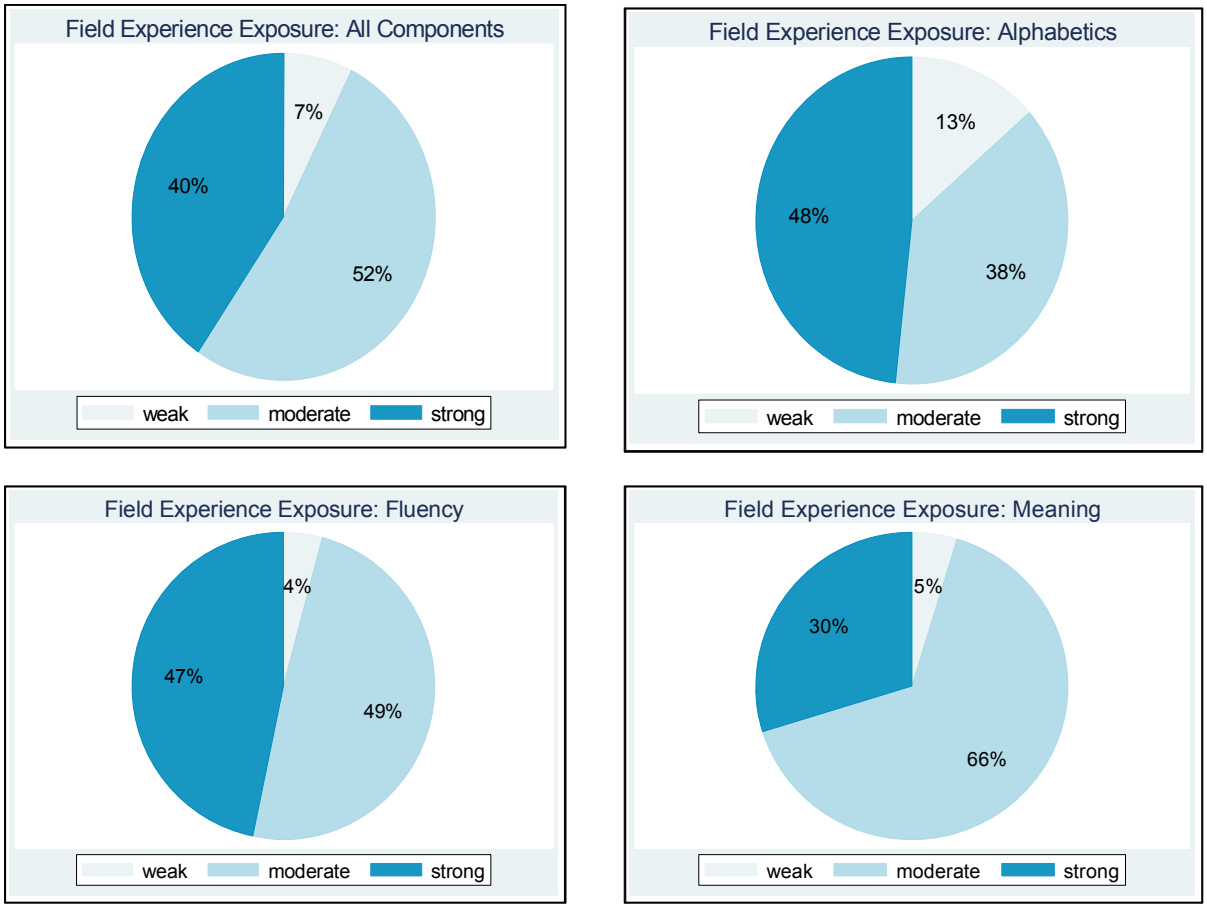
NOTE: Estimates of overall focus were based on coursework and field experience data combined. Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: Study of Teacher Preparation in Early Reading Instruction, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure ES-2. Percentage of pre-service teachers reporting weak, moderate, or strong coursework emphasis on the essential components of early reading instruction, by essential component



NOTE: Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure ES-3. Percentage of pre-service teachers reporting weak, moderate, or strong field experience exposure to the essential components of early reading instruction, by essential component



NOTE: Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Research Question 2

The second research question concerns the knowledge that pre-service teachers demonstrate in responding to multiple-choice knowledge items about the five essential components of early reading instruction. Analysis of data collected with the Knowledge Assessment answered this question. On average, pre-service teachers answered 57 percent of the Knowledge Assessment items correctly. Pre-service teachers were able to answer 53 percent of the alphabetics subscale questions correctly, 61 percent of the fluency subscale questions, and 58 percent of the meaning subscale.

CHAPTER 1:
STUDY OVERVIEW

A component of the *No Child Left Behind* Act (NCLB) (PL 107-110) is its emphasis on the importance of systematic and explicit instruction in early reading using practices that are grounded in scientific research. The Reading First legislation (Title I, Part B, Subpart 1)[5] within NCLB is designed to support state and local education agencies so that they can in turn base their early reading instruction on scientific research and focus on five "essential components" of early reading instruction, as defined by the legislation and informed by the National Reading Panel[6]: (1) phonemic awareness; (2) phonics; (3) vocabulary development; (4) reading fluency, particularly oral reading skills; and (5) reading comprehension strategies.

The *Study of Teacher Preparation in Early Reading Instruction* responds to a Congressional mandate in the Reading First legislation for "a measurement of how well students preparing to enter the teaching profession are prepared to teach the essential components of reading instruction" (No Child Left Behind Act, 2001, Section 1205(c)(8))[7]. The study was commissioned by the National Center for Education Evaluation and Regional Assistance at the U.S. Department of Education's Institute of Education Sciences. The study plan included a survey about teacher training programs and an assessment of pre-service teachers' knowledge about the essential components of early reading instruction.

Data for the study were gathered through administration of a two-part, paper-and-pencil instrument, *The Pre-Service Teacher Preparation Program and Knowledge Survey.* A nationally-representative sample of 2,237 pre-service teachers scheduled to graduate in spring or summer 2007 from 99 colleges and universities that prepare pre-service teachers to teach in elementary classrooms participated in the study. Analysis of data collected with this instrument provides a national estimate of the knowledge of pre-service teachers about the five essential components of reading instruction in aggregate. For analytic and reporting purposes, the study team collapsed the five components to three factors: alphabetics (phonemic awareness and phonics), fluency, and meaning (vocabulary and comprehension). (See Appendices G and I for a discussion of the process resulting in the use of a three-factor model for reporting.) Additional analyses provide answers to both primary and secondary research questions.

---

[5] See Title I, Part B, Subpart 1, Student Reading Skills Improvement Grants – Reading First, especially Section 1208(3). Downloadable from: http://www.ed.gov/policy/elsec/leg/esea02/pg4.html

[6] National Institute of Child Health and Human Development. (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* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office. Downloadable from: http://www.nationalreadingpanel.org/

[7] See Title I, Part B, Subpart 1, Student Reading Skills Improvement Grants – Reading First, especially Section 1205(c)(8). Downloadable from: http://www.ed.gov/policy/elsec/leg/esea02/pg4.html

Primary and Secondary Research Questions

Two primary research questions have guided the study.

- Research Question 1: To what extent does the content of teacher education programs focus on the essential components of early reading instruction?
- Research Question 2: To what extent are graduating pre-service teachers knowledgeable about the essential components of early reading instruction?

The study team also investigated relationships across the two questions and within subgroups of pre-service teachers. The following secondary research questions have guided these analyses.

- Research Question 3a: Which characteristics of teacher training institutions and programs are associated with their focus on the essential components of early reading instruction?
- Research Question 3b: To what extent are teacher training programs' focus on the essential components of early reading instruction associated with pre-service teachers' knowledge about these components?
- Research Question 3c: To what extent is pre-service teachers' knowledge about the essential components of early reading instruction related to these pre-service teachers' feelings of preparedness to teach various aspects of beginning reading?

The secondary research questions supplement the primary research questions through an investigation of potential relationships among the contexts in which pre-service teacher preparation occurs and the content offered to pre-service teachers; the content of the programs and pre-service teachers' knowledge as measured by the Knowledge Assessment; and knowledge about the essential components of early reading and pre-service teachers' feelings of preparedness to teach these components. Development of the questions was guided by research on pre-service teacher education, especially in early reading instruction and learning. The underlying premises of the questions are that pre-service teachers' professional knowledge accrues from coursework and clinical or field experiences[8] and that knowledge about the five essential components of early reading instruction is essential for high-quality teaching.[9] Further, correlational research on the characteristics of institutions housing teacher education programs has suggested that the level of preparedness of pre-service teachers differs depending on whether the institution offers only master's degrees or also includes doctoral programs.[10] Finally, positive relationships have been found between teachers' feelings of preparedness to teach a

[8] Darling-Hammond, L. (1997). *Doing what matters most: Investing in teacher quality.* Kutztown, PA: National Commission on Teaching and America's Future.

[9] National Institute of Child Health and Human Development, 2000, especially Chapter 5; Snow, C.E., Burns, M.S., & Griffin, P. (1997). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press; Snow, C.E., Griffin, P., & Burns, M.S. (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world.* New York: John Wiley & Sons.

[10] See especially Levine, A. (2006). *Educating school teachers.* Washington, DC: The Education Schools Project; also Gittomer, D., Latham, A., Ziomek, A, & Ziomeck, R. (1999). *The academic quality of prospective teachers: The impact of admissions and licensure testing.* Princeton, NJ: Educational Testing Service.

particular subject and the quality of their instruction.[11] This study investigates not the link between feelings of preparedness and instruction but these feelings and achievement.

Study Design

The study has consisted of four distinct phases: developing the sample frame and data collection tools; recruiting institutions and pre-service teachers; collecting data; and analyzing data to answer the primary and secondary research questions. Additional details regarding the study sample are discussed in Chapter 2, and details about data collection are discussed in Chapter 3.

Constructing the frame from which teacher training institutions were recruited was the first activity in conducting the study. Using data from the Integrated Postsecondary Education Data System (IPEDS),[12] the study team selected a random sample of 120 schools to be representative of the institutions that prepare elementary teacher candidates. The study team deemed this size adequate to estimate national averages with reasonable precision and to yield a final sample of 100 participating institutions, assuming an 85 percent response rate. (Ultimately, the initial sample consisted of 119 institutions.) Within participating institutions, the study team identified graduating pre-service teachers, drew a random sample of these pre-service teachers, and recruited individuals to sit for the questionnaire.

An additional preliminary step was to develop the data-gathering tool, *The Pre-Service Teacher Preparation Program and Knowledge Survey.*[13] This is a two-part, printed booklet designed to take approximately two hours to complete. The first section, termed the *Pre-Service Teacher Program Survey* (hereafter referred to as the "Program Survey") consists of 22 questions about pre-service teacher background characteristics. An additional 35 items elicit pre-service teachers' self-reports about the focus of their programs on the essential components of early reading instruction, in terms of both their perception of the emphasis within their courses and their exposure to the essential components during field experiences and student teaching. The survey also asked pre-service teachers about their feelings of preparedness to teach these instructional components.

The second section, *The Pre-Service Teacher Knowledge Assessment* (hereafter referred to as the "Knowledge Assessment"), consists of 56 multiple-choice questions about the essential components of reading instruction, especially as they are taught in kindergarten to grade three classrooms. The Program Survey and the Knowledge Assessment were administered to pre-service teachers together; the full data collection instrument is hereafter referred to as "the questionnaire."

---

[11] See U.S. Department of Education, National Center for Education Statistics. *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers,* NCES 1999-080, by Laurie Lewis, Basnat Parsad, Nancy Carey, Nicole Bartfai, Elizabeth Farris, & Becky Smerdon. Bernie Green, project officer. Washington, DC: Author.

[12] The Integrated Post Secondary Education Data System (IPEDS) is a searchable database of information about postsecondary education and is maintained by the National Center for Education Statistics. It includes information about enrollments, program completions, graduation rates, faculty, staff, finances, institutional prices, and student financial aid about major postsecondary institutions. It is accessible at http://nces.ed/gov.ipeds/.

[13] OMB 1850-0817

The second major study activity was recruitment of institutions that prepare pre-service teachers. Institutions that declined to participate were replaced until an initial sample of 100 institutions in 24 states was identified (ultimately, the sample included 99 institutions). From each institution in the sample, 30 pre-service teachers who would graduate in spring or summer 2007 were randomly selected for recruitment into the study.

Project staff proctored sessions at which the questionnaire was administered to pre-service teachers who agreed to participate in the study. By August 2007, 2,237 pre-service teachers from 99 institutions in 24 states had completed the questionnaire.

The study team has analyzed data collected from these pre-service teachers at participating institutions to answer the primary and secondary research questions. Research Question 1 concerns the pre-service teachers' perception of their teacher training programs' content related to early reading instruction. Specifically, it concerns the extent to which pre-service teacher training programs emphasize the five essential components of early reading instruction through formal pedagogy courses and provide opportunities through field experiences such as school-based practica or student teaching that will expose pre-service teachers to such instruction either through observation or actual practice teaching.[14] The report presents data on pre-service teachers' perceptions of their coursework *emphasis* on the essential components, their *exposure* during their field experience to the essential components, and their programs' *overall focus* on the essential components (the average of coursework emphasis and field experience exposure). Data from the Program Survey have been used to answer this question.

Research Question 2 concerns the knowledge that pre-service teachers demonstrate in responding to multiple-choice assessment items about the five essential components of early reading instruction. Data from the Knowledge Assessment have been used to answer this question.

For both research questions, the report presents findings for all the essential components combined, as well as alphabetics (the combination of phonemic awareness and phonics), fluency, and meaning (the combination of vocabulary and comprehension).

The secondary research questions have required analysis across various components of the Program Survey and the Knowledge Assessment. Research Question 3a concerns the relations across several factors measured by the questionnaire: the characteristics of the teacher training institutions and the programs they offer, and the focus on the essential components of early reading instruction reported by pre-service teachers at these institutions. Data from the Program Survey have been used to answer this question.

Research Question 3b addresses the relationship between pre-service teachers' perception of their programs' focus on the essential components of early reading instruction and their scores on

[14] According to *The Secretary's Fifth Annual Report on Teacher Quality*, there are 110 initial categories of initial state teaching certificates or licenses offered nationwide. Of these, 103 require pre-service teachers to complete a supervised practice teaching experience, and 101 require teachers to take specific pedagogy courses. See http://www/ed/gov/about/reports/annual.teachprep/2006-title2report.pdf, p. 27, Figure 3.1 and Appendix A2.

the Knowledge Assessment. Data from both components of the questionnaire have informed the answer to this question.

Finally, Research Question 3c concerns the relationship between pre-service teachers' feelings of preparedness to teach and their scores on the Knowledge Assessment. Again, the answer derives from analyses across the questionnaire, using survey items on pre-service teachers' perceptions of preparedness and the Knowledge Assessment data.

Organization of This Report

This report consists of the executive summary and this overview chapter, followed by three additional chapters that provide more details about the study. Chapter 2 provides information on the selection of the sample for the study. Chapter 3 discusses the study's data collection measures and procedures. Chapter 4 provides results that address the primary and secondary research questions. Appendices provide more technical information about the components of the study.

CHAPTER 2:
THE STUDY SAMPLE

The *Study of Teacher Preparation in Early Reading Instruction* investigated the extent to which pre-service teacher training programs include content on the five essential components of early reading instruction and pre-service teachers' knowledge of these essential components. The target student population for this study was pre-service teachers preparing to be elementary teachers in on-site (not primarily online) programs in colleges and universities. These pre-service teachers would graduate in spring or summer 2007, qualified to seek initial state certification or licensure so that they could enter an elementary classroom at the beginning of the 2007–2008 school year. To represent the target population adequately, the study team collected data from a nationally representative sample of 2,237 pre-service teachers attending 99 colleges and universities that prepare pre-service teachers. These institutions offered on-site elementary school teacher preparation programs that prepared students to apply for initial state certification or licensure. This chapter provides information about the sampling design, the sample selection, and the weighting procedures.

Three-Stage Sample Design

As a preparation for drawing the sample, the study team reviewed the teacher certification websites of states in the contiguous United States to develop an initial list of institutions in these states with on-site pre-service teacher preparation programs that might be eligible for the study. The review also resulted in a list of the terminology used to designate programs that prepare pre-service teachers. Next, the list of programs was cross-checked with the degree programs in the Integrated Postsecondary Education Data System (IPEDS; 2004), as defined by the Classification of Instructional Programs (CIP) published by National Center for Education Statistics (NCES). All programs listed within Education and Multi/Interdisciplinary Studies[15] were considered, but the final list included only programs whose descriptions indicated an emphasis on preparing teachers for work in kindergarten to grade six classrooms. Even though some of the programs prepare teachers for broader grade/age ranges (e.g., certification to teach pre-kindergarten to grade eight), these programs typically offer courses that included preparation in early reading instruction. The final list of identified program classifications is as follows:[16]

- General Education
- Elementary Education and Teaching
- Teacher Education, Multiple Levels
- Early Childhood Education and Teaching

[15] A web search indicated that Texas trained its teachers within interdisciplinary programs rather than in a school of education.

[16] The number of degrees conferred in these fields is greater than the number of elementary education teachers entering the workforce, since some of these students would be trained to become middle or high-school teachers. However, the focus of this level of sampling was on identifying institutions with elementary education teacher training. Institutions that offer middle or high school training may have an elementary education program as well. Institutions that did not offer an elementary education program were treated as ineligible.

- Reading Teacher Education
- Multi/Interdisciplinary Studies-Other

Analysis of 2004 IPEDS data revealed that the 50 largest programs in the contiguous United States accounted for over 25 percent of all reported elementary education graduates, and that 34 percent of institutions produced fewer than 25 graduates per year. In order for the study to better represent institutions producing the largest numbers of pre-service teachers, the study team restricted the institutional frame to colleges and universities that graduate at least 50 pre-service teachers at the bachelor's, post-bachelor's certificate, or master's level from one or more of the programs described above.[17]

To gain further information about potentially eligible institutions, the study team examined institution websites and state department of education websites to obtain program and state-specific certification requirements such as certification test(s), mandatory coursework, and minimum hours spent student teaching. Certification requirements listed in institution and state websites were generally complete; thus, information gained from these reviews was helpful in recruitment efforts, as few ineligible institutions were contacted.

The actual sample selection for the study followed a three-stage design, in which the study team first selected geographic areas (states), followed by higher-education institutions, and then pre-service teachers at participating institutions.

State Selection

Programs may be influenced by factors at the state level, such as state legislature initiatives, state funding sources, and student demographic characteristics. Every state in the contiguous United States has at least one institution that graduates 50 or more elementary education teachers a year. Thus, to control the cost and time for administering the questionnaire, the study team selected a sample of states as the first stage of sampling, using implicit stratification procedure. *Implicit* stratification, in contrast to *explicit* stratification, involves sorting the population of sampling units by some characteristic(s) and then selecting the sample from the sorted list, using a fixed sampling interval and a random start. In contrast, *explicit* stratification involves dividing the population of sampling units into strata and selecting a separate sample per stratum.[18] The implicit stratification variable used in the sampling of states in this study was the National Assessment of Educational Progress (NAEP) geographic cluster. Specifically, the frame of 48 continental states was sorted according to the NAEP geographic clusters and a measure of size, using a serpentine order (high to low, then low to high). The measure of size was the number of pre-service teachers expected to graduate from the programs listed from all eligible institutions in that state and who were eligible to obtain certification to teach elementary school classes. As a result, 24 of the 48 continental states were selected, with probability proportional to size (PPS),

[17] Based on the 2004 IPEDS database, 51 percent of the total number of colleges and universities that graduate any pre-service teachers met this criterion.

[18] For more information on explicit and implicit stratification, see Kish, L. (1995). *Survey sampling*. New York: John Wiley & Sons.

using a stratified systematic random sampling procedure.[19] This number represents 50 percent sampling at the state level, a compromise between limiting travel costs for data collection and having a sufficient sample size to investigate program focus at the state level.

Institution Selection

The sampling design for the second stage was also a stratified systematic random sample, with sampling probabilities proportionate to size (PPS). The measure of size at this stage was the number of graduating pre-service teachers preparing to seek initial certification to teach elementary school at an eligible institution in any of the programs listed above. The frame of institutions was sorted according to the following stratification variables: state, school type (public vs. private), and minority enrollment (high vs. low), as well as a measure of size, using a serpentine order (high to low, then low to high). This sorting ensured an adequate range of key characteristics across all institutions selected in the sample. The sample was systematically selected from the ordered frame, with the sampling interval calculated by dividing the cumulative measure of size by the sample size. In order to meet our recruitment goal of at least 2,000 pre-service teachers, we made the following assumptions: the pre-service teachers would participate at an 85 percent rate at each school, and schools would participate at an 85 percent rate. In order to have a 95 percent probability of having at least 100 participating institutions and 2,000 pre-service teachers, we determined that we would need to sample 119 schools and 25 pre-service teachers per school.

For each sampled institution, the next two institutions immediately following it in the sampling frame were designated as its replacement institutions. The use of implicit stratification variables, and the subsequent ordering of the institution sampling frame by size, ensured that replacements for any sampled institution had similar characteristics. When a sampled institution was the last institution listed, then the two institutions immediately above it were designated as its replacement institutions. If a sampled institution was the next to the last institution listed, then the institutions immediately above and below it were designated as its replacement institutions. However, a sampled institution cannot be designated as a replacement institution, and a replacement institution cannot be assigned to substitute for more than one sampled institution. The resulting sample included 119 institutions from the sampled states; they were selected based on their probability proportionate to size.[20]

Pre-Service Teacher Selection

After institutions were successfully recruited to participate in the study, project staff initiated the third and final stage of selection: identification of a sample of eligible pre-service teachers. Pre-service teachers were considered eligible if they were preparing to be elementary teachers, would graduate in spring or summer 2007, and would be qualified to seek certification and to enter a classroom at the beginning of the 2007–2008 school year. Pre-service teachers also needed to earn their degree and/or complete the teacher preparation program primarily on-site, not through

[19] To reduce the variability of the estimates, nine states with the largest measures of size were selected "with certainty." This means they were sure to be selected and would represent only themselves (i.e., had a selection probability of one and a sampling weight of one).

[20] Institutions that declined to participate were replaced by a similar institution.

courses taken online. Rosters of eligible pre-service teachers were obtained from each institution that agreed to participate in the study; the rosters were sorted alphabetically; and from the rosters, a stratified random sample of 30 pre-service teachers was selected for recruitment. The alphabet as applied to the last and first name of students was used as an implicit stratification variable because there was no reason to assume a correlation between one's last and first name and one's ability. When an institution had 30 or fewer eligible pre-service teachers, the sample consisted of all pre-service teachers.

Computation of Final Weights

The sample of pre-service teachers was collected at three stages with a differential probability of sample selection within each stage. Twenty-four states were selected proportionate to size, 119 institutions were then selected from these states based on their probability proportionate to size, and up to 30 pre-service teachers were randomly selected from each institution that agreed to participate regardless of size. Sample weights are necessary for all statistical analyses. Weights were applied to the pre-service teacher data to account for each institution's size; that is, results for pre-service teachers graduating from larger teacher training programs were weighted more heavily than results from pre-service teachers graduating from smaller programs.

Sampling weights for this study were calculated using the following formula:
$W_{ijk} = [1/(P_{ijk} * P_{jk} * P_k)]$,

where

- W is weight and P is probability,
- i is an individual, j is an institution, and k is a state.

The formulas below specify each probability shown in the equation above. Notice that stage 1 (i.e., selection of states) and stage 2 (i.e., selection of institutions) were based on probability proportionate to size; therefore, the probability calculation was based on the number of graduating pre-service teachers in states or in institutions. In this way, states or institutions with high numbers of graduating pre-service teachers were given due recognition that they had a larger probability of selection.

Stage 1: State levels

For the nine certainty states (see footnote 20), $P_k = 1$. For the remaining 15 sampled states:

$P_k = (N\_State * MOS_k / SUM\ of\ MOS)$,

where

- N_State is the number of states (= 15),
- MOS, measure of size, is the number of pre-service teachers graduating from eligible institutions.

Stage 2: Institution levels

$P_{jk} = (N\_Institution*MOS_{jk}/SUM\ of\ MOS_k)$,

where

- N_Institution is the number of institutions (= 119),
- MOS, measure of size, is the number of pre-service teachers graduating from eligible institutions. Sum of MOS_k is across the 24 selected states.

Stage 3: Pre-service teacher levels

$P_{ijk} = (N\_PST / N_{jk})$,

Where

- N_PST is the targeted number of pre-service teachers from each institution (= 30),
- N_{jk} is the number of pre-service teachers in relevant programs in each eligible institution. If N_{jk} is equal to or less than 30, the weight will be 1.

To account for nonresponse during data collection, nonresponse adjustment factors were applied to the formulas in stages 2 and 3. Specifically, the nonresponse adjustment factor for stage 2 was 0.83, which was equal to the number of institutions that responded to the survey (n = 99) divided by the number of institutions in the targeted sample (n = 119). The nonresponse adjustment factor for stage 3 was the number of pre-service teachers that responded to the survey in each selected institution divided by the number of pre-service teachers in the targeted sample (n = 30). Therefore, the stage 3 nonresponse adjustment factor varied for pre-service teachers selected from different institutions.

Exceptions

For three institutions, it was necessary to draw the pre-service teacher samples twice; hence, special weighting procedures were used. For all three institutions, a pre-service teacher frame was provided, the sample of 30 pre-service teachers was drawn, the pre-service teachers were recruited, and data collection was initiated. In these institutions, initial phases of recruitment or data collection determined that the list of graduating pre-service teachers provided to the study team included errors (i.e., some pre-service teachers did not meet the eligibility criteria). The study team subsequently obtained a corrected list for these institutions and could select the correct sample of graduating pre-service teachers. However, the number of pre-service teachers selected was adjusted to account for the number of eligible pre-service teachers from whom data had been collected before the errors were discovered. This process resulted in two samples being drawn from each of these schools with pre-service teachers in each sample having different probabilities of selection. To account for this difference, a composite weight was established from the two rounds of sample selection. The weight was a unique function of the sampling process required at each of these three institutions. The composite weight was used in the final analyses.

Some of the procedures used in the sample selection were assumed to bear no impact on probability of selection; therefore, they did not enter the algorithms for sample weight calculation. For example, no consideration was made to account for the fact that data frames were used to ensure diversity of the sample (e.g., geographic regions, sectors, minority concentration). Also, stage 2 used replacement schools when schools that had been originally sampled declined to participate. However, replacement schools were assumed to match the originally identified institution because of how the frames were ordered prior to selection.

Recruitment of Study Sample

Recruitment of institutions and pre-service teachers and administration of the questionnaire lasted from September 2006 to August 2007. In total, 99 institutions in 24 states agreed to participate in the study; 2,237 pre-service teachers at these institutions sat for the questionnaire administration. These students were pre-service teachers who would be eligible to graduate in spring or summer 2007 and to seek state certification as elementary school teachers.

Of the 119 original sampled institutions,[21] 63 institutions agreed to participate, 4 institutions were determined to be ineligible, and 52 institutions declined to participate for various reasons. This original, or initial, sample yielded a response rate of 55 percent unweighted. The weighted institution response rate before replacement was 58 percent. An additional 60 replacement institutions were contacted as substitutes for the institutions that declined to participate. Of the replacement institutions, 36 institutions agreed to participate and 24 institutions declined. The overall sample yielded a total of 99 institutions for a response rate of 86 percent unweighted and 94 percent weighted after replacement. Appendix D presents detailed information on institution recruiting, including the reasons cited by institutions that declined to participate.

From the 99 participating institutions, a total of 2,892 pre-service teachers were sampled for the study. Of these, 87 were deemed ineligible and removed from the sample. Of the remaining 2,805 pre-service teachers, 568 did not respond because of refusal or other reasons. For example, 49 of the students who registered to take the survey did not show up for an administration session at their Institutions. A total of 2,237 pre-service teachers completed the survey with a response rate of 80 percent unweighted and 78 percent weighted. Appendix E provides demographic and other pre-service teacher background characteristics from the Program Survey. Table 2-1 summarizes key characteristics of the pre-service teacher sample.

[21] Although 119 institutions were initially sampled, the study team contacted only 118. Prior to institution recruitment, preliminary reconnaissance indicated that one of the 119 institutions was ineligible because coursework at the institution was offered primarily online. This institution was not replaced.

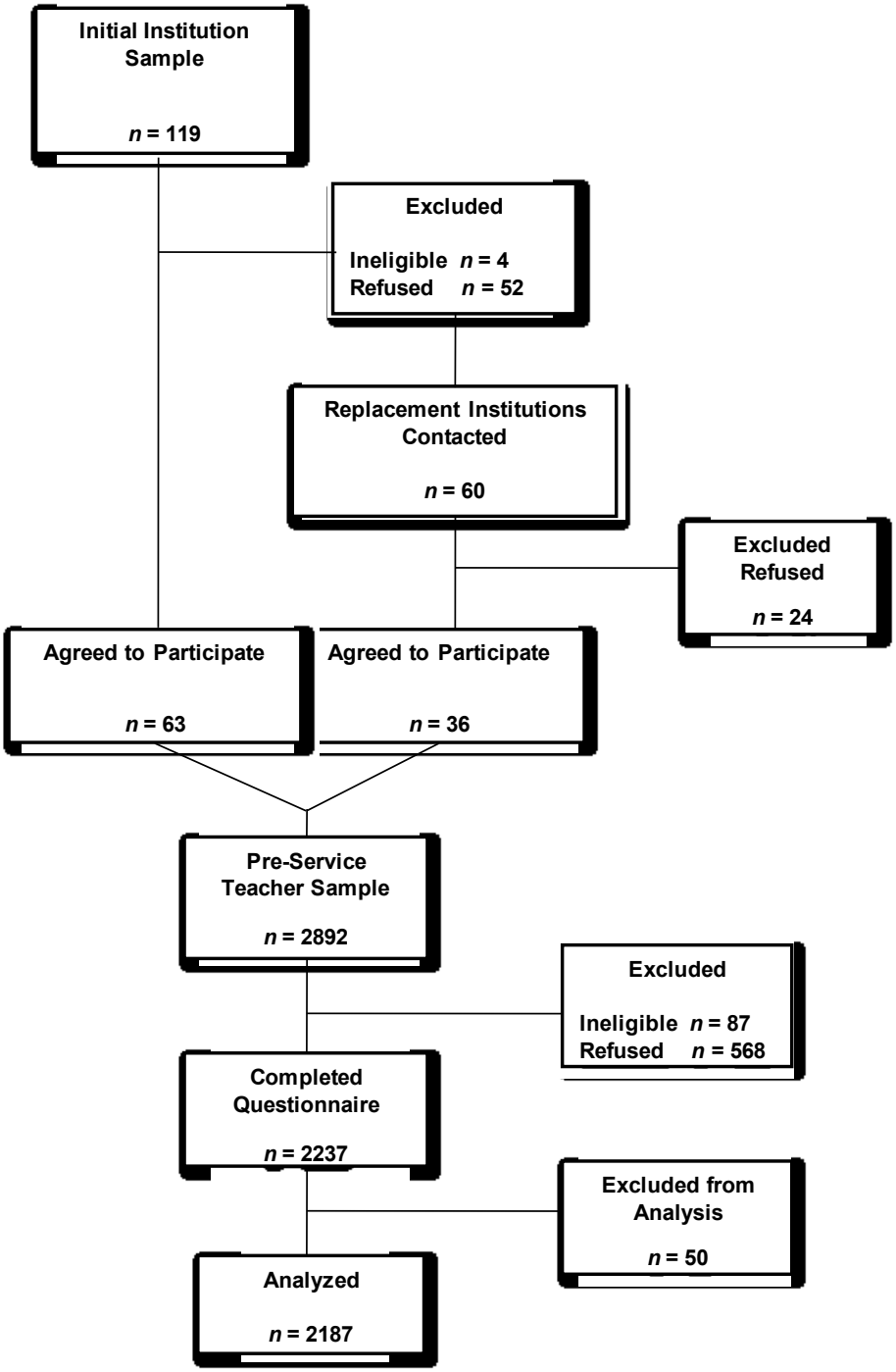
Table 2-1. Characteristics of final pre-service teacher sample

Pre-service teacher characteristic	Number	Percent
Female	2,011	92.0
25 years old or younger	1,568	71.7
White	1,823	83.4
Working toward an undergraduate degree	1580	72.2
Had elementary education major or concentration	1379	63.1
Had taken four or more courses with field experience (regardless of course focus)	1,604	73.3
Were completing student teaching in the spring 2007 semester	1,774	81.1
Had another college degree	803	36.7
Expected to graduate in spring or summer 2007	2,184	99.9
Had no prior teacher certification	2,028	92.7
Indicated they planned to teach in fall 2007	1,976	90.4
Total	*2,178*	*100*

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure 2.1 provides information on the flow of institutions and pre-service teachers from sampling through analysis. As is explained in Appendix F, although 2,237 pre-service teachers completed the *Pre-Service Teacher Preparation Program and Knowledge Survey*, data cleaning and initial psychometric analyses reduced the number of usable booklets to 2,187.

Figure 2-1. Flowchart of institution and student recruitment



SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

CHAPTER 3:
DATA COLLECTION

This chapter describes the steps taken to conceptualize, develop, and pilot test the *Pre-Service Teacher Preparation Program and Knowledge Survey* that was used to gather data to address the primary and secondary research questions presented in Chapter 1. This measure consists of two major components: the *Pre-Service Teacher Program Survey* and the *Pre-Service Teacher Knowledge Assessment.* This chapter also provides information on data collection procedures.

The Program Survey

The Program Survey gathered background and demographic characteristics about the pre-service teachers participating in the study and their perceptions of their teacher training programs. The background section consists of 22 multipart items that ask about pre-service teachers' gender, age, race/ethnicity, degree programs, type of degree and areas of study, and the number of reading courses required by their respective programs. Characteristics about pre-service teachers' academic history, such as SAT or ACT scores and cumulative grade point (GPA) average overall and in education courses, were also collected.

Items in the second section of the Program Survey ask about pre-service teachers' perceptions of the extent to which their coursework emphasized the five essential components of early reading instruction and their field experiences exposed them to such instruction in early reading. Items also ask about pre-service teachers' perceptions of their own preparedness[22] to teach the five essential components of beginning reading instruction. The study team fully recognizes that the survey data are self-reports of pre-service teachers' background characteristics, perceptions of their training programs, field experiences, and feelings of preparedness. Confirmation of background characteristics or of actual content covered in courses was beyond the scope of the work in this study. Appendix A presents the Program Survey.

In developing the second section of the Program Survey, the study team first conducted a literature review and identified four measures that had been developed to investigate teachers' thinking about their instruction. Three of the previous measures had been designed to differentiate pre- or in-service teachers according to their theoretical orientation toward instruction (constructivist/whole language vs. explicit code instruction) (Bos, Mather, Dickson, Podhajski, & Chard, 2002; DeFord, 1985; Evans, Fox, Cremaso, & McKinnon, 2004; Mather, Bos, & Babur, 2001). The fourth instrument (Baumann, Hoffman, Duffy, & Ro, 2000) was designed to investigate the general status of in-service teachers' reading instruction, including their philosophical orientation toward reading instructional methods and materials; its goal was to generate a better understanding of general instructional practices. These instruments proved to

[22] Although teachers' feelings of preparedness to teach aspects of the essential components of early reading instruction are not discussed in the *NRP Report*, the study team decided to include questions about this construct because it has been found to be a potential indicator of in-service teacher quality. See: U.S. Department of Education, National Center for Education Statistics. *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers,* NCES 1999-080, by Laurie Lewis, Basnat Parsad, Nancy Carey, Nicole Bartfai, Elizabeth Farris, & Becky Smerdon. Bernie Green, project officer. Washington, DC: Author.

be of minimal use because their focus differed from that of the study. The second step in the development process was to review recent scientific research on beginning reading, most specifically the *Report of the National Reading Panel* (NICHD, 2000), and work published by the National Research Council (Snow, Burns, & Griffin, 1998; Snow, Griffin, & Burns, 2005).

One set of yes or no items and three sets of forced-choice items were developed to measure pre-service teachers' perceptions of their programs' focus on the essential components and their preparedness to teach these components. The set of yes or no items consists of five questions that ask pre-service teachers to think about their coursework and field experiences in general as a total experience. Their task is to check "yes" or "no" to indicate whether or not they have "learned about what students need to know and be able to do" to demonstrate skills related to each of the five essential components of early reading instruction, such as associating "letters and the sounds they make to identify words" or understanding "what they read." The texts of the item's subparts were constructed by excerpting brief definitions of the five essential components of reading from the Report of the National Reading Panel.[23] To illustrate, the subpart referring to phonics is stated as "Associate letters and the sounds they make" to represent phonics. Thus, the item presents shortened definitions of the five essential components of early reading instruction as they are expressed in the *NRP Report*.

Two sets of identical items target perceptions of two aspects of their program—emphasis in coursework and exposure through field experience—by asking about specific research-based strategies employed in early reading instruction. The first set asks pre-service teachers to think about all their coursework in reading and literacy, and the second set asks them to think about their observations and own activities as part of field experiences and student teaching, focusing on aspects related to reading and literacy. This contrast allows for an investigation of perceived differences in emphasis and exposure provided by coursework versus field experience. Respondents rate each of these items on a 4-point scale (None, Little, Moderate, Considerable) to indicate their perceptions of exposure to or emphasis on these strategies. These items yielded national estimates of the overall focus of pre-service teacher training programs.

To obtain pre-service teachers' feelings of preparedness to teach reading, thirteen of the seventeen items from the coursework/field experience item sets are repeated. Here, pre-service teachers rate how prepared they felt to teach each concept or strategy. By repeating the items, it was possible to link pre-service teachers' reported programmatic focus on these concepts and strategies directly with their feelings of preparedness. All items in this group were given a 4-point scale (Not at all prepared, Somewhat prepared, Mostly prepared, and Definitely prepared).

Potential items for the Program Survey were subjected to extensive review and pilot testing with pre-service teachers similar to those who would be in the final sample. Two levels of pilot testing were employed: (1) focus groups and (2) cognitive laboratory interviews. The results of the pilot tests informed wording and format of the final instrument. More details about the Program Survey pilot testing are in Appendix C.

[23] National Institute of Child Health and Human Development, *op cit.*

Based on the pilot test data, the study team estimated that respondents would take approximately one hour to complete the final Program Survey. Table 3-1 shows the distribution of items in the

Based on the pilot test data, the study team estimated that respondents would take approximately one hour to complete the final Program Survey. Table 3-1 shows the distribution of items in the Program Survey and internal consistency reliability estimates of the clusters of items. Appendix B presents information about the variables measured in the questionnaire.

Table 3-1. Program survey item totals and internal consistency, by section

		Total items	Internal consistency
Part 1:	Background characteristics	22	†
Part 2:	Overall focus of programs	1	†
	Emphasis in coursework	17	.887
	Exposure through field experience	17	.878
	Feelings of preparedness	13	.880
Total		*70*	†

†Not applicable; internal consistencies are only available for the areas noted above.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

The Knowledge Assessment

The purpose of the Knowledge Assessment is to measure pre-service teachers' knowledge about the five essential components of early reading instruction. The foundation of the Knowledge Assessment is the *Survey of Teacher Knowledge of Reading Instruction in the Primary Grades* (Costigan, Baker, Day, Smith, & Salinger, 2005), which AIR had developed under a contract from the National Center for Education Statistics funded through the Educational Statistics Service Institute. Members of the study's Technical Working Group reviewed the teacher Knowledge Assessment items and recommended modifications and pilot testing with pre-service teachers.

Review and Modification of the Existing Assessment

The primary goals for revision of the *Survey of Teacher Knowledge of Reading Instruction in the Primary Grades* were to ensure that the resulting modified version would provide adequate and accurate coverage of the five essential components of early reading instruction and that the items were targeted correctly for pre-service teachers as opposed to in-service teachers.

Specifications for items within the *Survey of Teacher Knowledge of Reading Instruction in the Primary Grades* item bank extended beyond the five essential components of early reading instruction.[24] Removing irrelevant items from the existing bank yielded a first set of items about aspects of instruction that were systematically compared against the recommendations for scientifically based instruction included in the *NRP Report* (NICHD, 2000). Items that did not

[24] Items had been developed to assess components of a model that included student motivation and what was termed students' "physiological readiness" to participate in and benefit from instruction (see Costigan et al., 2005).

reflect the recommendations of the NRP were removed from consideration, leaving a set of items that could potentially be modified for use in the study. The resulting set of items was not large or comprehensive enough to meet the requirements of the Knowledge Assessment specifications, so new items were written to supplement those that could be revised. The writing/review/revision process included steps to ensure that the information in the items was aligned to the NRP recommendations and to modify the difficulty level or context for use with pre-service teachers.[25]

Pilot Testing of the Knowledge Assessment

A set of 106 items was assembled into two pilot test forms[26] for administration to a convenience sample of 142 pre-service teachers nearing completion of their programs at one university in a western state and several universities in the mid-Atlantic area. Additional information related to the Knowledge Assessment pilot test is provided in Appendix C.

Pilot test results were used to assemble a final form of the Knowledge Assessment consisting of 56 multiple-choice items that could be administered within approximately one hour. Table 3-2 shows the distribution of items across the five essential components of early reading instruction, with statistical results of the pilot test. Internal reliability ranged from 0.41 for phonics items to 0.59 for fluency items. To increase subscale reliabilities, the phonemic awareness and phonics subscales were combined to form the alphabetics subscale, with an internal consistency of .60, and the vocabulary and comprehension subscales were combined to form the meaning subscale, with an internal consistency of .67. Combining all items into a single scale resulted in a reliability of .78. Knowledge Assessment subscales reliabilities at these levels were a concern, in that they may have limited the ability to find significant relations among program focus measures and scores on the Knowledge Assessment subscales.

[25] Data from a pilot test of the *Survey of Teacher Knowledge of Reading Instruction in the Primary Grades* with a national sample of in-service teachers was used to gauge difficulty level (see Costigan et al., 2005).

[26] For the pilot test, two forms of the assessment, Form A and Form B, were created to guard against order and practice effects. Items were placed in a random order except when items pertained to a specific passage. When a group of items was related to a single passage, those items were clustered together on the assessment forms. Form B was created by reversing the order of the items on Form A; thus, the first items on Form A are at the end of Form B.

Table 3-2. Items in the Knowledge Assessment, with statistics from pilot test

Essential component of reading	Final # of items	Coefficient alpha	Mean % correct (final set)	SD % correct (final set)
Phonemic Awareness	12	.52	66.7	.170
Phonics	8	.41	55.4	.202
Fluency	12	.59	72.6	.166
Vocabulary	12	.58	72.3	.177
Comprehension	12	.44	60.1	.181
Alphabetics (Phonemic Awareness + Phonics)	20	.60	62.2	.150
Meaning (Vocabulary + Comprehension)	24	.67	66.3	.146
Total (all items)	*56*	*.78*	*66.1*	*.120*

NOTE: N of teachers = 142.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Additional analyses (detailed in Appendix C) examining the relations among scores on the Knowledge Assessment and other indicators of academic achievement suggested that scores were positively associated with SAT scores and GPAs, but not with ACT scores. Pre-service teachers at the master's level performed more poorly on the Knowledge Assessment than did those at the bachelor's level ($r = -.18$, $p < .05$), and teaching experience was not related to Knowledge Assessment scores ($r = .08$, $p > .05$; see Appendix C for complete results); however, this may have been an artifact of including only two schools in the pilot sample.

The Operational Data Collection Instrument

The final version of the questionnaire was constructed from items that had been thoroughly cross-referenced to reflect the recommendations in the *NRP Report*, had been subjected to extensive review and pilot testing, and were estimated to take about two hours to administer in total. Table 3-3 shows the distribution and reliability of items across the questionnaire.

The operational data collection instrument consisted of two sections: the Program Survey and the Knowledge Assessment. The Program Survey was made up of three sets of items: background characteristic items, exposure and emphasis items, and feelings of preparedness items. The background characteristic items were single-item measures and therefore did not have internal consistency reliability estimates. Internal consistency reliability estimates (coefficients alpha) for the exposure/emphasis items are reported separately for the Coursework and Field Experience sections of the questionnaire. The reliabilities for the Coursework section ranged from $\alpha=.83$ for the phonemic awareness subscale to $\alpha=.66$ for the fluency subscale. Combining the phonemic awareness subscale with the phonics ($\alpha=.74$) subscale resulted in an $\alpha=.85$ reliability for the alphabetics subscale. Similarly, combining the vocabulary ($\alpha=.77$) subscale with the comprehension subscale ($\alpha=.67$) resulted in a higher reliability ($\alpha=.81$) for the meaning subscale. For the field experience items, reliabilities ranged from $\alpha=.86$ for the phonemic awareness subscale to $\alpha=.64$ for the fluency subscale. As with the coursework items, combining the vocabulary subscale ($\alpha=.74$) with the comprehension subscale ($\alpha=.72$) improved the

reliability (α=.79). Combining phonemic awareness with the phonics subscale (α=.75) did not result in a reliability for the alphabetics subscale that exceeded that of the phonemic awareness items alone (α=.85). For the feelings of preparedness items, reliabilities ranged from α=.81 (phonemic awareness) to α=.51 (fluency). Combining phonemic awareness with phonics (α=.78) raised the reliability of the alphabetics scale (α=.86), as did combining vocabulary (α=.69) with comprehension (α=.67) to form the meaning subscale (α=.80).

The Knowledge Assessment consisted of 56 items, 53 of which were retained after dropping 3 items that did not correlate with the other items. Reliabilities of these scales ranged from α=.25 for the vocabulary subscale to α=.48 for the fluency and comprehension subscales. Combining the phonemic awareness (α=.35) and phonics (α=.37) subscales resulted in an alphabetics subscale with higher reliability (α=.50), as did combining the vocabulary (α=.25) and comprehension subscales to form the meaning subscale (α=.48). Summing all items into a single scale resulted in an overall reliability for the Knowledge Assessment of α=.71. While the overall assessment has a sufficiently high reliability, the subscales used in some analyses have insufficient reliability levels (0.25 – 0.50). Thus, the results of the analyses based on subscales should be interpreted with caution.

Table 3-3. Distribution of items in Pre-Service Teacher Preparation Program and Knowledge Survey in the main study sample

Component		Total items (items used in analyses)	Reliability			
Program Survey						
Part 1:	Background characteristic items	22				†
Part 2:	Exposure/Emphasis items		Coursework		Field Experience	
			n	α	n	α
	Phonemic Awareness	4	2,187	.83	2,184	.86
	Phonics	4	2,169	.74	2,174	.75
	Fluency	4	2,165	.66	2,167	.64
	Vocabulary	4	2,180	.77	2,177	.74
	Comprehension	4	2,137	.67	2,140	.72
	Alphabetics (Phonemic Awareness + Phonics)	8	2,169	.85	2,171	.85
	Meaning (Vocabulary + Comprehension)	8	2,131	.81	2,140	.79
	Total: Exposure/Emphasis	34	2,112	.89	2,124	.88
	Preparedness items				Preparedness	
					n	α
	Phonemic Awareness	2			2,183	.81
	Phonics	2			2,164	.78
	Fluency	2			2,173	.51
	Vocabulary	2			2,178	.69

See notes at end of table.

Table 3-3. Distribution of items in Pre-Service Teacher Preparation Program and Knowledge Survey in the main study sample—Continued

Component	Total items (items used in analyses)	Reliability	
Alphabetics (Phonemic Awareness + Phonics)	4	2,161	.86
Comprehension	2	2,176	.67
Meaning (Vocabulary + Comprehension)	4	2,168	.80
Total: Preparedness	13	2,131	.88
Total: Program Survey	*70*		†
Knowledge Assessment (n = 2,187)			
Phonemic Awareness	12 (11)		.35
Phonics	8 (7)		.37
Fluency	12		.48
Vocabulary	12 (11)		.25
Comprehension	12		.48
Alphabetics (Phonemic Awareness + Phonics)	20 (18)		.50
Meaning (Vocabulary + Comprehension)	24 (23)		.48
Total: Knowledge Assessment	*56 (53)*		*.71*
Total items to be completed in 2 hours	***126***		

† Not applicable; reliability estimates not applicable.
NOTE: Fourteen exposure and emphasis items and three preparedness items do not map on to a particular component of early reading instruction.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Counterbalanced Forms

For the administration of the questionnaire, the Program Survey and the Knowledge Assessment were counterbalanced to create Forms A and B. Form A presented the pre-service teacher background characteristic questions first, followed by the Program Survey questions and then the Knowledge Assessment. Form B presented the Knowledge Assessment first, the Program Survey items, and then the pre-service teacher background characteristic questions. Counterbalancing provided some degree of security during data collection and allowed investigation of any potential order or fatigue effects. Appendix I discusses the analysis of the comparability of the counterbalanced forms.

Data Collection

Data collection took place between late February 2007 and early August 2007. In the 99 participating institutions, 2,237 pre-service teachers completed the questionnaire. Members of the study team or institution staff proctored pre-service teachers during questionnaire administration. Questionnaire administration took place in various locations both on-site at

participating institutions and off-site to accommodate pre-service teachers who were no longer on-site. Participants who completed the questionnaire were given an honorarium of $100.00.

Summary

The study gathered data to answer the research questions guiding the *Study of Teacher Preparation in Early Reading Instruction* by using a two-part, printed questionnaire, the *Pre-Service Teacher Preparation Program and Knowledge Survey*. The two parts of the questionnaire are a Program Survey specifically designed to answer the first research question, and a Knowledge Assessment to answer the second question. The entire questionnaire took respondents approximately two hours to complete. It was administered in counterbalanced formats (Forms A and B) in an effort to investigate form order. During proctored administrations at the participating institutions, 2,237 pre-service teachers completed the questionnaire.

CHAPTER 4:
FINDINGS

This chapter reports findings related to both the primary and secondary research questions for the *Study of Teacher Preparation in Early Reading Instruction*. As discussed previously, two primary and three secondary research questions have guided the study.

- Primary Research Question 1: To what extent does the content of teacher education programs focus on the essential components of early reading instruction?
- Primary Research Question 2: To what extent are graduating pre-service teachers knowledgeable about the essential components of early reading instruction?
- Secondary Research Question 3a: Which characteristics of teacher training institutions and programs are associated with their focus on the essential components of early reading instruction?
- Secondary Research Question 3b: To what extent is teacher training programs' focus on the essential components of early reading instruction associated with pre-service teachers' knowledge about these components?
- Secondary Research Question 3c: To what extent is pre-service teachers' knowledge about the essential components of early reading instruction related to these pre-service teachers' feelings of preparedness to teach various aspects of beginning reading?

Definitions of Variables Used in the Report

Psychometric analyses of the Program Survey and the Knowledge Assessment (see Appendices G and I, respectively) indicated that the two most promising models for presenting the essential components of reading instruction are the three-factor model and the one-factor model. The report presents both; the one-factor model provides the highest reliability and the three-factor model provides information on whether responses vary by essential component and allows the study to examine the relationship between specific, conceptually meaningful components of early reading instruction within the Program Survey and the Knowledge Assessment. The psychometric analyses also suggested that coursework and field experience Program Survey items represent distinct aspects of pre-service teachers' training experiences. The report presents them combined as well as separately.

In discussing the results from both the Program Survey and the Knowledge Assessment, this report refers to variables related to the five essential components of early reading instruction (phonemic awareness, phonics, fluency, vocabulary, and comprehension) as follows:

- All components: information gathered from items on all five essential components of early reading instruction and analyzed as a single factor
- Alphabetics: information from items on phonemic awareness and phonics
- Fluency: information from oral reading fluency items

- Meaning: information from vocabulary and comprehension items

The study team measured variables related to the focus of pre-service teacher programs on the essential components in different aspects of programs, using the following metrics:

- Coursework emphasis on the essential components
  - None (0): not addressed in any of my courses
  - Little (1): addressed only briefly in one course
  - Moderate (2): addressed over several class periods in one or two courses
  - Considerable (3): took a course entirely devoted to this topic
- Field experience exposure to the essential components
  - None (0): did not observe
  - Little (1): observed one or two times
  - Moderate (2): observed many times (3–9)
  - Considerable (3): observed more than 10 times
- Overall focus of pre-service teacher programs on the essential components: the average of reported coursework emphasis and field experience exposure

Discussions of pre-service teachers' feelings of preparedness draw on self-reported answers to Program Survey questions that used the following metrics:

- Not at all prepared (0): do not know or do not understand activities well enough to use them with students
- Somewhat prepared (1): not completely sure how to use with students in all grades and at all reading levels
- Mostly prepared (2): understand how to use with some students but need to deepen understanding
- Definitely prepared (3): completely understand how to use with students in all grades and with all reading levels

Analyses to Answer the Study's Primary Research Questions

The findings presented in this chapter are based on analyses that incorporated non-response-adjusted sampling weights so that they are generalizable to the population of pre-service teachers in the nation from which the study sample was drawn. Thus, the findings are national estimates of the phenomena measured by the study, as determined by responses from pre-service teachers about to graduate from the 99 institutions that agreed to participate in the study.

Research Question 1: National Estimates of Teacher Education Program Focus on the Essential Components of Early Reading Instruction

The first primary research question for this study is intended to gauge the extent to which the content of teacher education programs focuses on the five essential components of early reading instruction: phonemic awareness, phonics, fluency, vocabulary, and comprehension. The hypothesis underlying this question is that the more opportunities pre-service teachers have to learn about the essential components of early reading instruction through coursework and through field-based experiences, the more they will learn about the essential components.

The answers to this question are presented as national estimates. As discussed in Chapter 3, the Program Survey section of the *Pre-Service Teacher Preparation Program and Knowledge Survey* included multiple sections that gathered data to address the question. For much of the analyses conducted to answer the first research question, the study team computed national estimates of program focus by using a three-level hierarchical linear model (HLM) that properly takes into account the nested data structure (i.e., items nested with teachers and teachers nested within states;[27] see Appendix K for details). Table 4-1 presents the national estimates of means, standard errors, and 95 percent confidence intervals of these estimates by essential component and aspect of program. These results are presented graphically in Figures 4-1a and 4-1b.

Table 4-1. National estimates of teacher education programs' focus on the essential components of early reading instruction, by essential component and aspect of program

	Weighted national estimates		95% confidence interval	
Scales	**Mean**	**Standard error**	**Lower bound**	**Upper bound**
Overall focus				
Alphabetics	1.8404	0.0320	1.7744	1.9064
Fluency	1.8523	0.0362	1.7776	1.9270
Meaning	1.6951	0.0313	1.6305	1.7597
All components	1.7571	0.0296	1.6960	1.8182
Coursework emphasis				
Alphabetics	1.7381	0.0440	1.6473	1.8289
Fluency	1.7502	0.0480	1.6511	1.8493
Meaning	1.5980	0.0372	1.5212	1.6748
All components	1.6581	0.0382	1.5793	1.7369
Field experience exposure				
Alphabetics	1.9427	0.0304	1.8800	2.0054
Fluency	1.9542	0.0282	1.8960	2.0124
Meaning	1.7919	0.0303	1.7294	1.8544
All components	1.8561	0.0263	1.8018	1.9104

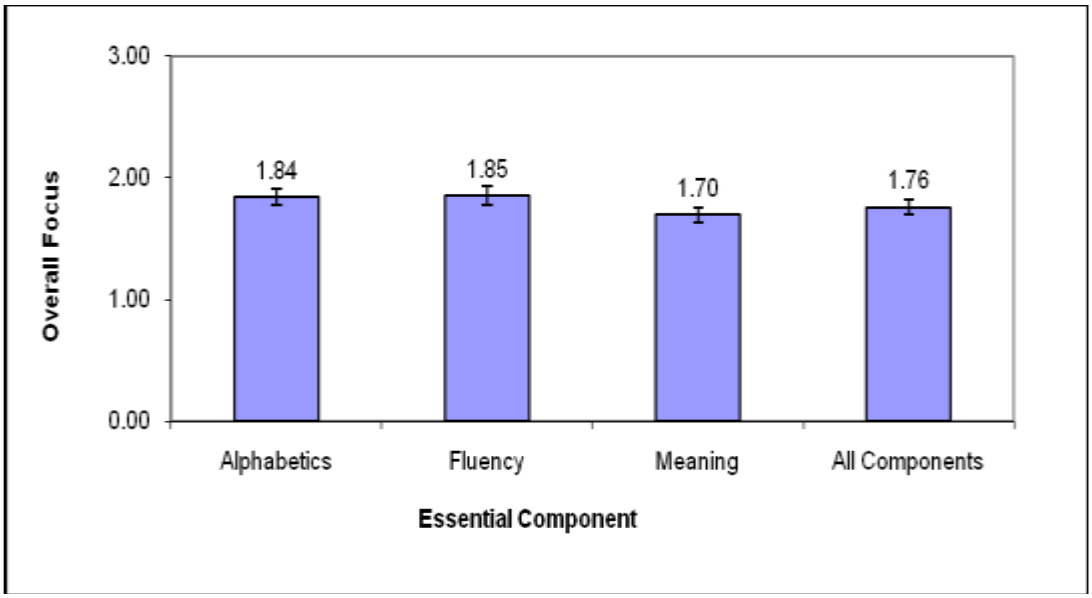
NOTE: Estimates of overall focus were based on coursework and field experience data combined. Program focus based on coursework and field experience data was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

These national estimates can be interpreted using the original metric of the survey questions. Estimates range between 0 and 3, with 0 representing no focus, 1 "little" focus, 2 "moderate" focus, and 3 "considerable" focus. As shown in Table 4-1, pre-service teachers rated the overall focus of their training programs on all the essential components of early reading instruction as

[27] Ideally, the study team would have liked to construct a four-level model, with items nested within pre-service teachers, pre-service teachers nested within institutions, and institutions nested within states. However, the current HLM software program can accommodate only up to three levels. Therefore, the study team omitted the institution level and used states as the level-3 units, as the standard error of estimates in a multilevel model depends primarily on the number of units at the highest level of aggregation (state in this case). Further, the state requirements for initial certification determine the main course offerings that pre-service teachers must take as part of their required programs.

being above "little" but below "moderate" (1.76), with the overall focus being significantly stronger in their field experience (1.86) than in their coursework (1.66) ($p < .001$). Pre-service teachers also reported that the overall focus of their training programs on alphabetics (1.84) and fluency (1.85) was significantly stronger than that on meaning (1.70) ($p < .001$) (see Figure 4-1a).

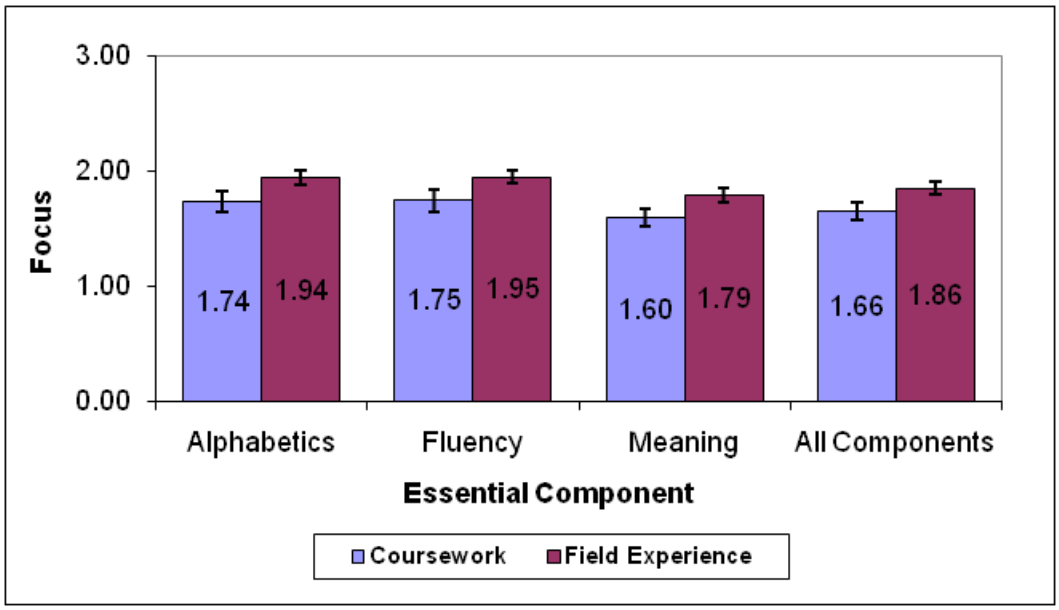
Figure 4-1a. National estimates of overall programmatic focus on the essential components of early reading instruction, by essential component



NOTE: Estimates of overall focus were based on coursework and field experience data combined. Program focus was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. The short vertical bar at the top of each bar represents the 95 percent confidence interval of the national estimate. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Data from the Program Survey also indicate that teacher preparation programs focused more heavily on alphabetics and fluency than on meaning, based on both coursework data and field experience data ($p < .001$) (see Figure 4-1b). Moreover, these programs' focus on all three essential components of early reading instruction (i.e., alphabetics, fluency, and meaning) was stronger in field experience than in coursework ($p < .001$).

Figure 4-1b. National estimates of coursework emphasis on and field experience exposure to the essential components of early reading instruction, by essential component



NOTE: Coursework emphasis and field experience exposure were measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. The short vertical bar at the top of each bar represents the 95 percent confidence interval of the national estimate. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

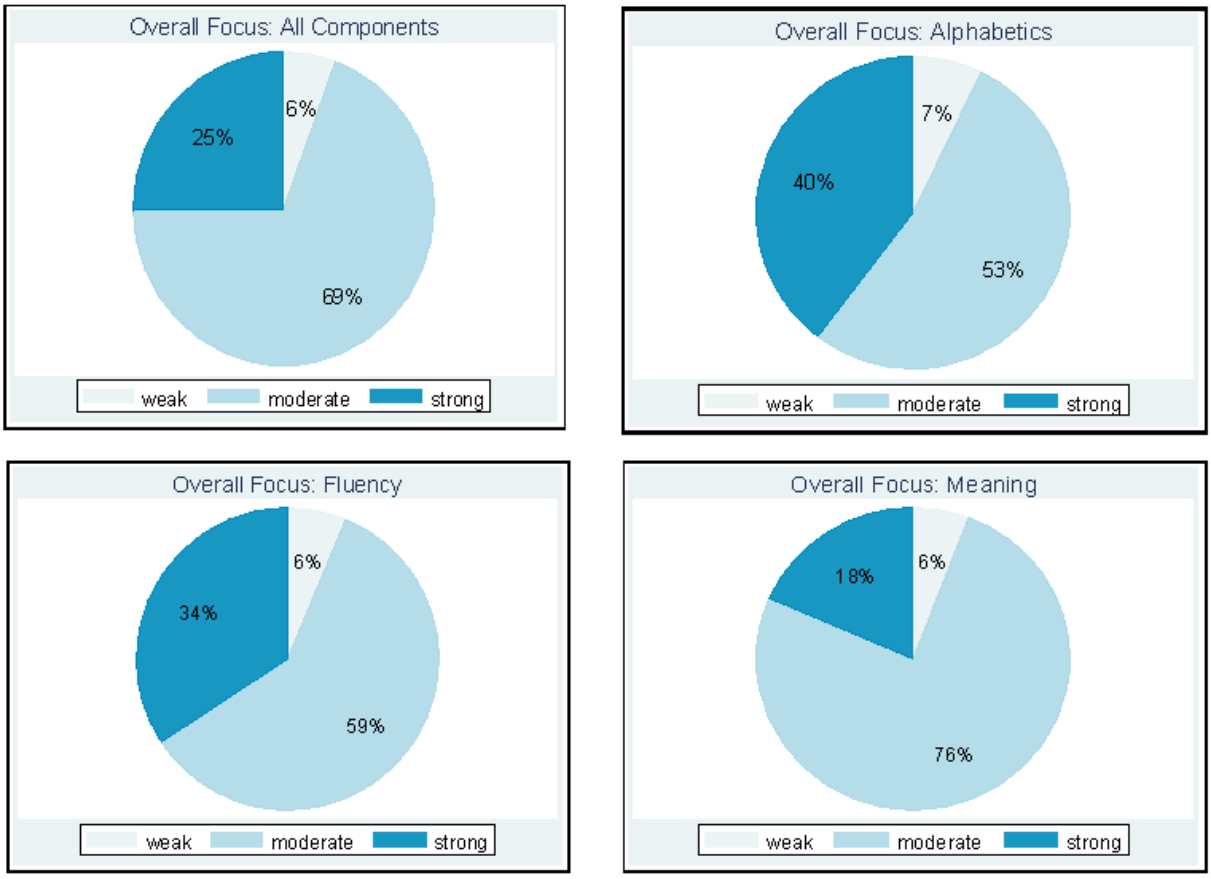
To further examine pre-service teachers' perceptions about the focus of their training program on the essential components of early reading instruction, the study team estimated the percentage of pre-service teachers in each of the three response categories—weak focus, moderate focus, and strong focus.[28] This presentation collapses the 4-point scale from the Program Survey items into three categories. Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2.

Figure 4-2a shows the distribution of pre-service teachers across the three categories, based on data on their overall program (both coursework and field experiences). Across all five essential components, 69 percent of pre-service teachers reported a moderate overall programmatic focus, while 25 percent of pre-service teachers reported a strong overall programmatic focus, and 6 percent reported a weak overall programmatic focus. Consistent with the national estimates presented in Figure 4-1a, the percentage of pre-service teachers reporting a strong overall programmatic focus on alphabetics (40 percent) or fluency (34 percent) was significantly higher

[28] As noted at the beginning of this chapter and in Table 4-1, four response categories were measured; however, the two lowest response categories ("None" and "Little") were selected by a small number of respondents. Thus, the study team elected to combine those two categories, creating the three category response measures that are reported.

than the percentage of pre-service teachers reporting a strong overall focus on meaning (18 percent) ($p < .001$).[29]

Figure 4-2a. Percentage of pre-service teachers reporting weak, moderate, or strong overall program focus on the essential components of early reading instruction, by essential component



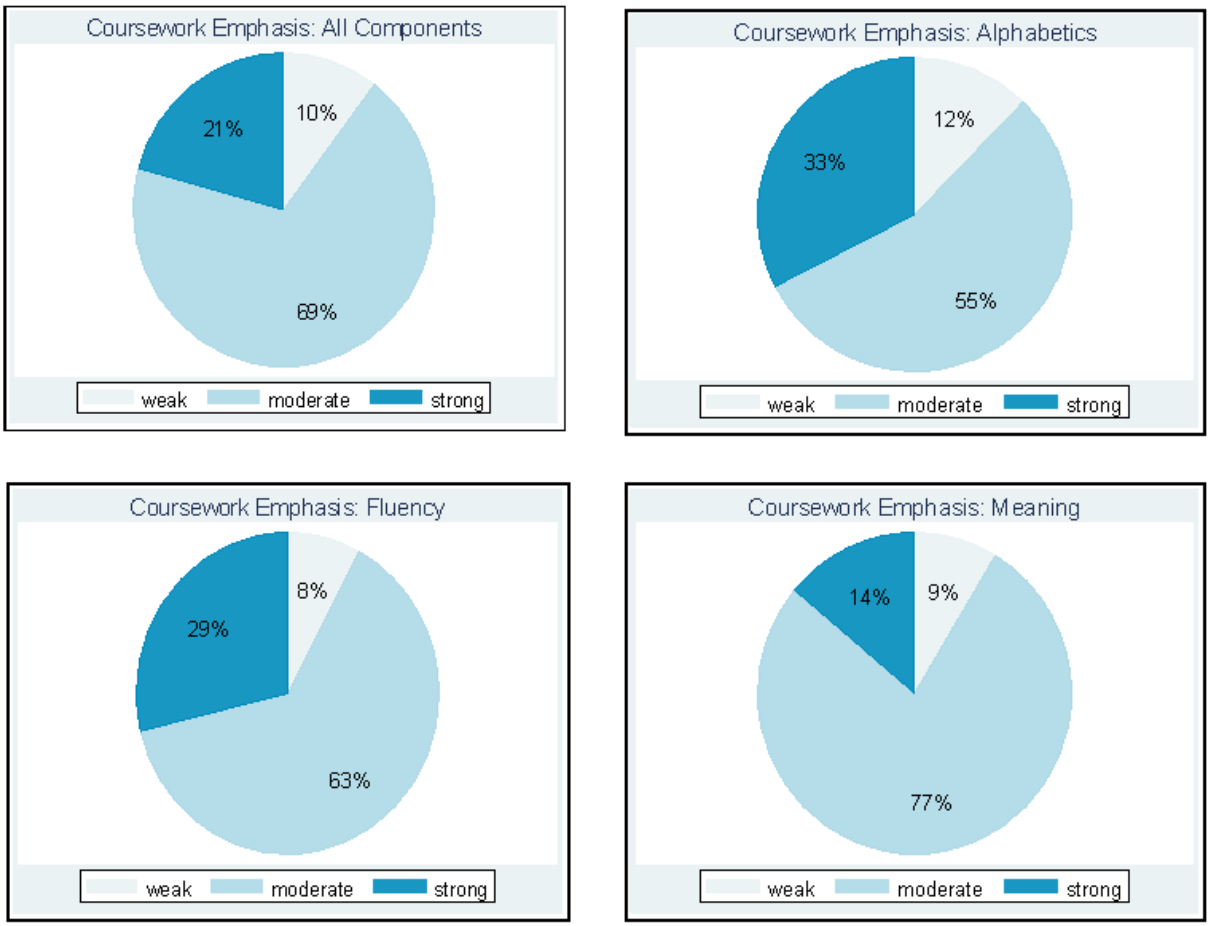
NOTE: Estimates of overall focus were based on coursework and field experience data combined. Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure 4-2b shows the distribution of pre-service teachers across the three response categories based on analysis of coursework emphasis data. Over two-thirds (69 percent) of the pre-service teachers perceived their coursework as having a "moderate" emphasis on all the essential components of early reading instruction. Ten percent of pre-service teachers reported a "weak"

[29] The statistical significance of these differences was tested using the McNemar test, which is a chi-square test of the difference in proportions from a single group. The McNemar test, however, cannot take into account the nested data structure and hence may have led to inflated statistical significance. Therefore, the statistical significance of the differences should be interpreted with caution.

coursework emphasis, and 21 percent reported a "strong" coursework emphasis on the essential components of early reading instruction. The distributions of pre-service teachers across different response categories for the alphabetics, fluency, and meaning scales generally follow a similar pattern, with the majority of pre-service teachers reporting a "moderate" coursework emphasis (55 percent to 77 percent), and fewer than one third of pre-service teachers reporting a "strong" coursework emphasis (14 percent to 33 percent) or a "weak" coursework emphasis (8 percent to 12 percent) on the different components. Again, the findings indicate a significantly stronger coursework emphasis on alphabetics and fluency than on meaning: pre-service teachers were more than twice as likely to report a "strong" emphasis on alphabetics (33 percent) or fluency (29 percent) than on meaning (14 percent) ($p < .001$).

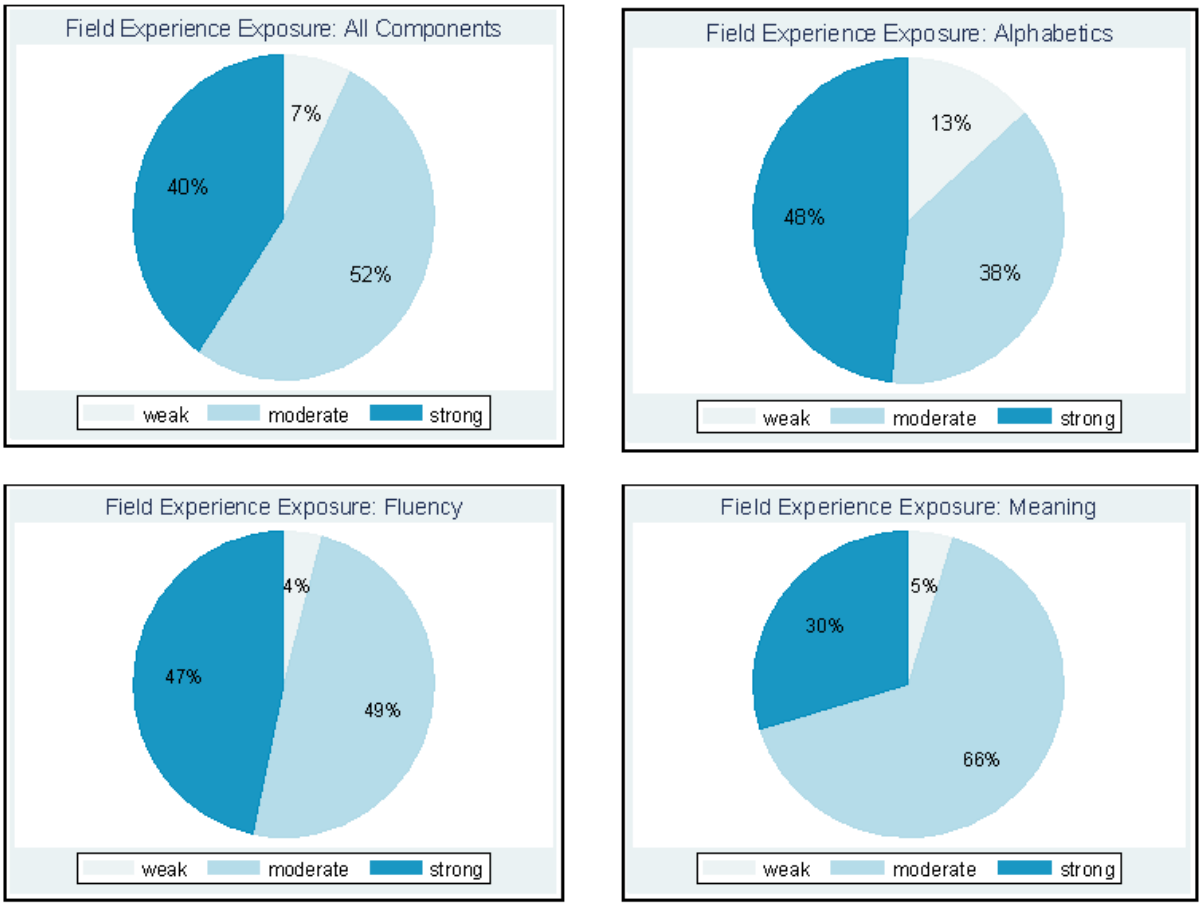
Figure 4-2b. Percentage of pre-service teachers reporting weak, moderate, or strong coursework emphasis on the essential components of early reading instruction, by essential component



NOTE: Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure 4-2c shows the percentage of pre-service teachers in each of the three response categories based on analysis of field experience data. A comparison of Figure 4-2b and Figure 4-2c suggests that pre-service teachers' field experience was more focused on the essential components than was their coursework. For example, the percentage of pre-service teachers reporting a strong exposure to all components as well as to meaning in their field experience was about twice that of coursework emphasis (40 percent vs. 21 percent across all components, and 30 percent vs. 14 percent for meaning) ($p < .001$). Moreover, nearly half of the pre-service teachers reported a "strong" exposure to alphabetics and fluency (48 percent and 47 percent, respectively) in their field experiences, compared with less than a third who reported "strong" emphasis in their coursework on alphabetics and fluency (33 percent and 29 percent, respectively) ($p < .001$).

Figure 4-2c. Percentage of pre-service teachers reporting weak, moderate, or strong field experience exposure to the essential components of early reading instruction, by essential component



NOTE: Pre-service teachers reporting a "weak" focus are those with an estimated value less than or equal to 1, pre-service teachers reporting a "moderate" focus include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting a "strong" focus are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

In summary, data collected with the Program Survey of the questionnaire provide answers to the first research question about the extent to which pre-service teacher education programs focus on the essential components of early reading instruction. Among the nationally representative sample of pre-service teachers nearing the end of their training, 25 percent reported a "strong" overall programmatic focus on the essential components of early reading instruction; 21 percent reported a strong focus when considering only their coursework; and 40 percent reported a strong focus when considering the field experiences required by their program.

In addition to the multicomponent coursework and field experience items for which pre-service teachers rated their programs' focus on the essential components, the survey included a broader question that addresses exposure as well as teachers' perceptions of the knowledge they gained from their program. This question[30] asked pre-service teachers to think about their program and then reply "yes" or "no" to indicate whether they have learned what students must know and be able to do related to each of the five essential components. As shown in Table 4-2, pre-service teachers perceived themselves to have learned the content well, as over 90 percent responded affirmatively for four of the five essential components. For phonemic awareness, 79.3 percent of pre-service teachers reported affirmatively.

Table 4-2. National estimates of pre-service teacher responses to question about having learned "what students need to know and be able to do" in the five essential components of reading

Essential component: activity	% Yes (CI)	SE
Phonemic Awareness: Manipulate phonemes in spoken words	79.3 (77.4–81.3)	1.94
Phonics: Associate letters and their sounds to identify words	94.8 (94.0–95.6)	0.77
Fluency: Read orally with appropriate speed, accuracy, expression	91.3 (89.6–93.0)	1.73
Vocabulary: Understand meanings of words and learn new words	92.6 (91.5–93.8)	1.15
Comprehension: Understand what is read	95.7 (94.8–96.6)	0.86

NOTE: N of teachers = 2,183 for phonemic awareness, 2,186 for phonics, 2,184 for fluency, 2,183 for vocabulary, and 2,184 for comprehension; N of institutions = 99; N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Research Question 2: National Estimates of Pre-Service Teachers' Knowledge about the Essential Components of Early Reading Instruction

The second primary research question for this study asks how knowledgeable pre-service teachers are about the essential components of reading instruction. The hypothesis underlying this question is that the more knowledgeable pre-service teachers are about the essential components, the better prepared they will be to teach their students.

[30] See the Program Survey in Appendix A, page A-5.

To calculate the national estimates from the Knowledge Assessment data, the study team accounted for clustering at the state level, consistent with what was done for the program focus estimates. Table 4-3 presents the raw scores and percent correct on the Knowledge Assessment.

On average, pre-service teachers responded correctly to 57 percent of the items on the Knowledge Assessment (30.2 of 53 items used in the analysis). On average, pre-service teachers correctly answered 61 percent of fluency items, 58 percent of meaning items, and 53 percent of alphabetics items. It is important to bear in mind the limited reliability of the Knowledge Assessment subscales when interpreting the national estimates for alphabetics, fluency, and meaning..

Table 4-3. Mean raw scores and percent correct for the Knowledge Assessment

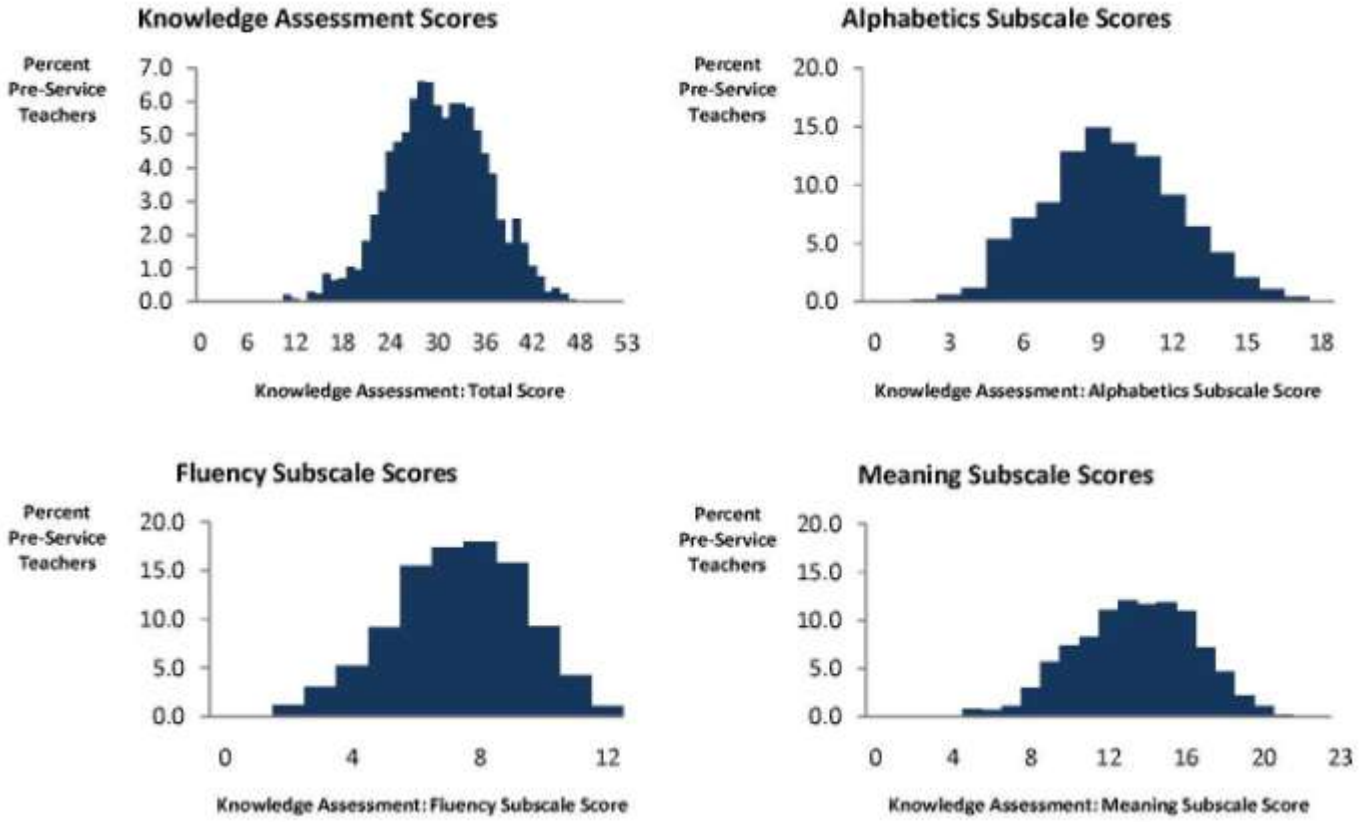
Essential Component	Mean raw score (SE)	Mean percent correct (SE)
All Components	30.21 (0.43)	57.00 (0.82)
Alphabetics	9.53 (0.17)	52.96 (0.93)
Fluency	7.31 (0.16)	60.93 (1.37)
Meaning	13.36 (0.13)	58.11 (0.57)

NOTE: N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Figure 4-3 shows the distribution of scores on the total Knowledge Assessment and for the alphabetics, fluency, and meaning subscales. The majority of pre-service teachers (84.7 percent) scored in the mid-range of the distribution on the Knowledge Assessment. With the exception of fluency, fewer than 2 percent of pre-service teachers scored in the top and bottom score bands in each component, and none scored in the top and bottom score bands on the overall assessment.

To examine the validity of the Knowledge Assessment scores further, the study team analyzed the assessment results by pre-service teacher self-reported characteristics, including SAT or ACT scores, their cumulative GPAs overall and in education courses, and the nature of their degree program. These analyses consisted of calculating simple descriptive statistics for the Knowledge Assessment and the alphabetics, fluency, and meaning subscales on each of the subgroups defined by pre-service teacher characteristics. The results of these analyses show that higher scores on the Knowledge Assessment correspond to self-reports of higher scores on the various measures of achievement, including overall GPA (polyserial $\rho = .35$, $p < .01$), education GPA (polyserial $\rho = .34$, $p < .01$), combined SAT (polyserial $\rho = .37$, $p < .01$), ACT (polyserial $\rho = .39$, $p < .01$), and GRE (polyserial $\rho = .40$, $p < .01$). Note that the results for overall GPA and education GPA are related. Likewise, scores on the Knowledge Assessment are higher for those pre-service teachers in post-baccalaureate/fifth-year programs and in master's programs (polyserial $\rho = .15$, $p < .01$). The results of these analyses are in Appendix J.

Figure 4-3. Histograms of Knowledge Assessment scores, by total and subscale



NOTE: N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Analyses to Answer the Study's Secondary Research Questions

The national estimates described in the previous section provide information about the extent to which teacher education programs focus on the essential components of early reading instruction and the extent to which pre-service teachers are knowledgeable about these essential components of early reading instruction. To gain further insight into these national estimates, the study team conducted three additional analyses guided by the following secondary research questions:

- Research Question 3a: Which characteristics of teacher training institutions and programs are associated with their focus on the essential components of early reading instruction?
- Research Question 3b: To what extent is teacher training programs' focus on the essential components of early reading instruction associated with pre-service teachers' knowledge about these components?
- Research Question 3c: To what extent is pre-service teachers' knowledge about the essential components of early reading instruction related to these pre-service teachers' feelings of preparedness to teach various aspects of beginning reading?

Although the study design does not support a causal analysis of these questions, the relationships described above were estimated using correlational techniques. The study team conducted correlational analyses for the scale denoting all of the essential components combined and for each of the three subscales of early reading instruction (i.e., alphabetics, fluency, and meaning). Sampling weights were applied to these analyses so that the findings can be generalized to all pre-service teachers in the nation. The Program Survey data were self-reports by pre-service teachers of their perceptions of their programs and feelings of preparation. Confirmation of these reports was beyond the scope of the work in this study. The data presented should not be construed as indicators of the content or quality of the teacher training programs that agreed to participate in the study.

Research Question 3a: Institutional Characteristics, Programmatic Concentration or Major, and Pre-Service Teachers' Perception of the Focus of Their Programs on the Essential Components of Early Reading Instruction

Research Question 3a addresses potential differences among pre-service teachers from institutions according to their sector designation (public or private funding), the highest degree offered, and the areas within the teacher training program in which the respondents have studied. The study team computed the national estimates for the various subpopulations and tested the differences between subpopulations, taking into account the sampling design of the study. To guard against inflated Type I errors (i.e., obtaining false findings due to chance) resulting from multiple pairwise comparisons of more than two groups, we first performed a global F test of each measure to determine whether there was a significant overall difference among the groups and conducted pairwise comparisons only if the global test indicated a significant overall difference. The results of these analyses are presented in Tables 4-4, 4-5, and 4-6. Although the study team confirmed institutional characteristics with public information, data from the pre-service teachers about their programs were gathered through self-report. In addition, the analysis to address this question is correlational in nature and thus no causal inferences can be drawn from the results.

Using data collected during the "reconnaissance" phase after the sample had been selected but before recruiting began, the study team categorized sampled institutions according to their public or private status and the highest degree offered.[31] The "major or concentration" of the pre-service teachers were determined through their response to question 13b in the Program Survey.[32]

---

[31] Depending on their states' certification requirements, their institutions, and at times their own prior education, pre-service teachers in the study might be working toward a bachelor's or a master's degree. Therefore, the study team decided to distinguish between institutions that do and do not offer doctor of education (EdD) or doctor of philosophy in education (PhD) degrees, even though the Carnegie classification makes finer distinctions (baccalaureate general, baccalaureate liberal education; masters I and masters II; and doctoral extensive and doctoral intensive) (McCormick, 2001).

[32] See question 13b in the Program Survey in Appendix A. Choices include early childhood education, elementary education, combined early childhood/elementary, combined early childhood/special education, combined elementary education/special education, combined early childhood/elementary/special education, curriculum and instruction, reading education, multi/interdisciplinary studies, and other. Multi/interdisciplinary studies is the term used to designate academic majors of pre-service teachers in states such as Texas that do not recognize an undergraduate program in education in conferring certification.

Table 4-4 presents the national estimates of pre-service teachers' perceptions of their teacher education programs' focus on the essential components of early reading instruction for public and private institutions, respectively. It shows that program focus in these two types of institutions was similar across different components of early reading. Statistical tests of group mean difference indicate that of the 12 measures of program focus (i.e., alphabetics, fluency, meaning, and all components, based on coursework emphasis, field experience exposure, and overall focus), the only statistically significant difference between public and private institutions was in the area of fluency, based on the aspect of coursework emphasis. These tests were performed using the HLM analysis described previously for Research Question 1. Further details on the model can be found in Appendix K. Specifically, pre-service teachers in public institutions reported a stronger focus on fluency in their coursework than pre-service teachers in private institutions (1.78 compared with 1.72, $p < .01$).

Table 4-4. National estimates of teacher education programs' focus on the essential components of early reading instruction, by essential component, aspect of program, and by sector of institution

	Public (n = 1,474)		Private (n = 763)	
Scales	Mean	SE	Mean	SE
Overall focus				
Alphabetics	1.84	0.02	1.85	0.02
Fluency	1.88	0.03	1.82	0.02
Meaning	1.70	0.02	1.66	0.04
All components	1.77	0.02	1.73	0.03
Coursework emphasis				
Alphabetics	1.75	0.04	1.77	0.03
Fluency**	1.78	0.05	1.72	0.02
Meaning	1.60	0.04	1.58	0.03
All components	1.66	0.04	1.65	0.02
Field experience exposure				
Alphabetics	1.94	0.03	1.94	0.02
Fluency	1.99	0.02	1.93	0.01
Meaning	1.80	0.01	1.74	0.05
All components	1.87	0.01	1.82	0.03

NOTE: Estimates of overall focus were based on coursework and field experience data combined. Program focus based on coursework and field experience data was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. Sample size in parentheses is the number of pre-service teachers in this study in each type of institution. Among the 99 institutions in the study sample, 65 are public institutions and 34 are private institutions. Statistically significant finding at the $p < .01$ level is indicated by **.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table 4-5 presents the national estimates of teacher education programs' focus on the essential components of early reading instruction by the highest degree offered by the institution. The asterisks in the first column of the table show that a significant overall difference among the three types of institutions existed for 10 of the 12 measures. Further pairwise comparisons for each of the 10 measures indicate that 24 of the 30 group differences tested were statistically significant. Most (20) of these significant differences point to a stronger program focus in institutions where the highest degree offered was the bachelor's degree than in other institutions.

The overall program focus across all components, for example, was 1.96 in institutions where the highest degree offered was the bachelor's degree, which was significantly higher than that for institutions where the highest degree offered was the master's degree (1.71) or doctoral degree (1.77) (see Table 4-5). There were four statistically significant differences in program focus between institutions where the highest degree offered was master's degree and institutions where the highest degree offered was doctoral degree. In all four cases, program focus was significantly stronger in institutions where the highest degree offered was the doctoral degree.

Table 4-5. National estimates of teacher education programs' focus on the essential components of early reading instruction, by essential component, aspect of program, and by the highest degree offered by institution

Scales	Bachelor's (n = 115)		Master's (n = 1,342)		Doctoral (n = 780)		Significance of difference		
	Mean	SE	Mean	SE	Mean	SE	BA vs. MA	BA vs. Doc	MA vs. Doc
Overall focus									
Alphabetics*	2.02	0.07	1.82	0.02	1.84	0.04	*	*	
Fluency**	2.13	0.06	1.81	0.03	1.89	0.03	***	**	*
Meaning***	1.86	0.03	1.65	0.03	1.71	0.01	***	***	
All components***	1.96	0.01	1.71	0.02	1.77	0.02	***	***	
Coursework emphasis									
Alphabetics***	2.17	0.01	1.68	0.04	1.77	0.05	***	***	
Fluency***	2.15	0.07	1.67	0.04	1.80	0.04	***	**	*
Meaning***	1.85	0.04	1.54	0.03	1.62	0.02	***	***	
All components***	2.00	0.04	1.59	0.03	1.68	0.03	***	***	
Field experience exposure									
Alphabetics	1.87	0.15	1.96	0.02	1.92	0.02			
Fluency*	2.08	0.06	1.94	0.01	1.97	0.02	*	*	*
Meaning*	1.85	0.02	1.76	0.03	1.80	0.01	*	*	*
All components	1.90	0.02	1.83	0.02	1.86	0.01			

NOTE: Estimates of overall focus were based on coursework and field experience data combined. Program focus based on coursework and field experience data was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. Sample size in parentheses is the number of pre-service teachers in this study in each type of institution. Among the 99 institutions in the study sample, 35 offer doctoral degrees, 59 offer master's degrees as the highest degree offered, and 5 only offer bachelor's degrees.
* $p < .05$; ** $p < .01$; ***$p < .001$.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table 4-6 presents the national estimates of teacher education programs' focus on the essential components of early reading instruction by program type. Again, statistically significant differences were found among pre-service teachers in different types of teacher training programs. As indicated by the asterisks in the first column of the table, significant overall differences among the three types of teacher training programs (i.e., early childhood education, elementary education, and combined programs)[33] existed for 9 of the 12 measures. The study

[33] Early childhood programs prepare pre-service teachers for certification to teach pre-school to grade three; elementary programs focus on preparation to teach kindergarten or grade one through grade six. Combined

team did not include "Other" programs in comparisons among different types of programs because the number of pre-service teachers in "Other" programs (53) is too small to draw reliable conclusions. Of the 27 pairwise comparisons conducted for the 9 measures, 22 were statistically significant. Most (16) of these significant differences favored early childhood education programs. The rating of overall program focus across all components, for example, was 1.83 among pre-service teachers in early childhood education programs, which was significantly higher than that among pre-service teachers in elementary education programs (1.73) or in combined programs (1.79) (see Table 4-7). Program focus on alphabetics and fluency was also significantly stronger in early childhood education programs than in elementary programs or combined programs (2.07 vs. 1.80 and 1.87, respectively, for alphabetics; 1.95 vs. 1.84 and 1.86, respectively, for fluency). Coursework focus on meaning, however, was significantly stronger in combined programs (1.64) than in early childhood education programs (1.57) or elementary programs (1.58). It is worth noting that the results in Tables 4-5, 4-6, and 4-7 are not independent to the extent that program type, degrees offered, and institutional control are related.

---

programs are those that include early childhood and/or elementary education, plus another major or concentration. See Appendix A, page A-2 for the list of combined programs.

Table 4-6. National estimates of teacher education programs' focus on the essential components of early reading instruction, by essential component, aspect of program, and by program type

Scales	Early childhood education (n = 334)		Elementary education (n = 1,399)		Combined programs (n = 448)		Other (n = 53)		Significance of difference		
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	ECE vs. EE	ECE vs. Comb.	EE vs. Comb.
Overall focus											
Alphabetics***	2.07	0.03	1.80	0.02	1.87	0.03	1.80	0.08	***	***	*
Fluency**	1.95	0.03	1.84	0.03	1.86	0.02	1.84	0.06	**	**	
Meaning	1.70	0.02	1.67	0.03	1.73	0.01	1.72	0.05			
All components*	1.83	0.02	1.73	0.03	1.79	0.02	1.76	0.07	**	*	*
Coursework emphasis											
Alphabetics*	1.90	0.05	1.72	0.04	1.76	0.04	1.82	0.10	**	*	
Fluency*	1.82	0.02	1.74	0.05	1.76	0.03	1.84	0.05	*	*	
Meaning**	1.57	0.02	1.58	0.04	1.64	0.02	1.67	0.07		**	*
All components	1.69	0.02	1.64	0.04	1.69	0.03	1.75	0.08			
Field experience exposure											
Alphabetics***	2.26	0.04	1.87	0.03	1.98	0.02	1.77	0.09	***	***	**
Fluency**	2.07	0.03	1.95	0.02	1.97	0.01	1.85	0.07	**	*	
Meaning	1.83	0.02	1.76	0.03	1.81	0.01	1.76	0.08			
All components***	1.98	0.02	1.82	0.02	1.88	0.01	1.76	0.10	***	***	*

NOTE: Estimates of overall focus were based on coursework and field experience data combined. Program focus based on coursework and field experience data was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. Sample size in parentheses is the number of pre-service teachers in this study in each type of program. Number of institutions represented by each program type: 58 for early childhood education, 94 for elementary education, 73 for combined programs, and 32 for other programs.
*$p < .05$; ** $p < .01$; *** $p < .001$.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Research Question 3b: Relationship between Program Focus and Knowledge of the Essential Components of Early Reading Instruction

Research Question 3b addresses the extent to which the focus of pre-teacher training programs on the essential components of early reading instruction is associated with pre-service teachers' knowledge about these components. Data from the Program Survey and the Knowledge Assessment informed the answer to this question. The Program Survey data were gathered through self-report. In addition, the analysis to address this question is correlational in nature and thus no causal inferences can be drawn from the results.

Analyses were conducted to assess whether teachers who reported a stronger focus in their pre-service training program on a given area scored higher on the items in the Knowledge

Assessment that addressed that particular area. The association between the teachers' perception of their training programs' focus on components of early reading instruction and pre-service teachers' knowledge was estimated for both coursework emphasis and field experience exposure. Based on the psychometric analyses (see Appendix G), estimates of coursework emphasis are at the institution level and field experience exposure is at the individual level.

A three-level Rasch-model HLM was used to take into account the nested nature of the data (i.e., teachers nested within institutions and institutions nested within states). The Rasch model estimates of pre-service teacher knowledge were used only in this phase of the analysis. In addition to the two main predictors (i.e., program emphasis as reflected in coursework and field experience), three control variables were included in the model to account for the background characteristics of the pre-service teachers: certification status, degree level, and race/ethnicity status.[34] Certification status indicates whether the pre-service teacher held any previous teaching certification (0 represents "not certified" and 1 represents "certified"). Degree level represents the degree a pre-service teacher was working toward (0 represents "undergraduate or post-baccalaureate"[35] and 1 represents "graduate"). Race/ethnicity status is based on pre-service teachers' self-reported racial/ethnic background and was coded such that 0 represents White/Asian, which includes White and Asian groups, and 1 represents non-White/Asian, which includes all other racial/ethnic groups. The grouping was based on whether the racial/ethnic group was considered historically underrepresented. Details on the HLM model used to answer this question are presented in Appendix K.

Table 4-7 presents the results for all components combined and for the three subscales of early reading instruction.

[34] The relationships between pre-service teachers' reports of background characteristics and of program focus may actually vary across different institutions. However, these were not the focus of this study; therefore, for simplicity, they were modeled as fixed effects at both the institution level and the state level.

[35] In some states, students attend a fifth-year teacher training program and receive a post-baccalaureate degree that qualifies them to seek initial certification. The coursework does not result in conferral of a master's degree.

Table 4-7. Regression coefficients between degree of program focus and pre-service teachers' knowledge, by essential component and aspect of program

Scale	Coefficient	p-value
All components		
Intercept	0.05	0.407
Field experience exposure	–0.07*	<.001
Coursework emphasis	0.16	0.204
Alphabetics		
Intercept	–0.09	0.263
Field experience exposure	–0.03	0.175
Coursework emphasis	0.11	0.518
Fluency		
Intercept	0.07	0.530
Field experience exposure	–0.03	0.174
Coursework emphasis	0.23*	0.024
Meaning		
Intercept	0.04	0.465
Field experience exposure	–0.05	0.075
Coursework emphasis	0.05	0.656

NOTE: Statistically significant findings at the $p \leq .05$ level are indicated by *. The analysis sample size is 2,156.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

As can be seen in Table 4-7, there is no statistically significant relationship between pre-service teachers' reports of field experience exposure to alphabetics, fluency, or meaning and their score on the Knowledge Assessment in these components. However, there is a statistically significant negative relationship (–0.07, $p < .05$) between pre-service teachers' reports of field experience exposure to all components combined and their overall Knowledge Assessment score. For coursework, there is no statistically significant relationship between emphasis across all components of early reading instruction and the overall score on the Knowledge Assessment. There is no statistically significant relationship between coursework emphasis on alphabetics or on meaning and pre-service teachers' Knowledge Assessment scores on these components. However, there is a statistically significant positive relationship (0.23, $p < .05$) between coursework emphasis on fluency and pre-service teachers' fluency Knowledge Assessment scores. Pre-service teachers who reported a stronger emphasis in their coursework on fluency scored higher on the fluency items in the Knowledge Assessment.

Research Question 3c: Relationship between Pre-Service Teachers' Knowledge and Their Feelings of Preparedness to Teach

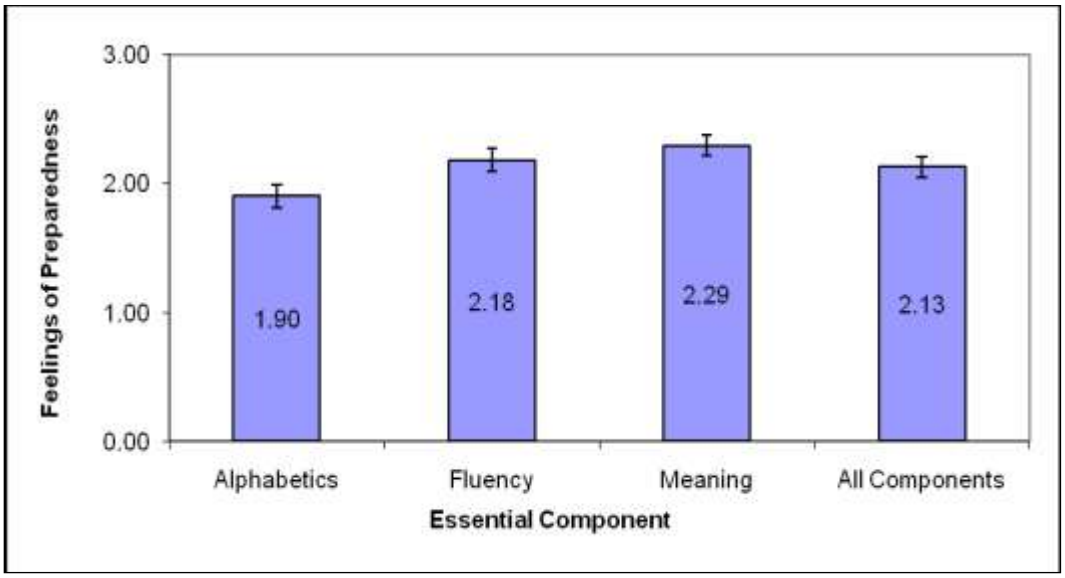
The final supplementary research question concerns the relationship between pre-service teachers' feelings of preparedness to teach the essential components of early reading instruction and their knowledge of the components. Discussions of pre-service teachers' feelings of preparedness draw on answers to Program Survey questions that presented the following metrics for these feelings:

- Not at all prepared (0): do not know or do not understand activities well enough to use them with students
- Somewhat prepared (1): not completely sure how to use with students in all grades and at all reading levels
- Mostly prepared (2): understand how to use with some students but need to deepen understanding
- Definitely prepared (3): completely understand how to use with students in all grades and with all reading levels

Specifically, analyses to answer this question investigated whether teachers who scored higher on the Knowledge Assessment reported feeling more or less prepared to teach the essential components of early reading instruction. The analysis to address this question is correlational in nature and thus no causal inferences can be drawn from the results.

Figure 4-4 presents national estimates of means and the standard errors and the 95 percent confidence intervals of pre-service teachers' feelings of preparedness to teach the essential components of early reading instruction.

Figure 4-4. National estimates of pre-service teachers' feelings of preparedness to teach the essential components of early reading instruction



NOTE: Feelings of preparedness was measured on a 4-point scale: 0 = not at all prepared, 1 = somewhat prepared, 2 = mostly prepared, and 3 = definitely prepared. The short vertical bar at the top of each bar represents the 95 percent confidence interval of the national estimate. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

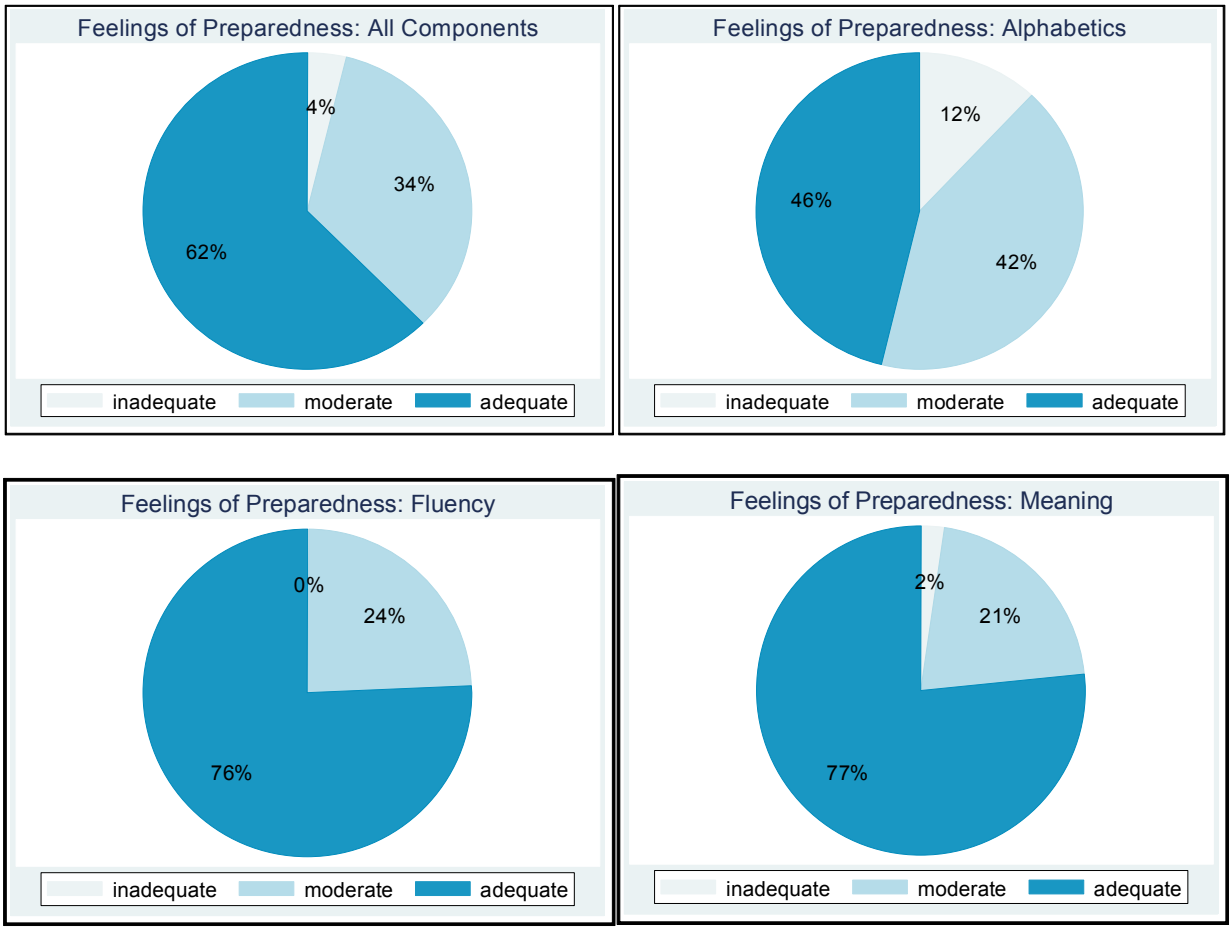
Like program focus, pre-service teachers' feelings of preparedness were measured on a 4-point scale, with 0 indicating "not at all prepared," 1 "somewhat prepared," 2 "mostly prepared," and 3 "definitely prepared." Findings presented in Figure 4-4 show that overall, pre-service teachers felt that they were above "mostly prepared" (2.13) for teaching the essential components of early reading instruction. What is notable is that pre-service teachers reported that they considered

themselves to be significantly better prepared to teach meaning (2.29) than to teach alphabetics (1.90, $p < .001$) or fluency (2.18, $p < .001$), even though their perceived program focus on meaning was weaker than that on alphabetics or fluency based on both their coursework emphasis and field experience exposure (see Figures 4-2a, 4-2b, and 4-2c).

To further examine pre-service teachers' feelings of preparedness, the study team estimated the percentage of pre-service teachers in each of the three response categories—inadequately prepared, moderately prepared, and adequately prepared. This presentation collapses the 4-point scale from the Program Survey items into three categories. Pre-service teachers reporting feeling "inadequately prepared" are those with an estimated value less than or equal to 1, pre-service teachers reporting feeling "moderately prepared" include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting feeling "adequately prepared" are those with an estimated value greater than 2.

As shown in Figure 4-5, the majority of pre-service teachers (62 percent) felt that across all components, they were "adequately" prepared to teach the essential components of early reading instruction, about a third (34 percent) felt that they were "moderately" prepared, and 4 percent felt that they were "inadequately" prepared. Based on their self report, pre-service teachers perceived themselves to be particularly well prepared to teach fluency and meaning, with over three quarters of pre-service teachers feeling "adequately" prepared to teach these components (76 percent and 77 percent, respectively). Less than half of the pre-service teachers (46 percent), however, felt that they were "adequately" prepared to teach alphabetics.

Figure 4-5. Percentage of pre-service teachers reporting feeling inadequately, moderately, or adequately prepared to teach the essential components of reading instruction, by essential component



NOTE: Pre-service teachers reporting feeling "inadequately prepared" are those with an estimated value less than or equal to 1, pre-service teachers reporting feeling "moderately prepared" include those with an estimated value greater than 1 but less than or equal to 2, and pre-service teachers reporting feeling "adequately prepared" are those with an estimated value greater than 2 on the 4-point scale of the Program Survey items. N of teachers = 2,187; N of institutions = 99; and N of states = 24.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

The study team also investigated the relationships between pre-service teachers' knowledge about the essential components of early reading instruction and their reported feelings of preparedness. The HLM used for addressing Research Question 3b was adapted to address this research question; it is described in Appendix K.

Table 4-8 presents the results for each of the three subscales as well as for all components combined.

Table 4-8. Regression coefficients between pre-service teachers' knowledge and their feelings of preparedness to teach, by essential component

Component	Coefficient	p-value
All components		
Intercept	0.04	0.637
Teacher knowledge	0.00	0.950
Alphabetics		
Intercept	-0.02	0.764
Teacher knowledge	0.04	0.104
Fluency		
Intercept	0.06	0.526
Teacher knowledge	0.01	0.751
Meaning		
Intercept	0.07	0.424
Teacher knowledge	0.07*	0.009

NOTE: Statistically significant findings at the $p \leq .05$ level are indicated by *. The analysis sample size is 2,156.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

The results show that after controlling for the pre-service teachers' background characteristics, pre-service teachers' knowledge is statistically significantly related to their feelings of preparedness (e.g., not at all prepared, somewhat prepared, mostly prepared, definitely prepared) in one area of early reading instruction: meaning (0.07, $p < .05$). Pre-service teachers' feelings of preparedness to teach across all the components combined or to teach alphabetics and fluency specifically are not significantly associated with their scores on the Knowledge Assessment in those areas. In other words, pre-service teachers who scored higher on the Knowledge Assessment did not feel more prepared about their abilities to teach in those areas than the pre-service teachers who scored lower and vice versa.

Summary

Data collected with the Program Survey provide national estimates of pre-service teachers' perceptions of the focus of their training programs on the essential components of early reading instruction.

- On average, pre-service teachers rated the overall focus (based on coursework and field experience data combined) of their training programs as being above "little" but below "moderate," or 1.76 on a zero-to-three scale. On average, pre-service teachers also rated coursework emphasis (1.66) and field experience exposure (1.86) as being above "little" but below "moderate" on a zero-to-three scale.

- Sixty-nine percent of pre-service teachers reported a moderate overall programmatic focus (rating greater than 1, but less than or equal to 2 on a zero-to-three scale) on the essential components of early reading instruction, 25 percent reported a strong focus

(rating greater than 2 on the scale), and 6 percent reported a weak focus (rating less than or equal to 1 on the scale).

- Pre-service teachers were almost twice as likely to report a strong overall programmatic focus on alphabetics (40 percent) and fluency (34 percent) than on meaning (18 percent). This was also the case for coursework emphasis (33 percent for alphabetics, 29 percent for fluency, and 14 percent for meaning).

When asked whether they have learned what students must know and be able to do related to each of the five essential components, the majority of pre-service teachers answered affirmatively (between 79.3 percent and 95.7, depending on the essential component). The study also included a Knowledge Assessment to provide national estimates of pre-service teachers' knowledge of the essential components of early reading instruction.

- On average, pre-service teachers responded correctly to 57 percent of the items on the Knowledge Assessment. On average, pre-service teachers correctly answered 61 percent of fluency items, 58 percent of meaning items, and 53 percent of alphabetics items.

Exploratory analyses examining relationships between (1) pre-service teachers' perceptions of program focus and characteristics of teacher training institutions and programs, (2) pre-service teachers' perceptions of program focus and their knowledge, and (3) pre-service teachers' knowledge and their feelings of preparedness did not provide consistent insight into the overall findings.

REFERENCES

Baumann, J. F., Hoffman, J. V., Duffy-Hester, A. M., & Ro, J. M. (2000). "The First R" Yesterday and Today: U.S. Elementary Reading Instruction Practices Reported by Teachers and Administrators. *Reading Research Quarterly, 35*(3), 338–377.

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.

Costigan, T. E., Baker, D. P., Day, R., Smith, M. A., & Salinger, T. S. (2005). *The Teacher Knowledge of Student Content Engagement (TK-SCE) Instrument: Pilot Test Results.* Washington, DC: American Institutes for Research.

Darling-Hammond, L. (1997). *Doing What Matters Most: Investing in Teacher Quality.* Kutztown, PA: National Commission on Teaching and America's Future.

DeFord, D. E. (1985). Validating the Construct of Theoretical Orientation in Reading Instruction. *Reading Research Quarterly, 20*, 351–367.

Evans, M. A., Fox, M., Cremaso, L., & McKinnon, L. (2004). Beginning Reading: The Views of Parents and Teachers of Young Children. *Journal of Educational Psychology*, *96*, 130–141.

Gittomer, D., Latham, A., Ziomek, A., & Ziomeck, R. (1999). *The Academic Quality of Prospective Teachers: The Impact of Admissions and Licensure Testing.* Princeton, NJ: Educational Testing Service.

Kish, L. (1995). *Survey Sampling.* New York: John Wiley & Sons.

Levine, A. (2006). *Educating School Teachers.* Washington, DC: The Education Schools Project.

Mather, N., Bos, C., & Babur, N. (2001). Perceptions and Knowledge of Preservice and Inservice Teachers About Early Literacy Instruction. *Journal of Learning Disabilities*, *34*(5), 472–482.

National Institute of Child Health and Human Development. (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* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Raudenbush, S. W., Rowan, B., & Kang, S. J. (1991). A Multilevel, Multivariate Model for Studying School Climate With Estimation via the EM Algorithm and Application to U.S. High-School Data. *Journal of Educational Statistics*, *16*(4), 295–330.

Rosnow, R. L., & Rosenthal, R. (1996). Computing Contrasts, Effect Sizes, and Counternulls on Other People's Published Data: General Procedures for Research Consumers. *Psychological Methods*, *1*, 331–340.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing Reading Difficulties in Young Children.* Washington, DC: National Academy Press.

Snow, C. E., Griffin, P. & Burns, M. S. (Eds.). (2005). *Knowledge to Support the Teaching of Reading: Preparing Teachers for a Changing World.* Washington, DC: National Academy Press.

U.S. Department of Education, National Center for Education Statistics. (1999). *Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers* (NCES 1999-080) by Laurie Lewis, Basnat Parsad, Nancy Carey, Nicole Bartfai, Elizabeth Farris, & Becky Smerdon. Bernie Green, project officer. Washington, DC: Author.

APPENDIX A.
PRE-SERVICE TEACHER PROGRAM SURVEY

SECTION ONE
PRE-SERVICE TEACHER PROGRAM SURVEY
(Approximately 1 hour)

Some of the following items will have only one best answer for your personal experiences in your teacher education program. For those items, please completely fill in the box that corresponds to the best answer. Some items may contain several possible answers for your teacher education program or your plans as a teacher. For those items, please mark all that apply by completely filling in the appropriate boxes

BACKGROUND INFORMATION

1. Gender: ☐ Male ☐ Female

2. Age:
   ☐ 20 to 21 ☐ 24 to 25 ☐ 30 +
   ☐ 22 to 23 ☐ 26 to 29

3. a. Are you Hispanic or Latino? ☐ Yes ☐ No

   b. Which of the following best describes you? Please select one or more.
   - ☐ American Indian or Alaska Native
   - ☐ Asian
   - ☐ Black or African American
   - ☐ Native Hawaiian or Other Pacific Islander
   - ☐ White

4. What is your overall GPA?
   - ☐ 3.7-4.0 (A or 90-100)
   - ☐ 3.3-3.6 (B+ or 87-89)
   - ☐ 3.0-3.2 (B or 83-86)
   - ☐ 2.7-2.9 (B- or 80-82)
   - ☐ 2.3-2.6 (C+ or 77-79)
   - ☐ 2.0-2.2 (C or 73-76)
   - ☐ 1.7-1.9 (C- or 70-72)
   - ☐ 1.6 or below (D-F or 0-69)
   - ☐ I do not recall my overall GPA.

5. What is your GPA for only the courses taken in the field of Education?
   - ☐ 3.7-4.0 (A or 90-100)
   - ☐ 3.3-3.6 (B+ or 87-89)
   - ☐ 3.0-3.2 (B or 83-86)
   - ☐ 2.7-2.9 (B- or 80-82)
   - ☐ 2.3-2.6 (C+ or 77-79)
   - ☐ 2.0-2.2 (C or 73-76)
   - ☐ 1.7-1.9 (C- or 70-72)
   - ☐ 1.6 or below (D-F or 0-69)
   - ☐ I do not recall my Education GPA.

6. What was your combined SAT score (verbal and math)?
   - ☐ 1400 or higher
   - ☐ 1300-1390
   - ☐ 1200-1290
   - ☐ 1100-1190
   - ☐ 1000-1090
   - ☐ 900-990
   - ☐ 800-890
   - ☐ 790 or lower
   - ☐ I do not recall my score.
   - ☐ I did not take the SAT.

7. What was your ACT score?
   - ☐ 32+
   - ☐ 29-31
   - ☐ 26-28
   - ☐ 24-25
   - ☐ 21-23
   - ☐ 19-20
   - ☐ 16-18
   - ☐ 15 or lower
   - ☐ I do not recall my score.
   - ☐ I did not take the ACT.

8. What was your combined GRE score (verbal and quantitative)?

☐ 1400 or higher
☐ 1300-1399
☐ 1200-1299
☐ 1100-1199
☐ 1000-1099
☐ 900-999
☐ 800-899
☐ 799 or lower
☐ I do not recall my score.
☐ I did not take the GRE.

9. How many times have you taken each of the following Praxis tests?

a. Praxis I: Pre-Professional Skills Assessments (PPST)

☐ 0 ☐ 1 ☐ 2 ☐ 3+

b. Praxis II: Subject Assessments, Principles of Learning and Teaching (PLT) Tests and/or Teaching Foundations Tests

☐ 0 ☐ 1 ☐ 2 ☐ 3+

c. Praxis III: Classroom Performance Assessments

☐ 0 ☐ 1 ☐ 2 ☐ 3+

10. Have you passed any of the following Praxis tests?

a. Praxis I: Pre-Professional Skills Assessments (PPST)

☐ Have not taken the tests ☐ Yes ☐ No

b. Praxis II: Subject Assessments, Principles of Learning and Teaching (PLT) Tests and/or Teaching Foundations Test(s)

☐ Have not taken the tests ☐ Yes ☐ No

c. Praxis III: Classroom Performance Assessments

☐ Have not taken the tests ☐ Yes ☐ No

11. I took a different and/or additional test for certification or licensure.

☐ Yes ☐ No

12. In what semester will you graduate from your current program?

☐ Spring 2007
☐ Summer 2007
☐ Fall 2007
☐ Spring 2008 or later

13. a. Select the degree that you are currently working toward:

☐ Undergraduate (e.g., BA, BS, BSEd)
☐ Graduate (e.g., MA, MS, MEd)
☐ Post-Baccalaureate (Postbac) (e.g., 5th year program, non-masters)

b. Select the major or concentration that is closest to the one you are currently working toward:

☐ Early Childhood Education
☐ Elementary Education
☐ Combined Early Childhood/Elementary Education
☐ Combined Early Childhood/Special Education
☐ Combined Elementary/Special Education
☐ Combined Early Childhood/Elementary/Special Education
☐ Curriculum and Instruction
☐ Reading Education
☐ Multi/Interdisciplinary Studies
☐ Other

14. a. Do you have another college degree, other than the one you're working toward?

☐ Yes ☐ No

b. If you answered YES to question 14a, indicate which one/s below:

☐ AA/AS
☐ BA/BS
☐ MA/MS/MEd

c. If you answered YES to question 14a, is your other degree in Education?

☐ Yes ☐ No

15. Prior to entering your current degree program, were you already certified or endorsed in any of these areas? (Mark all that apply.)

☐ Not currently certified
☐ Reading specialist
☐ 7th grade & up
☐ Speech/language therapy
☐ ESL, ESOL, ELL, or LEP
☐ Elementary education
☐ Early childhood education
☐ Special education

16. If you are already a certified teacher, how many years of teaching experience do you have?

☐ Not currently certified
☐ Less than 1 school year
☐ 1
☐ 2
☐ 3
☐ 4 or more

17. How many "methods for teaching reading" and/or "foundations of reading development" courses, not counting courses such as children's literature or teaching language arts, have you completed? Please count completed courses *and* those you are currently taking.

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4 or more

18. Have you had any courses on the following topics? (Mark all that apply.)

☐ Teaching early reading, emergent literacy, teaching reading in grades PreK-3 or K-3
☐ Teaching reading in the middle grades, teaching intermediate reading, teaching reading in grades 3-6
☐ Teaching reading in grades K-6 or 1-6
☐ Teaching reading in the content areas
☐ Children's literature
☐ Assessment, diagnosis, and/or evaluation of children's reading
☐ Remediation of children's reading problems
☐ Teaching reading using technology
☐ Foundations of oral language, oral language development, linguistics
☐ Teaching language arts
☐ Teaching writing
☐ Teaching reading to English language learners

19. In how many courses did you have assignments that you completed in an elementary classroom (e.g., observing a class, teaching a lesson, leading a small group activity, etc.)?

☐ 0
☐ 1
☐ 2
☐ 3
☐ 4 or more

20. Before you started your final student teaching placement, in what grade level(s) did you complete the following field experiences? (Mark all that apply.)

	Grade Levels				
	PreK	K - 1	2 - 3	4 - 6	Not Required
a. Observe in classrooms	☐	☐	☐	☐	☐
b. Tutor individual students	☐	☐	☐	☐	☐
c. Teach small groups of students	☐	☐	☐	☐	☐
d. Teach the whole class for part of a day	☐	☐	☐	☐	☐

21. a. Will you complete a student teaching placement as part of your program?

☐ Yes ☐ No

b. If you answered YES to question 21a, how many weeks long is your student teaching experience?

☐ 5 weeks or less
☐ 6 to 8 weeks
☐ 9 to 12 weeks
☐ 13 to 16 weeks
☐ 17 weeks or more

c. If you answered YES to question 21a, in what grades did you (or will you) complete your student teaching placement? (Mark all that apply.)

☐ K
☐ 1
☐ 2
☐ 3
☐ 4
☐ 5
☐ 6
☐ 7th grade or higher
☐ Other (e.g., librarian, reading specialist)

d. If you answered YES to question 21a, during which semester did you (or will you) finish your student teaching experience? (Mark all that apply.)

☐ Fall 2005 semester or earlier
☐ Spring 2006 semester
☐ Summer 2006 semester
☐ Fall 2006 semester
☐ Spring 2007 semester
☐ Summer 2007 semester
☐ Fall 2007 semester

22. a. Do you intend to work as an elementary school teacher next year?

☐ Yes ☐ No

b. If you answered YES to question 22a, what grade would you prefer to teach?

☐ K-1
☐ 2-3
☐ 4-6
☐ 7th grade or higher
☐ Other specialty area (e.g., librarian, reading specialist)

EXPOSURE TO AND EMPHASIS ON EARLY READING CONCEPTS

There are many components of learning to read and a variety of strategies for teaching reading. We would like to find out what you have learned about teaching reading from your coursework and your field experiences.

First, please think about your coursework and field experiences *in general.* Then, please read the following questions and mark the most appropriate square.

1. Have you learned about what students must *know and be able to do* in order to:

   a. Focus on and manipulate phonemes in spoken words?

   ☐ Yes ☐ No

   b. Associate letters and the sounds they make to identify words?

   ☐ Yes ☐ No

   c. Read orally with appropriate speed, accuracy, and expression?

   ☐ Yes ☐ No

   d. Understand the meanings of words and learn new words?

   ☐ Yes ☐ No

   e. Understand what they read?

   ☐ Yes ☐ No

COURSEWORK

Next think about courses you took in your current degree program that focused specifically on *reading and literacy*. Please rate the degree of emphasis that your program placed on the strategies listed below. Keep in mind that you will have the opportunity to rate the emphasis on these strategies in your Field Experiences next. Use the following scale to rate the emphasis in your coursework:

None	This was not addressed in any of my courses.
Little	This was addressed briefly in one course.
Moderate	This was addressed over several class periods in one or two of my courses.
Considerable	I took a course entirely devoted to this topic.

◆	None	Little	Moderate	Considerable
1. Teaching children how to isolate, identify, separate, and blend sounds in spoken words	☐	☐	☐	☐
2. Teaching children to use phonics skills to figure out how to pronounce unfamiliar words	☐	☐	☐	☐
3. Teaching children to monitor how well they understand what they read and to correct problems as they occur	☐	☐	☐	☐
4. Using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	☐	☐	☐	☐
5. Identifying the words in a text that your children do not know and using their background knowledge to help them figure out the words' meanings	☐	☐	☐	☐
6. Making instructional decisions based on evaluations of children's oral reading fluency	☐	☐	☐	☐
7. Teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the main ideas	☐	☐	☐	☐
8. Teaching phonics to children in a systematic way, with a series of skills and activities	☐	☐	☐	☐
9. Teaching children to recognize and name letters	☐	☐	☐	☐
10. Having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	☐	☐	☐	☐
11. Teaching reading with *both* fiction and nonfiction reading materials	☐	☐	☐	☐
12. Relationships between elements of reading and oral language	☐	☐	☐	☐
13. Relationships among elements of reading or different types of reading skills	☐	☐	☐	☐
14. Examined materials and/or participated in class discussions about using *core reading programs* (or basals), such as Harcourt Brace, Houghton Mifflin, McGraw Hill, Open Court, Scott Foresman, or SRA Reading Mastery	☐	☐	☐	☐
15. Examined materials and/or participated in class discussions about using *literature-based programs*, such as Fountas and Pinnell's Guided Reading, Rigby materials, Scholastic Guided Reading, or the Wright Group materials	☐	☐	☐	☐
16. Examined materials and/or participated in class discussions about using *supplemental programs*, such as Corrective Reading, Great Leaps, LiPS, Saxon Phonics, or Voyager	☐	☐	☐	☐
17. Examined materials and/or participated in class discussions about using *school-wide literacy models*, such as First Steps, Literacy Collaborative, or Success for All	☐	☐	☐	☐

FIELD EXPERIENCE

Now, think about the various experiences you had in elementary classrooms during your current degree program. These experiences may be times in which you observed, did a practicum, or did your student teaching. Please try not to focus on classroom experiences you have had outside your current degree program. Use the following scale to rate the emphasis in your field experience(s):

None	This was not addressed in any of my courses.
Little	This was addressed briefly in one course.
Moderate	This was addressed over several class periods in one or two of my courses.
Considerable	I took a course entirely devoted to this topic.

◆

	None	Little	Moderate	Considerable
1. Teaching children how to isolate, identify, separate, and blend sounds in spoken words	☐	☐	☐	☐
2. Teaching children to use phonics skills to figure out how to pronounce unfamiliar words	☐	☐	☐	☐
3. Teaching children to monitor how well they understand what they read and to correct problems as they occur	☐	☐	☐	☐
4. Using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	☐	☐	☐	☐
5. Identifying the words in a text that your children do not know and using their background knowledge to help them figure out the words' meanings	☐	☐	☐	☐
6. Making instructional decisions based on evaluations of children's oral reading fluency	☐	☐	☐	☐
7. Teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the main ideas	☐	☐	☐	☐
8. Teaching phonics to children in a systematic way, with a series of skills and activities	☐	☐	☐	☐
9. Teaching children to recognize and name letters	☐	☐	☐	☐
10. Having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	☐	☐	☐	☐
11. Teaching reading with *both* fiction and nonfiction reading materials	☐	☐	☐	☐
12. Relationships between elements of reading and oral language	☐	☐	☐	☐
13. Relationships among elements of reading or different types of reading skills	☐	☐	☐	☐
14. Examined materials and/or participated in class discussions about using *core reading programs* (or basals), such as Harcourt Brace, Houghton Mifflin, McGraw Hill, Open Court, Scott Foresman, or SRA Reading Mastery	☐	☐	☐	☐
15. Examined materials and/or participated in class discussions about using *literature-based programs*, such as Fountas and Pinnell's Guided Reading, Rigby materials, Scholastic Guided Reading, or the Wright Group materials	☐	☐	☐	☐
16. Examined materials and/or participated in class discussions about using *supplemental programs*, such as Corrective Reading, Great Leaps, LiPS, Saxon Phonics, or Voyager	☐	☐	☐	☐
17. Examined materials and/or participated in class discussions about using *school-wide literacy models*, such as First Steps, Literacy Collaborative, or Success for All	☐	☐	☐	☐

FEELINGS OF PREPAREDNESS

New teachers enter their own classrooms for the first time feeling prepared about their abilities to teach in certain areas and less prepared in others. Use the following scale to rate your feelings of preparedness:

Not at all prepared	I do not know about or do not understand these activities well enough to use them with students.
Somewhat prepared	I am not completely sure how to use these activities with students in all grades and at all reading levels.
Mostly prepared	I understand how to use these activities well with some students but still need to deepen my understanding of the activities.
Definitely prepared	I completely understand how to use these activities with students at all grades and at all reading levels.

◆

	Not at all prepared	Somewhat prepared	Mostly prepared	Definitely prepared
1. Teaching children how to isolate, identify, separate, and/or blend sounds in spoken words	☐	☐	☐	☐
2. Teaching children to use phonics skills to figure out how to pronounce unfamiliar words	☐	☐	☐	☐
3. Teaching children to monitor how well they understand what they read and to correct problems as they occur	☐	☐	☐	☐
4. Using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	☐	☐	☐	☐
5. Identifying the words in a text that children do not know and using their background knowledge to help them figure out the words' meanings	☐	☐	☐	☐
6. Making instructional decisions based on evaluations of children's oral reading fluency	☐	☐	☐	☐
7. Teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the main ideas	☐	☐	☐	☐
8. Teaching phonics to children in a systematic way, with a series of skills and activities	☐	☐	☐	☐
9. Teaching children to recognize and name letters	☐	☐	☐	☐
10. Having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	☐	☐	☐	☐
11. Teaching reading with *both* fiction and nonfiction reading materials	☐	☐	☐	☐
12. How prepared do you feel to teach Kindergartners and 1st-graders the essential skills of reading?	☐	☐	☐	☐
13. How prepared do you feel to teach 2nd and 3rd-graders the essential skills of reading?	☐	☐	☐	☐

APPENDIX B.
PROGRAM SURVEY VARIABLES AND ASSESSMENT ITEMS

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item

Background characteristic information on pre-service teachers	Item number
Demographics	
Respondent gender	Q01
Respondent age	Q02
Respondent race/ethnicity: Hispanic	Q03a
Respondent race/ethnicity: American Indian	Q03a_1
Respondent race/ethnicity: Asian	Q03a_2
Respondent race/ethnicity: African American	Q03a_3
Respondent race/ethnicity: Pacific Islander	Q03a_4
Respondent race/ethnicity: White	Q03a_5
College performance: GPA and entrance examination scores	
Overall GPA	Q04
Education GPA	Q05
Combined SAT score	Q06
ACT score	Q07
GRE score	Q08
Teacher certification test experience	
Times taken Praxis test: Pre-Professional Skills Assessment	Q09a
Times taken Praxis test: Subject Assessments	Q09b
Times taken Praxis test: Classroom Performance Assessments	Q09c
Passed Praxis test: Pre-Professional Skills Assessment	Q10a
Passed Praxis test: Subject Assessments	Q10b
Passed Praxis test: Classroom Performance Assessments	Q10c
Taken other certification test	Q11
Expected graduation date, degree level, and area of concentration	
Expected graduation semester	Q12
Expected degree level	Q13a
Major or concentration: early childhood education	Q13b_01
Major or concentration: elementary education	Q13b_02
Major or concentration: combined early childhood/elementary education	Q13b_03
Major or concentration: combined early childhood/special education	Q13b_04
Major or concentration: combined elementary/special education	Q13b_05
Major or concentration: combined early childhood/elementary/special education	Q13b_06
Major or concentration: curriculum and instruction	Q13b_07
Major or concentration: reading education	Q13b_08
Major or concentration: multi/interdisciplinary studies	Q13b_09
Major or concentration: other	Q13b_10

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Previous degree	
Other college degree	Q14a
Previous degree	Q14b
Previous degree in education	Q14c
Prior certification area, years of teaching experience	
Prior certification area: not currently certified	Q15_1
Prior certification area: reading specialist	Q15_2
Prior certification area: 7th grade and up	Q15_3
Prior certification area: speech/language therapy	Q15_4
Prior certification area: ESL, ESOL, ELL, or LEP	Q15_5
Prior certification area: elementary education	Q15_6
Prior certification area: early childhood education	Q15_7
Prior certification area: special education	Q15_8
Years of teaching experience	Q16
Course experience in literacy and reading	
Number of teaching reading courses taken	Q17
Courses in teaching early reading, emergent literacy, teaching reading in grades PreK–3 or K–3	Q18_01
Courses in teaching reading in the middle grades, teaching intermediate reading, teaching reading in grades 3-6	Q18_02
Courses in teaching reading in grades K–6 or 1–6	Q18_03
Courses in teaching reading in the content areas	Q18_04
Courses in children's literature	Q18_05
Courses in assessment, diagnosis, and/or evaluation of children's reading	Q18_06
Courses in remediation of children's reading problems	Q18_07
Courses in teaching reading using technology	Q18_08
Courses in foundations of oral language, oral language development, linguistics	Q18_09
Courses in teaching language arts	Q18_10
Courses in teaching writing	Q18_11
Courses in teaching reading to English language learners	Q18_12

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Fieldwork experience	
Number of courses taken in an elementary classroom	Q19
Grades pre-K: observe in classrooms	Q20a_1
Grades K-1: observe in classrooms	Q20a_2
Grades 2-3: observe in classrooms	Q20a_3
Grades 4-6: observe in classrooms	Q20a_4
Not required: observe in classrooms	Q20a_5
Grades pre-K: tutor individual students	Q20b_1
Grades K-1: tutor individual students	Q20b_2
Grades 2-3: tutor individual students	Q20b_3
Grades 4-6: tutor individual students	Q20b_4
Not required: tutor individual students	Q20b_5
Grades pre-K: teach small groups of students	Q20c_1
Grades K-1: teach small groups of students	Q20c_2
Grades 2-3: teach small groups of students	Q20c_3
Grades 4-6: teach small groups of students	Q20c_4
Not required: teach small groups of students	Q20c_5
Grades pre-K: teach the whole class for part of a day	Q20d_1
Grades K-1: teach the whole class for part of a day	Q20d_2
Grades 2-3: teach the whole class for part of a day	Q20d_3
Grades 4-6: teach the whole class for part of a day	Q20d_4
Not required: teach the whole class for part of a day	Q20d_5
Student teaching experience	
Complete student teaching	Q21a
Length of student teaching	Q21b
Student teaching in grade K	Q21c_1
Student teaching in grade 1	Q21c_2
Student teaching in grade 2	Q21c_3
Student teaching in grade 3	Q21c_4
Student teaching in grade 4	Q21c_5
Student teaching in grade 5	Q21c_6
Student teaching in grade 6	Q21c_7
Student teaching in grade 7 or higher	Q21c_8
Student teaching in other position (e.g., librarian)	Q21c_9
Final semester of student teaching: fall 2005 semester or earlier	Q21d_1
Final semester of student teaching: spring 2006 semester	Q21d_2
Final semester of student teaching: summer 2006 semester	Q21d_3
Final semester of student teaching: fall 2006 semester	Q21d_4
Final semester of student teaching: spring 2007 semester	Q21d_5
Final semester of student teaching: summer 2007 semester	Q21d_6
Final semester of student teaching: fall 2007 semester	Q21d_7

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Intent to teach in next year and preferred grade level	
Plan on teaching in an elementary school next year	Q22a
Preferred grade: K-1	Q22b_01
Preferred grade: 2-3	Q22b_02
Preferred grade: 4-6	Q22b_03
Preferred grade: 7 or higher	Q22b_04
Preferred grade: other specialty area	Q22b_05
Teacher perception of overall exposure to early reading components	
Exposure: phonemic awareness: focus on and manipulate phonemes in spoken words?	Overall exposure item Q01a
Exposure: phonics: associate letters and the sounds they make to identify words?	Overall exposure item Q01b
Exposure: fluency: read orally with appropriate speed, accuracy, and expression?	Overall exposure item Q01c
Exposure: vocabulary: understand the meanings of words and learn new words?	Overall exposure item Q01d
Exposure: comprehension: understand what they read?	Overall exposure item Q01e

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Teacher perception of exposure to early reading components in coursework	
Coursework: teaching children how to isolate, identify, separate, and blend sounds in spoken words	Coursework exposure item Q01
Coursework: teaching children to use phonics skills to figure out how to pronounce unfamiliar words	Coursework exposure item Q02
Coursework: teaching children to monitor how well they understand what they read and to correct problems as they occur	Coursework exposure item Q03
Coursework: using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	Coursework exposure item Q04
Coursework: identifying the words in a text that your children do not know and using their background knowledge to help them figure out the words' meanings	Coursework exposure item Q05
Coursework: making instructional decisions based on evaluations of children's oral reading fluency	Coursework exposure item Q06
Coursework: teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the main ideas	Coursework exposure item Q07
Coursework: teaching phonics to children in a systematic way, with a series of skills and activities	Coursework exposure item Q08
Coursework: teaching children to recognize and name letters	Coursework exposure item Q09
Coursework: having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	Coursework exposure item Q10
Coursework: teaching reading with both fiction and nonfiction reading materials	Coursework Exposure item Q11
Coursework: relationships between elements of reading and oral language	Coursework exposure item Q12
Coursework: relationships among elements of reading or different types of reading skills	Coursework exposure item Q13
Coursework: examined materials and/or participated in class discussions about using core reading programs (or basals), such as Harcourt Brace, Houghton Mifflin, McGraw Hill, Open Court, Scott Foresman, or SRA Reading Mastery	Coursework exposure item Q14
Coursework: examined materials and/or participated in class discussions about using literature-based programs, such as Fountas and Pinnell's Guided Reading, Rigby materials, Scholastic Guided Reading, or the Wright Group materials	Coursework exposure item Q15
Coursework: examined materials and/or participated in class discussions about using supplemental programs, such as Corrective Reading, Great Leaps, LiPS, Saxon Phonics, or Voyager	Coursework exposure item Q17

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Teacher perception of exposure to early reading components in fieldwork experiences	
Field experience: teaching children how to isolate, identify, separate, and blend sounds in spoken words	Fieldwork exposure item Q01
Field experience: teaching children to use phonics skills to figure out how to pronounce unfamiliar words	Fieldwork exposure item Q02
Field experience: teaching children to monitor how well they understand what they read and to correct problems as they occur	Fieldwork exposure item Q03
Field experience: using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	Fieldwork exposure item Q04
Field experience: identifying the words in a text that your children do not know and using their background knowledge to help them figure out the words' meanings	Fieldwork exposure item Q05
Field experience: making instructional decisions based on evaluations of children's oral reading fluency	Fieldwork exposure item Q06
Field experience: teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the main ideas	Fieldwork exposure item Q07
Field experience: teaching phonics to children in a systematic way, with a series of skills and activities	Fieldwork exposure item Q08
Field experience: teaching children to recognize and name letters	Fieldwork exposure item Q09
Field experience: having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	Fieldwork exposure item Q10
Field experience: teaching reading with both fiction and nonfiction reading materials	Fieldwork exposure item Q11
Field experience: relationships between elements of reading and oral language	Fieldwork exposure item Q12
Field experience: relationships among elements of reading or different types of reading skills	Fieldwork exposure item Q13
Field experience: examined materials and/or participated in class discussions about using core reading programs (or basals), such as Harcourt Brace, Houghton Mifflin, McGraw Hill, Open Court, Scott Foresman, or SRA Reading Mastery	Fieldwork exposure item Q14
Field experience: examined materials and/or participated in class discussions about using literature-based programs, such as Fountas and Pinnell's Guided Reading, Rigby materials, Scholastic Guided Reading, or the Wright Group materials	Fieldwork exposure item Q15
Field experience: examined materials and/or participated in class discussions about using supplemental programs, such as Corrective Reading, Great Leaps, LiPS, Saxon Phonics, or Voyager	Fieldwork exposure item Q16
Field experience: examined materials and/or participated in class discussions about using school-wide literacy models, such as First Steps, Literacy Collaborative, or Success for All	Fieldwork exposure item Q17

See note at end of table.

Table B-1. Variables measured by the Pre-Service Teacher Program Survey, by section of the survey and item—Continued

Variable description	Item number
Teacher feelings of preparedness to teach early reading components	
Feelings of preparedness: teaching children how to isolate, identify, separate, and/or blend sounds in spoken words	Preparedness to teach item Q01
Feelings of preparedness: teaching children to use phonics skills to figure out how to pronounce unfamiliar words	Preparedness to teach item Q02
Feelings of preparedness: teaching children to monitor how well they understand what they read and to correct problems as they occur	Preparedness to teach item Q03
Feelings of preparedness: using a variety of methods to teach children the meanings of words, including direct and indirect (conversational) instruction, and multiple exposures and repetition	Preparedness to teach item Q04
Feelings of preparedness: identifying the words in a text that children do not know and using their background knowledge to help them figure out the words' meanings	Preparedness to teach item Q05
Feelings of preparedness: making instructional decisions based on evaluations of children's oral reading fluency	Preparedness to teach item Q06
Feelings of preparedness: teaching children a variety of strategies for understanding the text they read, such as using graphic organizers, making predictions, asking questions, and identifying the	Preparedness to teach item Q07
Feelings of preparedness: teaching phonics to children in a systematic way, with a series of skills and activities	Preparedness to teach item Q08
Feelings of preparedness: teaching children to recognize and name letters	Preparedness to teach item Q09
Feelings of preparedness: having children repeatedly read the same text aloud to improve their speed, accuracy, and expression	Preparedness to teach item Q10
Feelings of preparedness: teaching reading with both fiction and nonfiction reading materials	Preparedness to teach item Q11
Feelings of preparedness: how prepared do you feel to teach Kindergartners and 1st-graders the essential skills of reading?	Preparedness to teach item Q12
Feelings of preparedness: how prepared do you feel to teach 2nd and 3rd-graders the essential skills of reading?	Preparedness to teach item Q13

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table B-2. Variables measured by the Pre-Service Teacher Knowledge Assessment, by scale

Variables	Item number
Alphabetics	
Assessment items: phonemic awareness	Assessment items 1 to 12
Assessment items: phonics	Assessment items 13 to 20
Fluency	
Assessment items: fluency	Assessment items 21 to 32
Meaning	
Assessment items: vocabulary	Assessment items 33 to 44
Assessment items: comprehension	Assessment items 45 to 56

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX C.
PROGRAM SURVEY AND KNOWLEDGE ASSESSMENT PILOT TESTING

Pilot Testing of Program Survey Items

Four focus groups were held with students who were nearing completion of pre-service teacher education programs in or near Washington, DC. Participants' responses to prototype items served as a check on overall comprehensibility, familiarity of language, and understanding of concepts presented in the items. Their comments suggested points where background characteristic and survey items needed improvement and confirmed the study team's assumptions about overall course content, availability of field experiences, and other aspects of pre-service teacher training programs.

During two rounds of cognitive laboratory interviews, individual pre-service teachers from programs in the Washington, DC, area were asked to "think aloud" about working drafts of the survey and to provide running commentaries on their interpretations of the items, the rationales for their answers, challenges encountered in trying to remember information, and aspects of their programs that they thought should be included on the survey. The study team analyzed feedback and modified survey items based upon students' input after each round of interviews.

The Program Survey was also reviewed at the Institute of Education Sciences (IES), by members of the study's Technical Working Group, and by the Office of Management and Budget (OMB). Revisions were made to the Program Survey drafts based on comments received from these sources.

Pilot Testing of Knowledge Assessment

The Pre-service Teacher Knowledge Assessment was subjected to pilot testing in spring and summer 2006 with students in teacher training programs in one western state and in the mid-Atlantic area. Although the pilot test was small, results provided data that were used to evaluate item performance, guide item revision, and assemble the final form of the Knowledge Assessment. In total, 106 items were pilot tested in two test forms.

Participants were pre-service teachers recruited from a university in a western state (n = 73) and from six programs in the mid-Atlantic area (n = 69), for a total of 142 students. Participants were both undergraduate and master's level students; the majority were majoring or concentrating in elementary education. Participants were compensated for participation.

Two forms of the knowledge assessment were created for the pilot test and assembled for administration at the pilot test sites. Proctors at all test sites followed similar testing protocols and administered the tests in standardized fashion, with booklets distributed to participants so that no two participants with the same test form were seated next to one another. Participants were encouraged to try their best on the assessment and to answer every question with their best

guess when they were uncertain. Participants were given two hours to complete the Knowledge Assessment and a brief background characteristic questionnaire, but most finished their tasks within an hour and 15 minutes.

Analysis of Pilot Test Results and Assembly of Final Knowledge Assessment

Across Forms A and B, 106 items were pilot tested. Using an approach based on classical test theory (CTT),[36] the study team evaluated the difficulty and discrimination of each item and used this information to improve the reliability of the assessment. At this stage, the study team analyzed data with a five-factor scale to represent each of the five essential components of early reading instruction. Two considerations determined the final set of items for the assessment. The first was statistical, that is, the degree of relations among the items in each of the potential five scales (i.e., internal consistency across items asking about the five essential components of early reading instruction). The second consideration was the content of the items (i.e., content validity).

The initial plan was to include approximately 60 items on the final pre-service teacher knowledge assessment, with approximately 12 items for each of the five essential components of early reading instruction. The pilot test confirmed that pre-service teachers could complete approximately one multiple-choice item per minute. After analysis of pilot test results, the final form of the knowledge assessment consists of 56 items. Selection of items for the final form was based on two factors: the amount of overlap between each item and the other items in the same subscale (to increase internal reliability) and the content of the item (to ensure adequate content coverage). Table C-1 presents the scale statistics from the pilot test, showing changes to the assessment after the analyses.

Table C-1. Summary of scale statistics on five factors, based on pilot test and assembly of final form of Pre-Service Teacher Knowledge Assessment

	Pilot test: # of items	Initial alpha	Final # of items	Final alpha	Mean % correct (final set)	SD % correct (final set)
Phonemic awareness	20	.495	12	.516	66.7	.170
Phonics	19	.254	8	.408	55.4	.202
Fluency	22	.523	12	.592	72.6	.166
Vocabulary	24	.304	12	.579	72.3	.177
Comprehension	21	.450	12	.440	60.1	.181

NOTE: N of teachers = 142.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

The initial (before revisions) internal consistency (alpha) estimates ranged from .254 to .523. After the revisions, alpha ranged from .408 to .579. This level of internal reliability (after revisions) is lower than is usually recommended for research settings.[37] Because the items were

[36] The study team's initial plan had been to conduct item response theory (IRT) analyses as a complement to the CTT. However, because there were fewer than 200 participants in the pilot test, IRT was not used.
[37] Nunnally (1978, p. 245) recommends that instruments used in basic research have reliability of about .70 or better; that instruments used for applied situations, reliability should be .80 or higher; and that for high-stakes decision making, reliabilities should at least .90, preferably .95 or better.

designed to be content valid measures of the five essential components of early reading instruction (or five factors), the reliabilities may have been underestimated due to range restriction in the pilot test sample. The study team determined it would be imprudent to make substantial changes to the instrument to further improve the alpha coefficients. As a result, the study team determined the best course of action would be to revisit scale refinement with the main study sample.

Because the internal consistency estimates were somewhat low for the five factors, the study team also considered a higher-order set of factors. The study team investigated a solution to combine the factors to create three subscales—alphabetics, fluency, and meaning.[38] *Alphabetics* covers phonemic awareness and phonics. These items address concepts such as predictors of reading acquisition, phoneme manipulation skills, and letter-sound correspondence. For the current assessment, this simply is a combination of the 20 items from phonics and phonemic awareness. Subscale 2 items focus on *fluency* and address concepts such as oral fluency, repeated reading, and automaticity. This is the same construct as the five-factor solution. Subscale 3 items focus on *meaning*, or vocabulary and comprehension. These items address concepts such as effective comprehension skill/strategy instruction, types of vocabulary knowledge, and approaches to vocabulary development.

Scale analyses show that using the three subscales structure yields more internally consistent scores, with estimates ranging from .579 to .666 (refer to Table C-2). This increase in internal reliability was expected because the scales for alphabetics and meaning include more items (reliability is related to test length). These estimates are still below those recommended in research settings. However, the implications of a low level of reliability for the knowledge assessment are that the study research questions may have a low level of statistical power if reliabilities are not better in the main sample study (reliabilities based on the study sample are provided in Appendix I). Low reliability levels are expected during the early phases of scale construction in an emerging research area.

Table C-2. Summary of scale statistics, by three subscales

	Initial # of items	Initial alpha	Final # of items	Final alpha	Mean % correct	SD % correct
Alphabetics	39	.579	20	.601	62.2	.150
Fluency	22	.523	12	.592	72.6	.166
Meaning	45	.578	24	.666	66.3	.146

NOTE: N of teachers = 142.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Finally, the assessment can be scored based on a single overall factor (refer to Table C-3). Clearly, this is a simpler way of scoring the assessment, but this would cause a loss in any important distinction between knowledge of the different factors. As expected, using a single factor of 56 items results in higher internal reliability (alpha = .782).

[38] The Reading First legislation defined the five essential components of reading instruction. The subgroups that assembled the research base were divided among alphabetics (phonemic awareness and phonics), meaning (vocabulary and comprehension), and fluency.

Table C-3. Summary of scale statistics, by one overall factor

	Initial # of items	Initial alpha	Final # of items	Final alpha	Mean % correct	SD % Correct
Total	106	.777	56	.782	66.1	.120

NOTE: N of teachers = 142.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Final Assessment: Descriptive Statistics

For the final 56-item *Pre-Service Teacher Knowledge Assessment*, the mean percent correct of the items was 66.1 percent, with the percent correct in each subscale varying from 55.4 percent to 72.6 percent.

The study team also calculated the correlations between the scores on the five different subscales. As Table C-4 shows, these correlations ranged from about 0.18 to 0.44. This level of statistical overlap between subscales is expected because each subscale deals with (1) knowledge of similar constructs, (2) knowledge that was likely to be learned in the same place (e.g., college classes, books, experience in the classroom), and (3) learning that was affected by stable traits, like cognitive ability. Thus, these correlations are aligned with expectations and do not show overlap that would cast doubt on whether the scales are assessing pre-service teacher knowledge of distinct elements.

Table C-4. Correlation matrix of subscales

	Phonemic awareness	Phonics	Fluency	Vocabulary	Comprehension
Sample size	*142*	*142*	*142*	*142*	*142*
Phonemic awareness	—				
Phonics	.343*	—			
Fluency	.371*	.185*	—		
Vocabulary	.293*	.249*	.285*	—	
Comprehension	.352*	.374*	.302*	.445*	1

NOTE: N of teachers = 142. * $p < .05$.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Validity Analyses

In addition, the study team ran analyses of the relationships between subscales scores and some of the teacher background characteristics for the pilot test participants (see Table C-5). In the first of these analyses, the study team analyzed the correlations between the subscales scores and SAT/ACT scores that the participants received before entry into college. Because some schools and some states require the ACT and others require the SAT, most students had just one score or the other. A total of 77 participants had ACT scores. Results showed that subscales scores were not statistically significantly correlated with ACT scores. Also, 61 participants reported SAT scores; results showed that the five subscales scores correlated with SAT scores from .03 to .38 (including some statistically significant relationships). These analyses demonstrate that there was some statistical overlap between the subscales scores on the teacher knowledge assessment

and SAT scores, with higher scores on the SAT associated with higher scores on the Knowledge Assessment. Results are also presented for the three subscale model and the overall Knowledge Assessment score, although these are related to the findings for the five subscales and should not be interpreted as distinct results.

The study team also considered the relationships between the five subscale scores and various aspects of pre-service teachers' college academic work (see Table C-5). Correlations between GPA and the five subscale scores ranged from .08 to .33; in other words, pre-service teachers with higher GPAs tended to do better on the teacher Knowledge Assessment. However, having classroom teaching experience was not associated with higher scores on the assessment (r's range from .00 to –.10). Also, correlations between subscale scores and the "Bachelor's v. Master's" variables ranged from –.01 to –.23. It is important to note that the pilot included a small number of universities; thus, it is not possible to determine if the difference in scores for bachelor's and master's pre-service teachers is meaningful or if this is the result of differences between universities. Again, Table C-5 includes results for the three subscales and the overall assessment score, but these results should not be interpreted as distinct from those relating to the five subscales.

Table C-5. Correlations between subscales and other scores

	SAT	ACT	GPA	Bachelor's/ Master's[A]	Teaching exp[B]	Sample[C]
Sample size	*61*	*77*	*141*	*140*	*140*	*140*
Phonemic awareness	.38*	.17	.24*	-.12	.00	-.12
Phonics	.22	.12	.23*	-.06	-.10	-.06
Fluency	.03	.02	.10	-.23*	-.03	-.23*
Vocabulary	.32*	.01	.08	-.20	-.03	-.20*
Comprehension	.24	.16	.33*	-.01	-.10	.01
Alphabetics	.36*	.18	.28*	-.11	-.05	-.11
Fluency	.03	.02	.10	-.23*	-.03	-.23*
Meaning	.32*	.11	.25*	-.11	-.08	-.11
Overall	.34*	.15	.29*	-.18*	-.08	-.18*

[A]Bachelor's degree = 1; Master's degree = 2
[B] No teaching experience = 1; teaching experience = 2
[C]University in a Western state = 1; mid-Atlantic area = 2
* indicates correlations significant at $p < .05$
NOTE: N of teachers = 142.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Summary of Analyses

Based on the results of the pilot test and the analyses, the study team was able to reduce the total number of items in the teacher Knowledge Assessment from 106 items to 56 items. In general, the scales measuring knowledge of the five essential components of early reading instruction showed low to moderate internal consistencies (ranging from α=.408 to α=.592) and were

moderately correlated with each other (ranging from r = .185 to r = .445). Also, overall Knowledge Assessment scores were significantly correlated with both SAT scores and GPAs.

APPENDIX D.
RECRUITMENT

Introduction

This appendix provides details about the numbers of institutions contacted and reasons given for refusal to participate. It also provides information on the recruitment success rate.

Table D-1. Percentage of institution disposition, by sample group

	Survey group				
	Main sample	Replacement sample group 1	Replacement sample group 2	All institutions	Number of institutions
Participating institutions	52.9	60.4	58.3	55.3	99
Ineligible institution	3.4	0.0	0.0	2.2	4
Refusal	43.7	39.6	41.7	42.5	76
Grand total	100.0	100.0	100.0	100.0	179
Total number of all institutions	**119**	**48**	**12**	**179**	

SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table D-2. Number (percent) of institutions providing specific reasons for refusal to participate, by reason and sample group

Counts	Survey group			
Reason	Main sample (in percent)	Replacement sample group 1 (in percent)	Replacement sample group 2 (in percent)	All institutions (in percent)
Faculty did not want to burden students	2 (3.8)	3 (15.7)	0 (0.0)	5 (6.6)
Faculty not interested	9 (17.3)	1 (5.3)	0 (0.0)	10 (13.2)
Institution going through accreditation	4 (7.7)	0 (0.0)	0 (0.0)	4 (5.6)
Institution not interested	3 (5.8)	1 (5.3)	1 (20.0)	5 (5.6)
Institution revising curriculum	2 (3.8)	1 (5.3)	0 (0.0)	3 (4.2)
Institution too busy	9 (17.3)	4 (21.1)	0 (0.0)	13 (18.3)
IRB burden	1 (1.9)	1 (5.3)	0 (0.0)	2 (2.8)
No contact	3 (5.8)	1 (5.3)	2 (40.0)	6 (4.2)
No response	9 (17.3)	3 (15.7)	2 (40.0)	14 (16.9)
Not specified	3 (5.8)	3 (15.7)	0 (0.0)	6 (8.5)
Philosophical differences	7 (13.5)	1 (5.3)	0 (0.0)	8 (11.3)
Grand total	52 (100)	19 (100)	5 (100)	76 (100)

SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table D-3. Recruitment success rates for pre-service teacher recruitment

Components of rate calculation	Completed survey (C) [1]	Declined to participate (D)	Did not agree or decline to participate (U) [2]	Registered, but did not complete the survey (NS)	Total denominator (C + D + U + NS)	Response rate (C/(C + D + U + NS))[3]
Pre-service teachers from all sample groups	2237	236	283	49	2805	80%

[1] Eighty-seven pre-service teachers were identified as ineligible during registration and were removed from the eligible sample. As explained in Appendix F, the analytic sample was further reduced. After examining the survey responses, an additional 50 pre-service teachers were excluded from the analysis sample.
[2] Pre-service teachers who did not agree or decline participation did not convey their disposition to the recruiters, or the recruiters were unable to make contact with the pre-service teachers.
[3] The student response rate was 80 percent unweighted and 78 percent weighted.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX E.
DEMOGRAPHIC INFORMATION GATHERED THROUGH PROGRAM SURVEY

Introduction

The program survey used in the *Study of Teacher Preparation in Early Reading Instruction* gathered background characteristics about the pre-service teachers who agreed to participate in the study. The tables presented in this appendix provide details about the sample.

Table E-1. Demographic data about pre-service teacher sample from the Program Survey: 2007

Variable	Total	Percent
Gender		
Male	174	8.0
Female	2,011	92.0
Missing	‡	‡
Race/Ethnicity		
White	1,823	83.4
Black	92	4.2
Hispanic	156	7.1
Asian	67	3.1
Pacific Islander	7	0.3
American Indian	14	0.6
Multiple	21	1.0
Missing	7	0.3
Age		
20 to 21	386	17.6
22 to 23	887	40.6
24 to 25	295	13.5
26 to 29	242	11.1
30+	375	17.1
Missing	‡	‡
Total	**2,187**	

‡ Reporting standards not met; cell counts suppressed where n < 3.
NOTE: Black includes African American, Pacific Islander includes Native Hawaiian, Hispanic includes Latino, and American Indian includes Alaska Native. Race categories exclude Hispanic origin unless specified.
SOURCE: Study of Teacher Preparation in Early Reading Instruction, "Pre-Service Teacher Program Survey," 2007, U.S. Department of Education, National Center for Education Statistics.

Table E-2. Self-reports of educational achievement of pre-service teacher sample: 2007

Variable	Value	N	Percent
Overall GPA	1.7–1.9 (C– or 70–72)	‡	‡
	2.0–2.2 (C or 73–76)	‡	‡
	2.3–2.6 (C+ or 77–79)	13	0.6
	2.7–2.9 (B– or 80–82)	96	4.4
	3.0–3.2 (B or 83–86)	258	11.8
	3.3–3.6 (B+ or 87–89)	836	38.2
	3.7–4.0 (A or 90–100)	964	44.1
	I do not recall my overall GPA.	15	0.7
	Missing	‡	‡
Education GPA	2.3–2.6 (C+ or 77–79)	5	0.2
	2.7–2.9 (B– or 80–82)	18	0.8
	3.0–3.2 (B or 83–86)	81	3.7
	3.3–3.6 (B+ or 87–89)	470	21.5
	3.7–4.0 (A or 90–100)	1,570	71.8
	I do not recall my Education GPA.	35	1.6
	Missing/Unscorable	8	0.4
Combined SAT Score	790 or lower	15	0.7
	800–890	57	2.6
	900–990	191	8.7
	1000–1090	291	13.3
	1100–1190	322	14.7
	1200–1290	188	8.6
	1300–1390	68	3.1
	1400 or higher	16	0.7
	I do not recall my score.	409	18.7
	Missing/Unscorable	9	0.4
Combined ACT Score	15 or lower	6	0.3
	16–18	63	2.9
	19–20	82	3.7
	21–23	205	9.4
	24–25	111	5.1
	26–28	176	8.0
	29–31	67	3.1
	32+	11	0.5
	I do not recall my score.	386	17.6
	I did not take the ACT.	1,070	48.9
	Missing/Unscorable	10	0.5

See note at the end of table.

Table E-2. Self-reports of educational achievement of pre-service teacher sample: 2007—Continued

Variable	Value	N	Percent
Combined GRE Score	790 or lower	4	0.2
	800–890	8	0.4
	900–990	31	1.4
	1000–1090	33	1.5
	1100–1190	39	1.8
	1200–1290	35	1.6
	1300–1390	15	0.7
	1400 or higher	6	0.3
	I do not recall my score.	248	11.3
	I did not take the GRE.	1,757	80.3
	Missing/Unscorable	11	0.5
Total		**2,187**	

‡ Reporting standards not met; cell counts suppressed where n < 3.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table E-3. Self-reports of Praxis experiences from pre-service teacher sample: 2007

Variable	Value	N	Percent
Times taken the Praxis Test: Pre-Professional Skills Assessment			
	0	1,255	57.4
	1	742	33.9
	2	103	4.7
	3+	71	3.2
	Missing	16	0.7
Times taken the Praxis Test: Subject Assessments			
	0	1,132	51.8
	1	929	42.5
	2	78	3.6
	3+	30	1.4
	Missing	18	0.8
Times taken the Praxis Test: Classroom Performance Assessments			
	0	1,982	90.6
	1	137	6.3
	2	7	0.3
	3+	6	0.3
	Missing	55	2.5
Passed Praxis: Pre-Professional Skills Assessment			
	Yes	915	41.8
	No	23	1.1
	Have not taken the tests	1,199	54.8
	Missing/Unscorable	49	2.3
Passed Praxis: Subject Assessments			
	Yes	958	43.8
	No	64	2.9
	Have not taken the tests	1,085	49.6
	Missing/Unscorable	80	3.7
Passed Praxis: Classroom Performance Assessments			
	Yes	143	6.5
	No	49	2.2
	Have not taken the tests	1,930	88.2
	Missing/Unscorable	65	3.0
Taken Other Certification Tests			
	No	1,011	46.2
	Yes	1,162	53.1
	Missing	14	0.6
Total		**2,187**	

SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table E-4. Degree programs sought by pre-service teacher sample: 2007

Variable	Value	N	Percent
Expected Graduation Semester			
	Spring 2007	1,932	88.3
	Summer 2007	252	11.5
	Missing	‡	‡
Expected Degree Level			
	Undergraduate (e.g. BA BS BSEd)	1,580	72.2
	Graduate (e.g. MA MS MEd)	510	23.3
	Post-Baccalaureate (Postbac) (e.g. 5th year program non-masters)	93	4.3
	Missing	4	0.2
Total		**2,187**	
Major Concentration (% endorsing)			
	Early Childhood Education	331	15.1
	Elementary Education	1,379	63.1
	Combined Early Childhood/Elementary Education	211	9.6
	Combined Early Childhood/Special Education	31	1.4
	Combined Elementary/Special Education	161	7.4
	Combined Early Childhood/Elementary/Special Education	27	1.2
	Curriculum and Instruction	6	0.3
	Reading Education	5	0.2
	Multi/Interdisciplinary Studies	34	1.6
	Other	20	0.9

‡ Reporting standards not met; cell counts suppressed where n < 3.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table E-5. Previous teaching experience reported by pre-service teacher sample: 2007

Variable	Value	N	Percent
Other College Degree			
	No	1,384	63.3
	Yes	803	36.7
Previous Degree			
	AA/AS	228	10.4
	BA/BS	533	24.4
	MA/MS/MEd	39	1.8
	Missing	‡	‡
	No Previous Degree	1,384	63.3
Previous Degree in Education			
	No	613	28.0
	Yes	182	8.3
	Missing/Unscorable	8	0.4
	No previous degree	1,384	63.3
Prior Certification Area (% endorsing)			
	Not currently certified	2,028	92.7
	Reading Specialist	4	0.2
	7th Grade and up	5	0.2
	Speech/language therapy	‡	‡
	ESL, ESOL, ELL or LEP	9	0.4
	Elementary education	40	1.8
	Early childhood education	28	1.3
	Special education	‡	‡
Years of Teaching Experience			
	Not currently certified	1,973	90.2
	Less than 1 school year	136	6.2
	1 Year	8	0.4
	2 Years	6	0.3
	3 Years or more	12	0.5
	Missing	52	2.4
Total		**2,187**	

‡ Reporting standards not met; cell counts suppressed where n < 3.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table E-6. Reports of classroom experience from pre-service teacher sample

Variable	Value	N	Percent
Number of "methods" or "foundations" of early reading courses taken			
	0	166	7.6
	1	470	21.5
	2	806	36.9
	3	380	17.4
	4 or more	363	16.6
	Missing/Unscorable	‡	‡
Foci of reading-related courses taken			
	Teaching early reading, emergent literacy, teaching reading in grades PreK–3 or K–3	1,396	63.8
	Teaching reading in the middle grades, teaching intermediate reading, teaching reading in grades 3–6	708	32.4
	Teaching reading in grades K–6 or 1–6	1,313	60.0
	Teaching reading in the content areas	612	28.0
	Children's literature	1,679	76.8
	Assessment, diagnosis, and/or evaluation of children's reading	1,107	50.6
	Remediation of children's reading problems	418	19.1
	Remediation reading using technology	243	11.1
	Foundations of oral language, oral language development, linguistics	597	27.3
	Teaching language arts	1,169	53.5
	Teaching writing	804	36.8
	Teaching reading to English language learners	500	22.9
Number of courses requiring assignments in an elementary classroom			
	0	24	1.1
	1	89	4.1
	2	199	9.1
	3	270	12.3
	4 or more	1,604	73.3
	Missing	‡	‡
Grades for courses requiring observations, prior to student teaching (% endorsing)			
	Pre-K Grades	603	27.6
	Grades K–1	1,466	67.0
	Grades 2–3	1,488	68.0
	Grades 4–6	1,250	57.2
	Not required	89	4.1

See note at end of table.

Table E-6. Reports of classroom experience from pre-service teacher sample—Continued

Variable	Value	N	Percent
Grades for courses that require tutoring individual students (% endorsing)			
	Pre-K Grades	135	6.2
	Grades K–1	822	37.6
	Grades 2–3	941	43.0
	Grades 4–6	796	36.4
	Not required	398	18.2
Grades for courses that require teaching small groups of students (% endorsing)			
	Pre-K Grades	328	15
	Grades K–1	1,167	53.4
	Grades 2–3	1,226	56.1
	Grades 4–6	979	44.8
	Not required	145	6.6
Courses that require teaching the whole class for part of the day (% endorsing)			
	Pre-K Grades	237	10.8
	Grades K–1	941	43.0
	Grades 2–3	1,013	46.3
	Grades 4–6	919	42.0
	Not required	261	11.9
Must complete student teaching			
	No	31	1.4
	Yes	2,154	98.5
	Missing	‡	‡
Length of student teaching	5 weeks or less	5	0.2
	6 to 8 weeks	139	6.4
	9 to 12 weeks	297	13.6
	13 to 16 weeks	1,290	59.0
	17 weeks or more	422	19.3
	Missing/Unscorable	‡	‡
	Have not completed student teaching	31	1.4
Student teaching grades (% endorsing)	K	447	20.4
	1	592	27.1
	2	626	28.6
	3	531	24.3
	4	494	22.6
	5	427	19.5
	6	203	9.3
	7 or higher	129	5.9
	Other Position (e.g., Librarian)	76	3.5

See note at end of table.

Table E-6. Reports of classroom experience from pre-Service teacher sample—Continued

Variable	Value	N	Percent
Final Semester of Student Teaching	Fall 2005 or earlier	7	0.3
	Spring 2006 semester	62	2.8
	Summer 2006 semester	7	0.3
	Fall 2006 semester	308	14.1
	Spring 2007 semester	1,774	81.1
	Summer 2007 semester	39	1.8
	Fall 2007 semester	62	2.8
Plan to teach in an elementary school in fall 2007	No	204	9.3
	Yes	1,976	90.4
	Missing	7	0.3
Preferred grades to teach	K–1	619	28.3
	2–3	898	41.1
	4–6	587	26.8
	7 or Higher	51	2.3
	Other Specialty Area	32	1.5
Total		**2,187**	

‡ Reporting standards not met; cell counts and percentages suppressed where n < 3.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX F.
DATA CLEANING PROCEDURES

To ensure data accuracy, various data cleaning procedures were conducted at several stages of data processing. Prior to submission of data files to Optimal, Questar Data Services, a subcontractor, examined the data to detect deviant response ranges, anomalous response patterns, excessive missing data, extreme outliers, and highly skewed or irregular distributions. Their procedures included:

- *Confirmation of the accuracy of machine scoring*. As a first check of the data file provided by Questar, 5 percent of the cases from the data file were selected at random and manually compared with program survey booklets.
- *Resolution of illegible marks and multiple entries.* Questar software coded illegible marks or multiple marks with an asterisk (*) in the raw record. Project analysts compared the coded errors with original questionnaire in an attempt to resolve any possible discrepancy. *Identification of aberrant response patterns.* Data were also checked for "straight-lining" that could indicate that respondents did not read the individual items.

AIR and Optimal conducted additional data checking and cleaning, in that approximately 10 percent of the scannable booklets were selected in a purposive way and manually checked against the data set for accuracy. The booklets were selected purposively, rather than randomly, based on rules that AIR defined to identify unlikely data patterns:

- Indicated Hispanic/Latino but checked any racial category other than White.
- Inconsistencies across the self-reported Overall GPA, Education GPA, SAT, ACT, and GRE scores. This was based on the assumption that respondents with high GPAs also have high test scores and vice versa.
- Checked "I do not recall" for more than two of the GPA/Test score questions.
- Inconsistencies between the number of times taking Praxis and whether passed the Praxis tests.
- Inconsistencies in skip patterns between a question and its follow-up questions.
- Indicated "currently working toward undergraduate degree" but said "already certified or endorsed in certain program areas."
- Indicated "YES" to feeling like having learned something for each of the five components on page 5 of the survey, yet endorsed "none" for the items on pages 6 and 7 that corresponded to each of the five components.
- Indicated "YES" to feeling like having learned something for each of the five components on page 5 of the survey, yet scored at bottom 5 percent for the corresponding component in the knowledge assessment.
- High GPAs/SAT/ACT/GRE scores but scored at bottom 10 percent in knowledge assessment.

- Top quartile in feelings of preparedness for a component (i.e., endorsed "Definitely prepared"), but bottom 5 percent on the knowledge assessment for the corresponding component.
- Indicated "not currently certified" in background characteristic questions 15 or 16 but checked "YES" to the Praxis III tests in questions 9c and 10c.

AIR applied the rules and created flags for any data records with violations of such rules. Depending on the number of flags assigned to a given data record, test booklets were pulled for a visual comparison of actual entries and the scanned data.

About 90 percent of the data records had fewer than 6 flags. One record had the highest number of flags, 17. Data records that had more than 6 flags (about 10 percent of the total sample) were manually checked. In addition, AIR also checked:

- Records with more than seven consecutive runs of the same response values (may indicate respondents did not read the questions). Seven such records were found.
- Records with more than seven missing responses. Fourteen such records were found.

As a result of the data checking procedures, 50 respondents were excluded from the analysis sample. Among the 50 respondents, one had missing responses to all survey questions, one had 14 consecutive runs of the same response values, and 48 were deemed ineligible to the survey because they indicated they would graduate in fall 2007 or later (see Chapter 2 for recruiting criteria).

When AIR and Optimal found data discrepancies between the actual survey booklets and the data file, further investigation showed that these discrepancies resulted from scanning issues. For example, discrepancies were identified when respondents checked answer boxes instead of filling in as instructed or marking an answer and then not thoroughly erasing it. When visual inspection confirmed that an apparent discrepancy resulted from a scanning error, the data file was corrected.

Additionally, an initial visual inspection of the booklets by AIR staff revealed an error that was introduced at some point in the production process. As part of the program survey, pre-service teachers were asked to rate the extent to which their courses place emphasis on specific aspects of the essential components using a 4-point scale (None, Little, Moderate, Considerable). They were also asked to use the same scale to rate the extent to which they observed these aspects of early reading instruction as part of their field experience and student teaching. Specific descriptors of the emphasis ratings were developed to differentiate between coursework and field experiences/student teaching. However, during the production of the booklets, the descriptors for coursework was inserted for *both* coursework and field experience.

To determine the extent to which pre-service teachers' responses may have been skewed by the error, AIR investigated responses to the two sections of the program survey and found that differential patterns of responses existed for coursework and field experiences, suggesting that pre-service teachers were not disconcerted by the error and were able to respond appropriately.

Further review of omitted responses aiming to identify individuals who may have been confused by this response scale did not reflect any serious issue either.

APPENDIX G.
PSYCHOMETRIC ANALYSIS OF THE PROGRAM SURVEY SCALES MEASURING COURSEWORK EMPHASIS ON, FIELD EXPERIENCE EXPOSURE TO, AND FEELINGS OF PREPAREDNESS TO TEACH THE ESSENTIAL COMPONENTS

Introduction

This appendix describes the psychometric analyses used to examine the factor structure of the Program Survey scales measuring coursework emphasis on, field experience exposure to, and pre-service teachers' feelings of preparedness to teach the essential components. Because the items in the Program Survey require pre-service teachers to describe the extent to which they had been exposed to essential components of early reading instruction in their training, their responses represent, to some extent, group-level constructs. The study team used hierarchical linear modeling (HLM) to examine scale reliabilities at both the individual level and the institution level, based on alternative factor structures. Descriptive statistics and intra-class correlations for raw scores are also included.

Analysis of Program Survey Items

A major decision for the study team was the number of factors with which to report results. The study team conducted unweighted reliability analyses of the 17 coursework items, the 17 field experience items, and the 13 feelings of preparedness items on the Program Survey. The sections that follow will first present reliability results based on pre-service teacher-level analyses that ignore the institution membership of the pre-service teachers and then present results based on multivariate hierarchical linear models (HLM) that take into account the nested structure of the Program Survey data (i.e., items nested within teachers and teachers nested within institutions). Results are presented for scales corresponding to the following alternative factor models:

- Five-factor model: phonemic awareness, phonics, fluency, comprehension, and vocabulary
- Three-factor model: alphabetics, fluency, and meaning
- Two-factor model: word and meaning
- One-factor model: all five essential components of early reading instruction combined

In the three-factor model, alphabetics is formed by collapsing phonemic awareness and phonics, and meaning is formed by collapsing comprehension and vocabulary. In the two-factor model, the word scale is formed by collapsing alphabetics and fluency. The one-factor model collapses across all five scales.

Results from preliminary pre-service teacher-level reliability analyses

As a first step, the study team estimated reliability by using teacher-level analyses that ignored the nested data structure (and hence the dependence among pre-service teachers within the same institutions). Although this analysis is not ideal, it did provide simple reliability estimates that are often reported by other researchers, allowing limited (in the sense that the coefficient alpha reliabilities do not account for the clustering of the data) comparisons across studies. Table G-1 presents the Cronbach's alpha reliability estimates for scales measuring the essential components of reading instruction based on survey data on pre-service teachers' coursework, field experience, the sum of coursework and field experience (treating the coursework and field experience sections as one combined section instead of two), and feelings of preparedness to teach these components. The reliabilities for the five-factor model range between 0.507 and 0.860. Phonemic awareness and phonics have, on average, higher reliabilities across sections of the Program Survey, whereas fluency is lower, on average, than the rest of the scales.

Table G-1. Internal consistency (Cronbach's alpha) of Program Survey scales, by aspect of program and the components of early reading instruction

Essential Component	Coursework		Field Experience		Coursework and Field Experience		Feelings of Preparedness	
Five scales	n	α	n	α	n	α	n	α
Phonemic Awareness	2,187	0.826	2,184	0.860	2,184	0.747	2,183	0.809
Phonics	2,169	0.741	2,174	0.750	2,157	0.739	2,164	0.781
Fluency	2,165	0.664	2,167	0.637	2,167	0.722	2,173	0.507
Comprehension	2,180	0.767	2,177	0.737	2,149	0.819	2,178	0.688
Vocabulary	2,137	0.673	2,140	0.720	2,177	0.695	2,176	0.672
Three scales								
Alphabetics	2,169	0.849	2,171	0.850	2,157	0.844	2,161	0.862
Fluency	2,165	0.664	2,167	0.637	2,167	0.722	2,173	0.507
Meaning	2,131	0.806	2,140	0.787	2,089	0.850	2,168	0.795
Two scales								
Word	2,147	0.849	2,152	0.840	2,117	0.870	2,155	0.850
Meaning	2,131	0.806	2,140	0.787	2,089	0.850	2,168	0.795
One scale	2,112	0.887	2,124	0.878	2,062	0.913	2,131	0.880

NOTE: Program focus based on coursework and field experience data was measured on a 4-point scale in the Program Survey: 0 = none, 1 = little, 2 = moderate, and 3 = considerable. Feelings of preparedness was measured on a 4-point scale: 0 = not at all prepared, 1 = somewhat prepared, 2 = mostly prepared, and 3 = definitely prepared.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Collapsing the five scales into three scales raises the reliability of the combined factor above the higher of the two reliabilities of the scales from which the combined factor was formed. For example, the reliability for the meaning scale for coursework items (in the three-factor model) is higher (0.806) than the reliability for either comprehension (0.767) or vocabulary (0.673). The one exception to this finding is for alphabetics field experience items. This trend did not continue when the three scales were collapsed into two scales. The highest reliabilities corresponded to a one-factor model. In addition, combining coursework and field experience

does not always result in an increase in reliability. Since this suggests that coursework and field experience may represent slightly different facets of pre-service teachers' experiences, the report presents findings for coursework and field experience separately, in addition to combined.

Results from Multivariate HLM Analyses

As a further step to understand the reliability of the pre-service teachers' responses to the Program Survey items, the study team examined the responses using a multivariate hierarchical linear model (HLM). This model explicitly takes into account the nested data structure and allows the study team to assess scale reliabilities at both the pre-service teacher level and the institution level. From the results of these analyses, it may be clearer which aspects of program emphasis are experienced similarly by pre-service teachers within the same institutions and therefore should be treated as group-level constructs and which aspects of program emphasis experienced by pre-service teachers vary substantially, even among pre-service teachers within the same institutions, and therefore should be treated as individual-level constructs in subsequent analyses.

Following the methods proposed by Raudenbush, Rowan, and Kang (1991), the study team conducted three-level multivariate analyses that derive estimates of the scale scores as well as scale reliabilities at both the pre-service teacher level and the institution level. The reliability estimates from the multilevel analyses are presented in Tables G-2a and G-2b.

Table G-2a. Reliability estimates for Program Survey scales based on multivariate HLM analyses, by aspect of program and the components of early reading instruction

Essential Component	Coursework			Field Experience			Coursework and Field Experience		
		Teacher level	Institution level		Teacher level	Institution level		Teacher level	Institution level
Five scales	n	ICC	ICC	n	ICC	ICC	n	ICC	ICC
Phonemic Awareness	2,187	0.795	0.790	2,184	0.854	0.494	2,184	0.716	0.728
Phonics	2,169	0.693	0.792	2,174	0.732	0.628	2,157	0.700	0.770
Fluency	2,165	0.610	0.768	2,167	0.627	0.464	2,167	0.694	0.699
Comprehension	2,180	0.744	0.703	2,177	0.729	0.580	2,149	0.807	0.670
Vocabulary	2,137	0.649	0.634	2,140	0.715	0.321	2,177	0.677	0.560
Three scales									
Alphabetics	2,169	0.821	0.814	2,171	0.845	0.602	2,157	0.822	0.775
Fluency	2,165	0.610	0.768	2,167	0.626	0.466	2,167	0.693	0.700
Meaning	2,131	0.792	0.716	2,140	0.788	0.567	2,089	0.844	0.675
Two scales									
Word	2,147	0.817	0.816	2,152	0.829	0.585	2,117	0.849	0.767
Meaning	2,131	0.792	0.716	2,140	0.787	0.570	2,089	0.844	0.676
One scale	2,112	0.873	0.772	2,124	0.875	0.592	2,062	0.905	0.725

NOTE: N of institutions = 99.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table G-2b. Reliability estimates for program focus scales based on multivariate HLM analyses of Program Survey items related to feelings of preparedness

Essential Component	Feelings of Preparedness		
	n	Teacher level	Institution level
Five scales	2,183	0.790	
Phonemic Awareness	2,164	0.759	0.679
Phonics	2,173	0.486	0.682
Fluency	2,178	0.681	0.643
Comprehension	2,176	0.661	0.494
Vocabulary			0.508
Three scales	2,161	0.847	
Alphabetics	2,173	0.470	0.703
Fluency	2,168	0.788	0.651
Meaning			0.521
Two scales	2,155	0.828	
Word	2,168	0.787	0.707
Meaning	2,131	0.865	0.517
One scale	2,183	0.790	0.677

NOTE: N of institutions = 99.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

A number of findings in Table G-2a are noteworthy. First, moving from five scales to three, the two combined essential component scales (i.e., alphabetics and meaning) both have higher reliabilities than the individual scales (exceptions being for pre-service teacher–level phonemic awareness and institution-level phonics, both for field experience). For example, with regard to coursework, comprehension and vocabulary have a reliability of 0.744 and 0.649, respectively, at the pre-service teacher level when examined as two separate scales and have a reliability of 0.792 when analyzed as a combined scale (meaning). Moving to two scales does not uniformly increase reliability over three scales. Given that some of the five individual scales have relatively low reliabilities, the study reports findings for the more reliable collapsed scales under a one-factor framework. The report also provides findings under the three-factor framework to identify any differences among the essential components.

Another noticeable finding in Tables G-2a and G-2b is that there are differences in the reliability estimates between the teacher-level and institution-level. Whereas the institution-level scale reliabilities are consistent with the pre-service teacher–level reliabilities based on responses to coursework items, they are, on average, lower than pre-service teacher–level reliabilities across all scales based on responses to field experience items and across all the feelings of preparedness scales with the exception of fluency. One explanation for such differences is that they reflect differences in the intra-class correlations (ICC) among scales that refer to different aspects of pre-service teachers' experiences (experience referents). As Raudenbush, Rowan, and Kang (1991) note, the group-level reliability of a scale depends on four factors: the number of items making up the scale; the level of intercorrelation among these items at the individual level; the level of "intersubjective agreement" among individuals within the same group (i.e., the ICC); and the number of individuals sampled within the group. All else being equal, the higher the

level of agreement about the scale ("intersubjective agreement") among individuals within the same group, the more reliable is the group-level estimate of the scale based on individual-level data.

The Program Survey data indicate that the level of intersubjective agreement among pre-service teachers within the same institutions varies both across scales and experience referents. As shown in Table G-3, the ICCs are lower for field experience and feelings of preparedness scales than for coursework scales. This finding is consistent with the finding that the institution-level scales based on coursework items are, on average, more reliable than those based on field experience and feelings of preparedness. As an extreme example, the ICC for the vocabulary scale based on field experience items is only 0.03, which indicates that most of the variation in this measure is between pre-service teachers within institutions as opposed to between institutions (97 percent vs. 3 percent). In other words, pre-service teachers from the same institution vary in their perceptions of the focus on vocabulary in instruction they observed during their field experience. It is therefore not surprising that the institution-level measure of this scale contains a large amount of measurement error and thus has a low reliability (0.321, see Table G-2a).

Table G-3. Institution-level intra-class correlations for Program Survey scales based on coursework, field experience, and feelings of preparedness items

Essential Component	Coursework	Field Experience	Coursework and Field Experience	Feelings of Preparedness
Five scales				
Phonemic Awareness	0.18	0.05	0.15	0.11
Phonics	0.20	0.10	0.18	0.12
Fluency	0.20	0.06	0.14	0.15
Comprehension	0.13	0.08	0.11	0.06
Vocabulary	0.11	0.03	0.08	0.07
Three scales				
Alphabetics	0.20	0.08	0.16	0.12
Fluency	0.20	0.06	0.14	0.16
Meaning	0.13	0.07	0.10	0.06
Two scales				
Word	0.20	0.07	0.15	0.12
Meaning	0.13	0.07	0.10	0.06
One scale	0.07	0.03	0.12	0.05

NOTE: N of institutions = 99. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Next, the study team examined the correlations among scales at each level to determine if there was sufficient justification to collapse the five scales into three scales. The correlations among the five scales based on the HLM analyses seem to support collapsing the five-factor model into three scales. As shown in Tables G-4a, G-4b, G-4c, and G-4d, at the institution level, phonemic awareness and phonics are correlated at 0.899, 0.949, 0.920, and 0.883 based on data on coursework, field experience, combined coursework and field experience, and feelings of

preparedness, respectively. Comprehension and vocabulary likewise are correlated at levels ranging from 0.869 to 0.960. These findings suggest that these two pairs of components (phonemic awareness and phonics as well as comprehension and vocabulary) may be appropriately combined into single, more reliable scales.

Table G-4a. Institution-level correlations among five scales based on coursework items

Essential Component	Phonemic Awareness	Phonics	Fluency	Comprehension	Vocabulary
Phonemic Awareness	1.000				
Phonics	0.899	1.000			
Fluency	0.779	0.760	1.000		
Comprehension	0.536	0.570	0.770	1.000	
Vocabulary	0.663	0.666	0.861	0.869	1.000

Table G-4b. Institution-level correlations among five scales based on field experience items

Essential Component	Phonemic Awareness	Phonics	Fluency	Comprehension	Vocabulary
Phonemic Awareness	1.000				
Phonics	0.949	1.000			
Fluency	0.708	0.812	1.000		
Comprehension	0.718	0.705	0.744	1.000	
Vocabulary	0.709	0.767	0.759	0.960	1.000

Table G-4c. Institution-level correlations among five scales based on coursework and field experience items combined

Essential Component	Phonemic Awareness	Phonics	Fluency	Comprehension	Vocabulary
Phonemic Awareness	1.000				
Phonics	0.920	1.000			
Fluency	0.799	0.726	1.000		
Comprehension	0.573	0.617	0.741	1.000	
Vocabulary	0.726	0.638	0.824	0.885	1.000

Table G-4d. Institution-level correlations among five scales based on feelings of preparedness items

Essential Component	Phonemic Awareness	Phonics	Fluency	Comprehension	Vocabulary
Phonemic Awareness	1.000				
Phonics	0.883	1.000			
Fluency	0.587	0.691	1.000		
Comprehension	0.673	0.863	0.901	1.000	
Vocabulary	0.697	0.784	0.879	0.958	1.000

NOTE: N of institutions = 99. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

As shown in Tables G-5a through G-5d, the correlations among the three resulting scales (i.e., alphabetics, fluency, and meaning) range from 0.609 to 0.899. It is possible to further collapse the three scales into two scales. Although fluency is correlated with both alphabetics and meaning, there is no clear pattern among the correlations, suggesting that collapsing scales further might not provide any additional benefit.

Table G-5a. Institution-level correlations among three scales based on coursework items

Essential Component	Alphabetics	Fluency	Meaning
Alphabetics	1.000		
Fluency	0.791	1.000	
Meaning	0.609	0.811	1.000

Table G-5b. Institution-level correlations among three scales based on field experience items

Essential Component	Alphabetics	Fluency	Meaning
Alphabetics	1.000		
Fluency	0.769	1.000	
Meaning	0.732	0.746	1.000

Table G-5c. Institution-level correlations among three scales based on coursework and field experience items combined

Essential Component	Alphabetics	Fluency	Meaning
Alphabetics	1.000		
Fluency	0.778	1.000	
Meaning	0.641	0.772	1.000

Table G-5d. Institution-level correlations among three scales based on feelings of preparedness items

Essential Component	Alphabetics	Fluency	Meaning
Alphabetics	1.000		
Fluency	0.643	1.000	
Meaning	0.799	0.899	1.000

NOTE: N of institutions = 99. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

In summary, results from the reliability analyses support using a three-factor framework of early reading instruction (i.e., alphabetics, fluency, and meaning). The reliabilities of the scales differ across scales, across levels of data (teacher level vs. institution level), and across item referents (coursework, field experience, and feelings of preparedness). The institution-level scales, which conceptually are more appropriate measures of program focus on the components of early reading instruction than the pre-service teacher-level scales, seem to have appropriate reliability

for the coursework items. They are less reliable, however, for the field experience items and feelings of preparedness items, suggesting that there is less uniformity in pre-service teachers' field experience and perception of preparedness than in their coursework experience. As a result, the report analyses consider coursework a state-level construct because pre-service teacher programs are preparing students to meet certification requirements and testing mandates determined at the state level. The analyses consider field experience and feelings of preparedness individual-level constructs because field experiences may differ according to grade level, content of instruction observed or taught, and other factors such as the quality of the teachers who are observed. Further, pre-service teachers' feelings of preparedness to teach the essential components of early reading instruction will derive from distinct personal as well as experiential factors. The raw score frequencies for each of these scales is presented in Appendix H.

APPENDIX H.
DESCRIPTIVE STATISTICS FOR PROGRAM SURVEY ITEMS

This appendix presents item frequencies discussed in Appendix G. Tables H-1 through H-3 show frequencies for items for the coursework, field experience, and feelings of preparedness sections, respectively.

Table H-1. Coursework frequencies

Item	N, None	%, None	N, Little	%, Little	N, Moderate	%, Moderate	N, Considerable	%, Considerable
CRS_Q01	84	3.84	641	29.31	1112	50.85	350	16
CRS_Q02	73	3.34	535	24.46	1132	51.76	447	20.44
CRS_Q03	73	3.35	554	25.39	1131	51.83	424	19.43
CRS_Q04	57	2.61	463	21.21	1079	49.43	584	26.75
CRS_Q05	29	1.33	389	17.81	1205	55.17	561	25.69
CRS_Q06	76	3.48	554	25.39	991	45.42	561	25.71
CRS_Q07	12	0.55	151	6.91	975	44.64	1046	47.89
CRS_Q08	142	6.51	751	34.42	884	40.51	405	18.56
CRS_Q09	212	9.75	754	34.68	829	38.13	379	17.43
CRS_Q10	153	7.02	741	34.01	939	43.09	346	15.88
CRS_Q11	67	3.07	425	19.49	986	45.21	703	32.23
CRS_Q12	118	5.42	726	33.33	988	45.36	346	15.89
CRS_Q13	112	5.18	711	32.9	1011	46.78	327	15.13
CRS_Q14	353	16.16	807	36.95	711	32.55	313	14.33
CRS_Q15	522	23.93	737	33.79	546	25.03	376	17.24
CRS_Q16	1073	49.09	803	36.73	231	10.57	79	3.61
CRS_Q17	1052	48.15	777	35.56	268	12.27	88	4.03

NOTE: The Item column references the Program Survey items as presented in the Coursework section (page A-6).
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table H-2. Fieldwork frequencies

Item	N, None	%, None	N, Little	%, Little	N, Moderate	%, Moderate	N, Considerable	%, Considerable
FLD_Q01	109	4.98	506	23.14	872	39.87	700	32.01
FLD_Q02	82	3.75	375	17.17	936	42.86	791	36.22
FLD_Q03	39	1.79	370	16.94	1013	46.38	762	34.89
FLD_Q04	48	2.2	347	15.88	922	42.2	868	39.73
FLD_Q05	21	0.96	292	13.4	965	44.29	901	41.35
FLD_Q06	68	3.12	441	20.22	873	40.03	799	36.63
FLD_Q07	26	1.19	175	8	749	34.25	1237	56.56
FLD_Q08	210	9.62	525	24.04	801	36.68	648	29.67
FLD_Q09	301	13.83	488	22.42	633	29.08	755	34.68
FLD_Q10	160	7.32	495	22.65	766	35.06	764	34.97
FLD_Q11	47	2.15	285	13.04	782	35.77	1072	49.04
FLD_Q12	140	6.44	671	30.85	907	41.7	457	21.01
FLD_Q13	129	5.97	623	28.84	929	43.01	479	22.18
FLD_Q14	300	13.73	418	19.13	594	27.19	873	39.95
FLD_Q15	533	24.39	584	26.73	580	26.54	488	22.33
FLD_Q16	1105	50.62	607	27.81	290	13.28	181	8.29
FLD_Q17	1093	50.07	615	28.17	293	13.42	182	8.34

NOTE: The Item column references the Program Survey items as presented in the Fieldwork section (page A-7).
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table H-3. Feelings of preparedness frequencies

Item	N, Not at All Prepared	%, Not at All Prepared	N, Somewhat Prepared	%, Somewhat Prepared	N, Mostly Prepared	%, Mostly Prepared	N, Definitely Prepared	%, Definitely Prepared
PREP_Q01	73	3.34	599	27.41	974	44.58	539	24.67
PREP_Q02	68	3.11	486	22.25	978	44.78	652	29.85
PREP_Q03	35	1.6	426	19.51	1015	46.5	707	32.39
PREP_Q04	37	1.69	374	17.12	878	40.18	896	41.01
PREP_Q05	12	0.55	218	10	868	39.82	1082	49.63
PREP_Q06	66	3.03	515	23.68	936	43.03	658	30.25
PREP_Q07	4	0.18	113	5.18	621	28.49	1442	66.15
PREP_Q08	157	7.2	697	31.96	854	39.16	473	21.69
PREP_Q09	83	3.82	332	15.3	659	30.37	1096	50.51
PREP_Q10	39	1.78	250	11.44	737	33.73	1159	53.04
PREP_Q11	15	0.69	151	6.94	596	27.38	1415	65
PREP_Q12	171	7.82	620	28.36	891	40.76	504	23.06
PREP_Q13	60	2.74	446	20.4	1061	48.54	619	28.32

NOTE: The Item column references the Program Survey items as presented in the Feelings of Preparedness section (page A-8).
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX I.
KNOWLEDGE ASSESSMENT SCORING, SCALE, DISTRACTOR, AND FORMS ANALYSIS

This appendix examines pre-service teacher responses to the Knowledge Assessment. Although it is expected that students within the same institution will score more similarly to one another than to students from other institutions (on average), knowledge is an individual-level construct. For the Knowledge Assessment, the study team used scale and distractor analyses to examine the quality of the scales. This appendix presents raw score descriptive statistics and intra-class correlation coefficients to indicate the extent to which responses were similar across pre-service teachers within institutions. The last section describes the analysis performed to determine if responses differed by form (A or B).

Analysis of Knowledge Assessment Items

To analyze the reliability of the Knowledge Assessment, the study team conducted scale and distractor analyses. The purpose of these analyses was to determine (a) the internal consistency reliability of the scales and (b) whether the scales could be improved by eliminating some items. In addition, distractor analysis allowed checks of the keyed responses to determine whether any errors in scoring had occurred.

Scale and Distractor Analysis

As a first step in this process, the study team examined the relation of the items to each scale using the five-factor model, the three-factor model, the two-factor model, and the one-factor model. The team also calculated the coefficient alpha reliability for each scale, including the correlation of each item to the scale score and the scale alpha if the item was deleted.

The alpha coefficients are summarized in Table I-1, with scale means and standard deviations. For the five-factor model, alpha coefficients are between 0.25 and 0.48, and the longer scales in the three-factor model exhibit reliabilities between 0.45 and 0.52.

Table I-1. Initial scale descriptive statistics for the Knowledge Assessment

Essential Component	Items	N	Mean	SD	Alpha
Five scales					
Phonemic Awareness	12	2187	6.74	1.82	.32
Phonics	8	2187	3.68	1.55	.31
Fluency	12	2187	7.33	2.06	.48
Vocabulary	12	2187	6.61	1.76	.25
Comprehension	12	2187	7.65	2.08	.48
Three scales					
Alphabetics	20	2187	10.42	2.67	.45
Fluency	12	2187	7.33	2.06	.48
Meaning	24	2187	14.26	3.08	.52
Two scales					
Word	32	2187	17.75	3.90	.59
Meaning	24	2187	14.26	3.08	.52
One scale	56	2187	32.01	5.94	.69

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

In an attempt to improve the reliability of the scales, the study team conducted a distractor analysis to determine whether many respondents who scored well on the Knowledge Assessment selected an incorrect option ("distractor"), rather than the correct answer ("key"). This kind of answer pattern may suggest either that the item itself was confusing or that the computer program performing the scoring had an error for this item. The study team flagged items for four reasons in the distractor analysis: (1) the item raised the scale coefficient alpha when deleted; (2) the item-total score correlation was below .10; (3) the average score for a distractor was higher than the average score for the keyed (i.e., correct) response; and (4) the keyed response was not the most common response.[39] The results of this analysis are at the end of this appendix (see Tables I-6 and I-7; complete descriptions of the table contents are included following the Forms Analysis section).

Using these rules, the study team reviewed 11 items: 4, 5, 13, 16, 33, 36, 37, 38, 40, 42, and 44. These items functioned similarly across both the five-factor model and the three-factor model. After reviewing the keyed responses and documentation about the development of each item, the team decided to drop three items from the analysis. The first asked about "the most effective format for delivering phonemic awareness instruction." Analysis of distractors showed that the item did not correlate to other items in the assessment. The second dropped item asked about the grade level through which phonics instruction would benefit the spelling performance of "normally progressing" students. Analysis of distractors showed that the upper-grade options were selected by high-performing test-takers, and again the item did not correlate with other phonics items. The correct response for both these items was supported by scientifically based research (NICHD, 2000), but the study team reasoned that pre-service teachers might not be

[39]Additional rules were not useful. Specifically, there were no items with omit rates (percentage of pre-service teachers not answering the question) above 15%, and there were no items with difficulties (proportion of pre-service teachers answering the question correctly) below .05 or above .95.

familiar with this information. The final item that was dropped targeted morphological knowledge by asking for the "root word" of a common word. The psychometric properties of this item suggested that test-takers were confused about whether the English root word or the etymological root was being requested.[40]

Using the results of the distractor analysis of the remaining items, the study team recomputed the scale descriptive statistics. These results are summarized in Table I-2, which shows some improvement for phonemic awareness and phonics in the five-factor model and for alphabetics in the three-factor model; the findings, however, are generally consistent with the results in Table I-1. The scale revisions did not result in any additional items being flagged for review.

Table I-2. Refined scale descriptive statistics for the Knowledge Assessment

Essential Component	Items	N	Mean	SD	Alpha
Five scales					
Phonemic Awareness	11	2187	6.14	1.76	.35
Phonics	7	2187	3.48	1.52	.37
Fluency	12	2187	7.33	2.06	.48
Vocabulary	11	2187	5.83	1.70	.25
Comprehension	12	2187	7.65	2.08	.48
Three scales					
Alphabetics	18	2187	9.63	2.64	.50
Fluency	12	2187	7.33	2.06	.48
Meaning	23	2187	13.48	3.04	.52
Two scales					
Word	30	2187	16.95	3.89	.62
Meaning	23	2187	13.48	3.04	.52
One scale	53	2187	30.43	5.92	.71

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Tables I-3 through I-5 present the intercorrelations for the Knowledge Assessment scales for the five-scale model, the three-scale model, and the two-scale model, respectively. Alpha coefficients are reported in the diagonals. These findings suggest that using three scales has an advantage over the five-scale model. Most notably, the vocabulary subscale does not appear to be adequately reliable (α=.25) by itself and tends to correlate with other subscales at a level similar to its internal consistency estimate (r=.19 to r=.23). When combined with the comprehension subscale (α=.48), the reliability improves (α=.52). Similarly, the phonemic awareness (α=.35) and phonics scales (α=.37) are more reliable when combined (α=.50).

Although the within-scale correlations (alpha coefficients) are slightly higher than the across-scale correlations, these results might argue that a single Knowledge Assessment score might be

[40] In general, the study team was reluctant to drop items solely for psychometric purposes, considering that the item may have different psychometric properties in future administrations. The study team retained items that had strong theoretical bases in some cases where the item statistics were equivocal.

adequate and more reliable. This finding is not uncommon in knowledge assessments, where a general factor tends to run throughout the scale. The reliability of the single-scale Knowledge Assessment score is 0.71. Similar to the decision for the Program Survey constructs (see Appendix G), the report presents Knowledge Assessment findings for both the one-factor model as well as the three-factor models.

Table I-3. Knowledge Assessment subscale intercorrelations for the five-factor model

Essential Component	Phonemic Awareness	Phonics	Fluency	Vocabulary	Comprehension
1. Phonemic Awareness (PA)	(.35)				
2. Phonics (PH)	.29	(.37)			
3. Fluency (FL)	.30	.29	(.48)		
4. Vocabulary (VO)	.22	.19	.23	(.25)	
5. Comprehension (CO)	.30	.29	.33	.29	(.48)

NOTE: Alpha coefficients are presented in the diagonal in parentheses. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table I-4. Knowledge Assessment subscale intercorrelations for the three-factor model

Essential Component	Alphabetics	Fluency	Meaning
1. Alphabetics (AL)	(.45)		
2. Fluency (FL)	.36	(.48)	
3. Meaning (ME)	.39	.35	(.52)

NOTE: Alpha coefficients are presented in the diagonal in parentheses. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table I-5. Knowledge Assessment subscale intercorrelations for the two-factor model

Essential Component	Word	Meaning
1. Word (WO)	(.62)	
2. Meaning (ME)	.45	(.52)

NOTE: Alpha coefficients are presented in the diagonal in parentheses. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Forms Analysis

As a final step in the psychometric analysis process, the study team examined the potential score differences between participants who responded to different forms of the instrument. In Form A, the Program Survey items preceded the Knowledge Assessment items. In Form B, the order was reversed. This analysis was important because large group differences between respondents on each form would suggest that the order of presentation may have influenced scores on the Knowledge Assessment or pre-service teacher ratings of their programs.

Overall, the study team found no form differences that would have suggested that the presentation order for Forms A and B had a difference on pre-service teachers' responses to the program survey or score on the knowledge assessment. The only consistent difference was found for the coursework section of the Program Survey, which had statistically significantly[41] higher ratings in Form B as compared with Form A, although the difference was small.[42] Specifically, the fluency and comprehension scales were about a sixth of a standard deviation higher in Form B than in Form A, and vocabulary was about a tenth of a standard deviation higher in Form B than in Form A. The differences for phonemic awareness and phonics were in the same direction, but not statistically significant. There were no statistically significant differences between Forms A and B for the field experience portion of the survey, and the differences were not in a consistent direction. For the feelings of preparedness items, no single scale was statistically significantly different across forms, and although means for Form A scales were consistently higher than Form B means, the multivariate test was not statistically significant (Wilks' λ 5, 2114 = 1.19, $p > .01$).

For the Knowledge Assessment, only the phonics scale showed any form difference, with Form A being significantly higher than Form B ($t = 2.58$, $p < .01$, $d = 0.11$). Across the other scales, the direction of the difference was not consistent. The Form A means were higher for phonemic awareness ($t = 1.19$, $p > .01$), fluency ($t = 2.13$, $p > .01$), and vocabulary ($t = 0.72$, $p > .01$), although the Form B mean was higher for comprehension ($t = -0.54$, $p > .01$).

Taken together, these findings do not provide strong evidence that the order of presentation significantly influenced pre-service teacher responses to either the Program Survey or the Knowledge Assessment.

Description of Item Statistics

This section describes the item statistics presented in Tables I-6 and I-7. Table I-6 presents the classical item statistics used to evaluate the item functioning of the Knowledge Assessment for the three-factor model. Table I-6 contains basic psychometric information, including the item-total score point-biserial correlation ("Correlation with Total"), the effect on coefficient alpha if the item is deleted, and two flags to indicate potential psychometric problems with the items. The first flag, labeled "Non-Modal Key Flag," indicates that the keyed response was not the most commonly chosen response option. The second flag, labeled "Average Score Flag," indicates that the adjusted mean score for pre-service teachers selecting one of the non-keyed response options is higher than the adjusted mean score for the keyed response options.

Additionally, the tables present the number of pre-service teachers who selected each response option ("e.g., N, Response A" for option A) and those who did not answer the question ("N, Missing"). Corresponding to each count for the response options is the mean subscale score for

[41]The study team used $p < .01$ as the cutoff for statistical significance given the number of contrasts run. An alternative procedure would have been to use a multivariate test (e.g., MANOVA), but the study team thought that it was sufficiently conservative to conduct multiple t-tests for the ease of the reader.

[42] The study team used Cohen's *d* as a measure of effect size (Cohen, 1988), using the common variant using the pooled standard deviation, which works well when group standard deviations are similar (Rosnow & Rosenthal, 1996). Cohen's (1988) suggested criterion for "small" effect sizes was $d < 0.20$.

each pre-service teacher selecting that option, adjusted for whether he or she got credit for that question ("e.g., "Adjusted Mean, Response A" for response option A) and for missing responses. The item key and factor (from the three-factor model) are also included. Table I-7 presents the same information for the five-factor model.

Table I-6. Item analysis for three-factor model

Item	Factor	Correlation with Total	Alpha if Deleted	Non-Modal Key Flag	Average Score Flag	N, Key	Adjusted Mean, Key	N, Distractor 1	Adjusted Mean, Distractor 1
A_Q01	Alphabetics	**0.08**	**0.45**			1263	10.01	27	8.52
A_Q02	Alphabetics	0.20	0.42	*		532	11.07	112	9.71
A_Q03	Alphabetics	0.28	0.40			1138	10.56	508	9.4
A_Q04	Alphabetics	0.11	0.44	*		292	11.00	1326	10.27
A_Q05	Alphabetics	**–0.02**	**0.47**		*	1297	9.77	352	9.89
A_Q06	Alphabetics	**0.08**	0.45			1909	9.63	83	9
A_Q07	Alphabetics	0.14	0.44			1941	9.66	229	8.57
A_Q08	Alphabetics	0.12	0.44			847	10.41	392	9.97
A_Q09	Alphabetics	0.20	0.42			1628	9.97	177	9.09
A_Q10	Alphabetics	0.15	0.43			1188	10.22	702	9.71
A_Q11	Alphabetics	0.10	0.44			1696	9.79	168	9.3
A_Q12	Alphabetics	**0.09**	0.45			1000	10.21	205	9.51
A_Q13	Alphabetics	**–0.07**	**0.47**	*	*	439	9.85	56	9.48
A_Q14	Alphabetics	0.21	0.42			1697	9.93	14	8
A_Q15	Alphabetics	0.31	0.40	*		683	11.27	453	9.02
A_Q16	Alphabetics	**0.04**	**0.46**		*	843	10.16	747	10.31
A_Q17	Alphabetics	0.13	0.44			894	10.42	500	10.1
A_Q18	Alphabetics	0.22	0.42			1046	10.53	116	8.97
A_Q19	Alphabetics	0.16	0.43		*	1485	10.02	177	9.71
A_Q20	Alphabetics	0.14	0.44			972	10.39	683	9.81
A_Q21	Fluency	0.15	0.47	*		458	7.7	532	6.67
A_Q22	Fluency	0.16	0.47			878	7.3	21	6.43
A_Q23	Fluency	0.28	0.43			1429	7.05	117	6.08
A_Q24	Fluency	**0.08**	**0.49**	*		744	7.2	1044	6.9
A_Q25	Fluency	0.19	0.46			1449	6.92	156	6.31
A_Q26	Fluency	0.26	0.44	*		939	7.46	1088	6.48
A_Q27	Fluency	0.11	**0.48**			1781	6.61	111	5.76
A_Q28	Fluency	0.18	0.46			1734	6.71	44	5.43
A_Q29	Fluency	0.23	0.45			1422	7	201	5.97
A_Q30	Fluency	0.18	0.46		*	144	5.41	144	5.41
A_Q31	Fluency	0.19	0.46			1579	6.83	58	5.66

See note at end of table.

Table I-6. Item analysis for three-factor model – Continued

Item	Factor	Correlation with Total	Alpha if Deleted	Non-Modal Key Flag	Average Score Flag	N, Key	Adjusted Mean, Key	N, Distractor 1	Adjusted Mean, Distractor 1
A_Q32	Fluency	0.19	0.46			1730	6.72	139	5.43
A_Q33	Meaning	**0.05**	**0.52**	*		762	14.1	140	13.66
A_Q34	Meaning	0.18	0.50			951	14.44	830	13.51
A_Q35	Meaning	0.25	0.49			1471	14.11	72	11.93
A_Q36	Meaning	**0.02**	**0.52**	*		208	14.38	325	13.72
A_Q37	Meaning	**0.01**	**0.53**	*	*	523	14.05	156	12.4
A_Q38	Meaning	**0.03**	**0.53**		*	1173	13.81	607	13.21
A_Q39	Meaning	0.22	0.49		*	1438	14.06	715	12.77
A_Q40	Meaning	**0.03**	**0.52**		*	1705	13.52	122	13.5
A_Q41	Meaning	0.26	0.49			1693	13.91	224	12.59
A_Q42	Meaning	**0.01**	**0.53**		*	1550	13.57	125	14.02
A_Q43	Meaning	0.18	0.50			1171	14.21	72	11.9
A_Q44	Meaning	**0.04**	**0.52**		*	1804	13.49	34	10.71
A_Q45	Meaning	0.20	0.50			1193	14.25	131	13.21
A_Q46	Meaning	0.16	0.50			1572	13.84	188	12.09
A_Q47	Meaning	0.21	0.50			1680	13.83	113	12.43
A_Q48	Meaning	0.22	0.49			1546	13.96	351	12.79
A_Q49	Meaning	0.20	0.50			1766	13.74	118	12.11
A_Q50	Meaning	0.23	0.49			1139	14.39	364	13.14
A_Q51	Meaning	0.19	0.50			1985	13.53	91	11.58
A_Q52	Meaning	0.21	0.50		*	1565	13.93	200	12.85
A_Q53	Meaning	0.15	0.51			1673	13.75	200	12.85
A_Q54	Meaning	**0.08**	0.52			737	14.25	154	13.38
A_Q55	Meaning	0.20	0.50			1156	14.29	20	11.8
A_Q56	Meaning	0.20	0.50	*		724	14.78	394	13.34

See note at end of table.

Table I-6. Item analysis for three-factor model – Continued

Item	N, Distractor 2	Adjusted Mean, Distractor 2	N, Distractor 3	Adjusted Mean, Distractor 3	N, Missing	Adjusted Mean, Missing
A_Q01	496	9.63	305	9.75	96	9.4
A_Q02	789	9.48	752	10.35	.	.
A_Q03	146	8.4	381	9.22	7	8.29
A_Q04	542	10.04	25	8.64	1	6
A_Q05	494	9.99	40	9.03	1	6
A_Q06	56	8.32	138	9.24	.	.
A_Q07	15	8.33	2	6.5	.	.
A_Q08	351	9.71	592	9.73	.	.
A_Q09	335	8.72	43	8.4	.	.
A_Q10	142	8.85	152	8.93	.	.
A_Q11	309	9.11	13	8.08	.	.
A_Q12	306	9.37	670	10.01	.	.
A_Q13	699	10.24	991	10.42	1	6
A_Q14	292	8.74	179	8.61	1	6
A_Q15	902	9.93	147	9.16	.	.
A_Q16	259	9.56	332	9.48	.	.
A_Q17	555	9.34	233	9.86	.	.
A_Q18	456	9.51	565	9.41	.	.
A_Q19	299	8.79	221	9.17	1	13
A_Q20	181	9.42	349	9.47	.	.
A_Q21	664	6.93	525	7.3	.	.
A_Q22	513	6.46	769	6.83	1	3
A_Q23	92	5.95	547	5.93	1	3
A_Q24	145	6.85	251	6.81	1	3
A_Q25	344	5.79	236	6.63	1	3
A_Q26	37	6.24	117	6.53	5	5.6
A_Q27	283	6.22	10	4.9	1	3
A_Q28	212	5.73	194	6.08	2	4.5
A_Q29	163	6.41	399	6.01	.	.
A_Q30	1879	6.61	119	5.76	4	7
A_Q31	235	5.62	314	6.38	1	4

See note at end of table.

Table I-6. Item analysis for three-factor model – Continued

Item	N, Distractor 2	Adjusted Mean, Distractor 2	N, Distractor 3	Adjusted Mean, Distractor 3	N, Missing	Adjusted Mean, Missing
A_Q32	100	5.72	215	6.15	3	4.67
A_Q33	793	13.98	489	13.56	1	10
A_Q34	219	12.63	185	13.5	.	.
A_Q35	608	12.59	34	12.44	.	.
A_Q36	98	12.44	1555	14.34	1	8
A_Q37	476	13.49	1029	14.51	.	.
A_Q38	287	14.31	112	14.13	.	.
A_Q39	23	11.52	9	11.78	1	15
A_Q40	309	13.39	46	12.26	1	17
A_Q41	103	11.64	163	11.59	1	12
A_Q42	55	11.85	454	13.6	.	.
A_Q43	252	12.13	687	13.69	1	13
A_Q44	62	11.63	283	13.78	3	12.67
A_Q45	310	12.93	549	13.12	2	11.5
A_Q46	346	13.11	79	13.08	.	.
A_Q47	83	12.08	309	12.39	1	10
A_Q48	118	12.18	166	12.3	.	.
A_Q49	202	12.1	99	12.67	2	10
A_Q50	210	12.35	472	13.27	.	.
A_Q51	1139	14.39	96	11.41	.	.
A_Q52	284	12.36	13	10.77	1	16
A_Q53	280	12.65	33	11.7	.	.
A_Q54	581	13.49	714	14.06	.	.
A_Q55	1007	13.13	4	13.25	.	.
A_Q56	305	13.3	764	13.67	.	.

NOTE: Item = order of item in assessment. Factor =component of early reading from three-factor model. Correlation with total = point-biserial correlation of the item with the total score on the other items on the factor subscale. Alpha if item deleted = Coefficient alpha for the factor subscale if the item is deleted. Non-modal key flag indicates whether the key is not the most common response options, indicated with an asterisk (*). Average score flag indicates an item where the keyed response is not the response with the highest mean score on the other items on the subscale, indicated with an asterisk (*). N, Distractor 1 is the number of examinees who selected the first non-keyed distractor (as are N, Distractor 2 and N, Distractor 3, respectively). Adjusted Mean, Distractor 1 is the mean score of pre-service teachers who selected the first non-keyed response option (as are Adjusted Mean, Distractor 2 and Adjusted Mean, Distractor 3, respectively). N, Missing and Adjusted Mean, Missing are the corresponding statistics for missing responses (omissions). Numbers may not be equal for all rows because of double responses or erasure errors that could not be resolved. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table I-7. Item analysis for five-factor model

Item	Factor	Correlation with Total	Alpha if Deleted	Non-Modal Key Flag	Average Score Flag	N, Key	Adjusted Mean, Key
A_Q01	Ph. Aware	**0.10**	0.30			1263	10.01
A_Q02	Ph. Aware	0.15	0.28	*		532	11.07
A_Q03	Ph. Aware	0.22	0.25			1138	10.56
A_Q04	Ph. Aware	**0.08**	0.31	*		292	11.00
A_Q05	Ph. Aware	**–0.01**	**0.35**		*	1297	9.77
A_Q06	Ph. Aware	**0.09**	0.31			1909	9.63
A_Q07	Ph. Aware	0.12	0.30			1941	9.66
A_Q08	Ph. Aware	0.12	0.29			847	10.41
A_Q09	Ph. Aware	0.16	0.28			1628	9.97
A_Q10	Ph. Aware	**0.09**	0.31			1188	10.22
A_Q11	Ph. Aware	**0.07**	0.31			1696	9.79
A_Q12	Ph. Aware	0.10	0.30			1000	10.21
A_Q13	Phonics	**–0.07**	**0.37**	*	*	439	9.85
A_Q14	Phonics	0.15	0.27			1697	9.93
A_Q15	Phonics	0.27	0.20	*		683	11.27
A_Q16	Phonics	**0.03**	**0.34**		*	843	10.16
A_Q17	Phonics	0.10	0.30			894	10.42
A_Q18	Phonics	0.25	0.21			1046	10.53
A_Q19	Phonics	0.15	0.27		*	1485	10.02
A_Q20	Phonics	0.12	0.29			972	10.39
A_Q21	Fluency	0.15	0.47	*		458	7.7
A_Q22	Fluency	0.16	0.47			878	7.3
A_Q23	Fluency	0.28	0.43			1429	7.05
A_Q24	Fluency	**0.08**	**0.49**	*		744	7.2
A_Q25	Fluency	0.19	0.46			1449	6.92
A_Q26	Fluency	0.26	0.44	*		939	7.46
A_Q27	Fluency	0.11	**0.48**			1781	6.61
A_Q28	Fluency	0.18	0.46			1734	6.71
A_Q29	Fluency	0.23	0.45			1422	7
A_Q30	Fluency	0.18	0.46		*	144	5.41
A_Q31	Fluency	0.19	0.46			1579	6.83

See note at the end of the table.

Table I-7. Item analysis for five-factor model – Continued

Item	Factor	Correlation with Total	Alpha if Deleted	Non-Modal Key Flag	Average Score Flag	N, Response A	Adjusted Mean, Response A
A_Q32	Fluency	0.19	0.46			1730	6.72
A_Q33	Vocab	**0.02**	**0.26**	*		762	14.1
A_Q34	Vocab	0.14	0.20			951	14.44
A_Q35	Vocab	0.20	0.17			1471	14.11
A_Q36	Vocab	**0.01**	**0.25**	*		208	14.38
A_Q37	Vocab	**0.01**	**0.26**	*	*	523	14.05
A_Q38	Vocab	**0.04**	**0.25**		*	1173	13.81
A_Q39	Vocab	0.17	0.19			1438	14.06
A_Q40	Vocab	**0.03**	**0.25**		*	1705	13.52
A_Q41	Vocab	0.20	0.18			1693	13.91
A_Q42	Vocab	**0.02**	**0.26**		*	1550	13.57
A_Q43	Vocab	**0.09**	0.23		*	1171	14.21
A_Q44	Vocab	**0.01**	**0.26**		*	1804	13.49
A_Q45	Comp	0.19	0.46			1193	14.25
A_Q46	Comp	0.16	0.47			1572	13.84
A_Q47	Comp	0.19	0.46			1680	13.83
A_Q48	Comp	0.22	0.45			1546	13.96
A_Q49	Comp	0.19	0.46			1766	13.74
A_Q50	Comp	0.25	0.44			1139	14.39
A_Q51	Comp	0.19	0.47			1985	13.53
A_Q52	Comp	0.22	0.45		*	1565	13.93
A_Q53	Comp	0.16	0.47			1673	13.75
A_Q54	Comp	**0.08**	**0.49**			737	14.25
A_Q55	Comp	0.18	0.47			1156	14.29
A_Q56	Comp	0.20	0.46	*		724	14.78

See note at the end of the table.

Table I-7. Item analysis for five-factor model – Continued

Item	N, Distractor 1	Adjusted Mean, Distractor 1	N, Distractor 2	Adjusted Mean, Distractor 2	N, Distractor 3	Adjusted Mean, Distractor 3	N, Missing	Adjusted Mean, Missing
A_Q01	27	8.52	496	9.63	305	9.75	96	6.08
A_Q02	112	9.71	789	9.48	752	10.35	.	.
A_Q03	508	9.4	146	8.4	381	9.22	7	5.14
A_Q04	1326	10.27	542	10.04	25	8.64	1	5
A_Q05	352	9.89	494	9.99	40	9.03	1	5
A_Q06	83	9	56	8.32	138	9.24	.	.
A_Q07	229	8.57	15	8.33	2	6.5	.	.
A_Q08	392	9.97	351	9.71	592	9.73	.	.
A_Q09	177	9.09	335	8.72	43	8.4	.	.
A_Q10	702	9.71	142	8.85	152	8.93	.	.
A_Q11	168	9.3	309	9.11	13	8.08	.	.
A_Q12	205	9.51	306	9.37	670	10.01	.	.
A_Q13	56	9.48	699	10.24	991	10.42	1	1
A_Q14	14	8	292	8.74	179	8.61	1	1
A_Q15	453	9.02	902	9.93	147	9.16	.	.
A_Q16	747	10.31	259	9.56	332	9.48	.	.
A_Q17	500	10.1	555	9.34	233	9.86	.	.
A_Q18	116	8.97	456	9.51	565	9.41	.	.
A_Q19	177	9.71	299	8.79	221	9.17	1	5
A_Q20	683	9.81	181	9.42	349	9.47	.	.
A_Q21	532	6.67	664	6.93	525	7.3	.	.
A_Q22	21	6.43	513	6.46	769	6.83	1	3
A_Q23	117	6.08	92	5.95	547	5.93	1	3
A_Q24	1044	6.9	145	6.85	251	6.81	1	3
A_Q25	156	6.31	344	5.79	236	6.63	1	3
A_Q26	1088	6.48	37	6.24	117	6.53	5	5.6
A_Q27	111	5.76	283	6.22	10	4.9	1	3
A_Q28	44	5.43	212	5.73	194	6.08	2	4.5
A_Q29	201	5.97	163	6.41	399	6.01	.	.
A_Q30	144	5.41	1879	6.61	119	5.76	4	7
A_Q31	58	5.66	235	5.62	314	6.38	1	4

See note at the end of the table.

Table I-7. Item analysis for five-factor model – Continued

Item	N, Distractor 1	Adjusted Mean, Distractor 1	N, Distractor 2	Adjusted Mean, Distractor 2	N, Distractor 3	Adjusted Mean, Distractor 3	N, Missing	Adjusted Mean, Missing
A_Q32	139	5.43	100	5.72	215	6.15	3	4.67
A_Q33	140	13.66	793	13.98	489	13.56	1	3
A_Q34	830	13.51	219	12.63	185	13.5	.	.
A_Q35	72	11.93	608	12.59	34	12.44	.	.
A_Q36	325	13.72	98	12.44	1555	14.34	1	3
A_Q37	156	12.4	476	13.49	1029	14.51	.	.
A_Q38	607	13.21	287	14.31	112	14.13	.	.
A_Q39	715	12.77	23	11.52	9	11.78	1	6
A_Q40	122	13.5	309	13.39	46	12.26	1	7
A_Q41	224	12.59	103	11.64	163	11.59	1	5
A_Q42	125	14.02	55	11.85	454	13.6	.	.
A_Q43	72	11.9	252	12.13	687	13.69	1	7
A_Q44	34	10.71	62	11.63	283	13.78	3	5.33
A_Q45	131	13.21	310	12.93	549	13.12	2	7
A_Q46	188	12.09	346	13.11	79	13.08	.	.
A_Q47	113	12.43	83	12.08	309	12.39	1	5
A_Q48	351	12.79	118	12.18	166	12.3	.	.
A_Q49	118	12.11	202	12.1	99	12.67	2	4.5
A_Q50	364	13.14	210	12.35	472	13.27	.	.
A_Q51	91	11.58	1139	14.39	96	11.41	.	.
A_Q52	200	12.85	284	12.36	13	10.77	1	8
A_Q53	200	12.85	280	12.65	33	11.7	.	.
A_Q54	154	13.38	581	13.49	714	14.06	.	.
A_Q55	20	11.8	1007	13.13	4	13.25	.	.
A_Q56	394	13.34	305	13.3	764	13.67	.	.

NOTE: Item = order of item in assessment. Factor =component of early reading from three-factor model. Correlation with total = point-biserial correlation of the item with the total score on the other items on the factor subscale. Alpha if item deleted = Coefficient alpha for the factor subscale if the item is deleted. Non-modal key flag indicates whether the key is not the most common response options, indicated with an asterisk (*). Average score flag indicates an item where the keyed response is not the response with the highest mean score on the other items on the subscale, indicated with an asterisk (*). N, Distractor 1 is the number of examinees who selected the first non-keyed distractor (as are N, Distractor 2 and N, Distractor 3, respectively). Adjusted Mean, Distractor 1 is the mean score of pre-service teachers who selected the first non-keyed response option (as are Adjusted Mean, Distractor 2 and Adjusted Mean, Distractor 3, respectively). N, Missing and Adjusted Mean, Missing are the corresponding statistics for missing responses (omissions). Numbers may not be equal for all rows because of double responses or erasure errors that could not be resolved. N of teachers = 2,187.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction,* 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX J.
KNOWLEDGE ASSESSMENT MEANS TABLES

The tables in Appendix J were produced using SAS Proc SurveyFreq, which generates weighted mean scores for complex samples. Assessment means are provided by total and each subscale.

Table J-1a. Assessment means by overall GPA

		Assessment Total				Alphabetics			
Overall GPA	N	Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
1.7–1.9 (C– or 70–72)	‡	—	—	—	—	—	—	—	—
2.0–2.2 (C or 73–76)	‡	—	—	—	—	—	—	—	—
2.3–2.6 (C+ or 77–79)	13	49.35	1.95	44.31	54.38	42.28	2.93	34.74	49.82
2.7–2.9 (B– or 80–82)	96	51.96	1.63	48.53	55.38	50.45	1.92	46.42	54.49
3.0–3.2 (B or 83–86)	258	50.97	1.19	48.51	53.44	45.93	1.27	43.30	48.55
3.3–3.6 (B+ or 87–89)	836	55.75	0.88	53.92	57.58	51.53	1.02	49.42	53.64
3.7–4.0 (A or 90–100)	964	61.12	0.88	59.29	62.95	57.21	1.22	54.69	59.72

		Fluency				Meaning			
	N	Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
1.7–1.9 (C– or 70–72)	‡	—	—	—	—	—	—	—	—
2.0–2.2 (C or 73–76)	‡	—	—	—	—	—	—	—	—
2.3–2.6 (C+ or 77–79)	13	57.69	3.61	48.40	66.98	50.53	2.67	43.66	57.40
2.7–2.9 (B– or 80–82)	96	53.13	1.82	49.31	56.95	52.52	2.14	48.02	57.02
3.0–3.2 (B or 83–86)	258	56.17	1.78	52.49	59.85	52.21	1.51	49.08	55.34
3.3–3.6 (B+ or 87–89)	836	59.56	1.74	55.97	63.15	57.07	0.65	55.72	58.41
3.7–4.0 (A or 90–100)	964	65.12	1.37	62.30	67.95	62.10	0.59	60.87	63.33

‡ Reporting standards not met; cell counts suppressed where n < 3.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table J-1b. Assessment means by education GPA

Education GPA	N	Assessment Total Mean	Std err	Lower CI	Upper CI	Alphabetics Mean	Std err	Lower CI	Upper CI
1.7–1.9 (C– or 70–72)	0	—	—	—	—	—	—	—	—
2.0–2.2 (C or 73–76)	0	—	—	—	—	—	—	—	—
2.3–2.6 (C+ or 77–79)	5	55.38	0.91	52.86	57.89	53.87	1.52	49.65	58.10
2.7–2.9 (B– or 80–82)	18	48.95	2.99	42.29	55.60	47.66	2.97	41.04	54.29
3.0–3.2 (B or 83–86)	81	50.48	1.36	47.62	53.34	45.71	2.30	40.87	50.55
3.3–3.6 (B+ or 87–89)	470	52.52	0.84	50.78	54.26	47.90	1.26	45.29	50.50
3.7–4.0 (A or 90–100)	1,570	59.13	0.89	57.30	60.97	55.28	1.03	53.15	57.40

	N	Fluency Mean	Std err	Lower CI	Upper CI	Meaning Mean	Std err	Lower CI	Upper CI
1.7–1.9 (C– or 70–72)	0	—	—	—	—	—	—	—	—
2.0–2.2 (C or 73–76)	0	—	—	—	—	—	—	—	—
2.3–2.6 (C+ or 77–79)	5	72.69	3.85	61.99	83.35	47.53	3.64	37.42	57.64
2.7–2.9 (B– or 80–82)	18	50.13	3.88	41.48	58.78	49.33	3.43	41.70	55.96
3.0–3.2 (B or 83–86)	81	52.13	2.09	47.75	56.52	53.35	1.70	49.79	56.91
3.3–3.6 (B+ or 87–89)	470	57.33	1.42	54.39	60.27	53.63	0.80	51.97	55.29
3.7–4.0 (A or 90–100)	1,570	62.87	1.51	59.76	65.99	60.20	0.62	58.93	61.48

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table J-1c. Assessment means by combined SAT score

Combined SAT Score	N	Assessment Total Mean	Assessment Total Std err	Assessment Total Lower CI	Assessment Total Upper CI	Alphabetics Mean	Alphabetics Std err	Alphabetics Lower CI	Alphabetics Upper CI
790 or lower	15	50.57	2.15	45.48	55.66	54.34	2.77	47.79	60.89
800–890	57	49.16	2.27	44.30	54.03	44.91	3.12	38.22	51.60
900–990	191	53.84	1.06	51.63	56.05	48.58	1.35	45.76	51.40
1000–1090	291	56.11	1.06	53.89	58.33	51.18	1.29	48.49	53.88
1100–1190	322	58.68	1.30	55.97	61.38	54.03	1.52	50.87	57.19
1200–1290	188	60.05	1.78	56.33	63.78	56.60	2.25	51.89	61.30
1300–1390	68	68.47	1.28	65.73	71.20	66.64	2.00	62.34	70.94
1400 or higher	16	71.53	2.54	65.67	77.40	62.16	4.22	52.44	71.88

		Fluency Mean	Fluency Std err	Fluency Lower CI	Fluency Upper CI	Meaning Mean	Meaning Std err	Meaning Lower CI	Meaning Upper CI
790 or lower	15	51.09	2.66	44.80	57.38	47.35	3.81	38.34	56.35
800–890	57	52.31	3.68	44.41	60.21	50.85	2.14	46.26	55.43
900–990	191	59.04	1.90	55.06	63.01	55.24	0.90	53.36	57.13
1000–1090	291	62.14	1.44	59.13	65.16	56.81	1.37	53.94	59.67
1100–1190	322	62.29	1.97	58.18	66.39	60.43	1.08	58.17	62.69
1200–1290	188	60.99	2.61	55.51	66.46	62.27	1.63	58.86	65.67
1300–1390	68	73.62	1.81	69.75	77.50	67.21	2.32	62.24	72.19
1400 or higher	16	70.62	3.33	62.95	78.30	79.34	2.27	74.11	84.57

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table J-1d. Assessment means by ACT score

		Assessment Total				Alphabetics			
ACT Score	N	Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
15 or lower	6	47.08	2.15	37.83	56.34	36.17	3.13	22.68	49.67
16–18	63	48.66	3.12	41.86	55.45	44.25	3.45	36.73	51.78
19–20	82	54.11	1.32	51.33	56.89	50.28	2.00	46.05	54.51
21–23	205	56.53	1.24	53.93	59.12	53.58	1.88	49.65	57.51
24–25	111	56.28	2.38	51.31	61.25	51.75	2.51	46.52	56.98
26–28	176	61.73	1.39	58.83	64.63	58.77	1.88	54.87	62.67
29–31	67	65.32	1.96	61.20	69.44	62.66	4.26	53.70	71.62
32+	11	62.30	5.83	48.03	76.57	60.99	4.57	49.82	72.16
		Fluency				Meaning			
		Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
15 or lower	6	62.02	5.04	40.34	83.70	47.82	5.14	25.68	70.00
16–18	63	58.48	3.32	51.26	65.71	46.98	3.36	39.66	54.30
19–20	82	56.80	2.41	51.72	61.88	55.70	1.37	52.81	58.60
21–23	205	59.77	1.99	55.62	63.93	57.14	0.95	55.15	59.12
24–25	111	59.59	3.27	52.77	66.41	58.10	2.18	53.56	62.65
26–28	176	63.95	2.27	59.23	68.66	62.89	1.09	60.62	65.16
29–31	67	67.42	3.36	60.36	74.48	66.31	1.07	64.07	68.56
32+	11	61.15	8.90	39.37	82.93	63.94	6.02	49.21	78.66

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table J-1e. Assessment means by combined GRE score

Combined GRE Score	N	Assessment Total				Alphabetics			
		Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
790 or lower	4	59.73	3.51	48.56	70.90	54.81	6.03	35.63	74.00
800–890	8	58.33	2.54	52.12	64.55	50.23	2.94	43.03	57.44
900–990	31	59.08	1.12	56.62	61.55	56.76	1.85	52.69	60.83
1000–1090	33	63.71	1.66	59.87	67.55	60.58	3.40	52.75	68.42
1100–1190	39	65.31	2.55	59.81	70.81	62.81	3.73	54.74	70.88
1200–1290	35	69.41	2.47	63.89	74.92	67.07	2.80	60.82	73.32
1300–1390	15	71.09	2.01	66.17	76.00	67.79	2.23	62.34	73.24
1400 or higher	6	69.70	4.81	56.36	83.04	67.37	4.18	55.75	79.00

		Fluency				Meaning			
		Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
790 or lower	4	67.48	7.36	44.06	90.90	59.53	6.83	37.78	81.28
800–890	8	67.97	3.54	59.32	76.63	59.64	3.81	50.32	68.96
900–990	31	64.68	3.43	57.13	72.22	57.98	2.84	51.72	64.24
1000–1090	33	68.71	2.65	62.60	74.82	63.55	1.18	60.82	66.28
1100–1190	39	70.46	2.28	65.53	75.38	64.58	2.77	58.60	70.57
1200–1290	35	72.32	4.15	63.08	81.56	69.72	1.92	65.43	74.00
1300–1390	15	71.06	5.01	58.79	83.32	73.68	1.74	69.41	77.95
1400 or higher	6	72.47	2.80	64.69	80.25	70.08	6.77	51.29	88.86

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Table J-1f. Assessment means by expected degree level

Expected Degree Level	N	Assessment Total				Alphabetics			
		Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
Undergraduate (e.g., BA BS BSEd)	1,580	56.48	0.87	54.68	58.28	52.26	1.02	50.16	54.37
Post-Baccalaureate (Postbac) (e.g., 5th year program non-masters)	93	57.97	1.41	54.94	61.00	53.99	1.98	49.75	58.24
Graduate (e.g., MA MS MEd)	510	59.29	1.32	56.55	62.04	55.76	1.59	52.45	59.07

		Fluency				Meaning			
		Mean	Std err	Lower CI	Upper CI	Mean	Std err	Lower CI	Upper CI
Undergraduate (e.g., BA BS BSEd)	1,580	60.55	1.44	57.58	63.52	57.66	0.66	56.29	59.02
Post-Baccalaureate (Postbac) (e.g., 5th year program non-masters)	93	62.28	1.82	58.37	66.19	58.83	1.43	55.77	61.89
Graduate (e.g., MA MS MEd)	510	62.64	2.08	58.29	66.99	60.31	0.96	58.31	62.31

SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

APPENDIX K.
HIERARCHICAL LINEAR MODEL USED TO ANSWER THE PRIMARY AND SECONDARY RESEARCH QUESTIONS

This appendix provides additional information about the hierarchical linear models used in the study.

Primary Research Questions

- Primary Research Question 1: To what extent does the content of teacher education programs focus on the essential components of early reading instruction?
- Primary Research Question 2: To what extent are graduating pre-service teachers knowledgeable about the essential components of early reading instruction?

Primary Research Question 1: National Estimates of Programs Focus on the Essential Components of Early Reading Instruction

To answer the first primary research question of this study, the study team used data from the 99 institutions that agreed to participate in the study to compute national estimates of program focus based on a multilevel model that explicitly takes into account the nested data structure (items nested within teachers and teachers nested within states). Ideally the study team would like to construct a four-level HLM model that nests survey items within pre-service teachers, pre-service teachers within teacher training institutions, and institutions of teacher training within states. However, the current HLM software program can accommodate only up to three levels. Therefore, the study team omitted the institution level and used states as the level-3 units, because the standard error of estimate in a multilevel context depends primarily on the number of units at the highest level of aggregation (state in this case). The study team constructed separate HLM models for analyzing data related to program focus on the essential components of early reading instruction, data related to pre-service teachers' feelings of preparedness to teach these components as described below.

To generate the national estimates of program focus, the study team combined coursework data and field experience data and created a set of dummy indicator variables to distinguish different aspect of program (i.e., coursework vs. field experience), different components (i.e., alphabetics, fluency, and comprehension), and different data-by-component combinations (i.e., coursework_alphabetics, coursework_fluency, coursework_comprehension, field_alphabetics, field_fluency, and field_comprehension). This set of dummy indicator variables allowed the study team to not only obtain national estimates of program focus on different components and based on different aspects of the program but also test whether the differences between different national estimates were statistically significant. The study team specified the following model, for instance, to obtain the national estimate of program focus on comprehension based on coursework data, and at the same time test whether it is significantly different from program focus on other components represented in the model.

Level 1 (item level)

$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{coursework\_alphabetics})_{ijk} + \pi_{2jk}(\text{coursework\_fluency})_{ijk} + \pi_{3jk}(\text{field\_alphabetics})_{ijk} + \pi_{4jk}(\text{field\_ fluency})_{ijk} + \pi_{5jk}(\text{field\_ comprehension})_{ijk} + e_{ijk}$

Where

- Y_{ijk} is the response of teacher j in state k to survey item i.
- π_{0jk} is the average response of teacher j in state k to items related to comprehension based on coursework data.
- $\pi_{1jk} \sim \pi_{5jk}$ are the differences between the response of teacher j in state k to items related to a given component based on a given aspect of program and her response to items related to comprehension based on coursework data.
- e_{ijk} is a random error associated with the response of teacher j in state k to item i.

Level 2 (teacher level)

$\pi_{0jk} = \beta_{00k} + r_{0jk}$
$\pi_{1jk} = \beta_{10k}$
$\pi_{2jk} = \beta_{20k}$
$\pi_{3jk} = \beta_{30k}$
$\pi_{4jk} = \beta_{40k}$
$\pi_{5jk} = \beta_{50k}$

Where

- β_{00k} is the average response to items related to comprehension based on coursework data across all teachers in state k.
- $\beta_{10k} \sim \beta_{50k}$ are the differences between the average response to items related to a given component based on a given aspect of program and the average response to items related to comprehension based on coursework data across all teachers in state k.
- r_{0jk} is a random error associated with teacher j in state k on the response to items related to comprehension based on coursework data.

Level 3 (state level)

$\beta_{00k} = \gamma_{000} + u_{00k}$
$\beta_{10k} = \gamma_{100}$
$\beta_{20k} = \gamma_{200}$
$\beta_{30k} = \gamma_{300}$
$\beta_{40k} = \gamma_{400}$
$\beta_{50k} = \gamma_{500}$

Where

- γ_{000} is the average response to items related to comprehension based on coursework data across all teachers in all states in the study sample.

- $\gamma_{100} \sim \gamma_{500}$ are the differences between the average response to items related to a given component based on a given aspect of program and the average response to items related to comprehension based on coursework data across all teachers in all states in the study sample.
- u_{00k} is a random error associated with state k on the average teacher response to items related to comprehension based on coursework data.

The level-3 intercept γ_{000} from the model above represents the national estimate of program focus on comprehension based on coursework data. The other five level-3 fixed effects ($\gamma_{100} \sim \gamma_{500}$) represent the differences between this estimate and the estimates of program focus on other components based on coursework or field experience data. By specifying different data-by-component combinations as the omitted reference in the level-1 model, the study team was able to both obtain the national estimate of program focus on individual components based on either coursework or field experience data and test the differences between different national estimates.

Similarly, by specifying different components as the omitted reference in the level-1 model, the study team was able to both obtain the national estimate of program focus on individual components and test the differences between these national estimates based on coursework and field experience data combined. The level-3 intercept γ_{000} from the following model, for example, represents the national estimate of program focus on comprehension, and γ_{100} and γ_{200} represent the differences between program focus on comprehension and program focus on alphabetics and fluency, respectively, based on coursework and field experience data combined.

Level 1 (item level)

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{alphabetics})_{ijk} + \pi_{2jk}(\text{fluency})_{ijk} + e_{ijk}$$

Level 2 (teacher level)

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k}$$
$$\pi_{2jk} = \beta_{20k}$$

Level 3 (state level)

$$\beta_{00k} = \gamma_{000} + u_{00k}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$

The study team further modified the model by including either the dummy indicator for coursework or the indicator for field experience as the only predictor in the level-1 model, which generated both the national estimate of program focus across all components based on either coursework or field experience data and the difference between the two estimates. The study team used a similar approach to obtain the national estimates of pre-service teachers' feelings of preparedness to teach the essential components of early reading instruction and to test the differences between the estimates.

Primary Research Question 2: National Estimates of Pre-service Teachers' Scores on the Knowledge Assessment

The national estimates of pre-service teacher knowledge were estimated using a model similar to the model used for the national estimates of coursework and feelings of preparedness, although these estimates were obtained from SAS (using Proc Surveymeans) as opposed to HLM. These procedures yield consistent estimates, although we used a Rasch-based model in the HLM analyses, we used raw scores to generate the national estimates so that the results would be interpretable in the raw score metric. This model is specified below.

$$National\ mean = \frac{\sum_S w_k y_k}{\sum_S w_k}.$$

Where

- w_k is the weight for each of S prime sampling units (PSUs; in this case, states),
- y_k is the mean for each PSU, and
- the national estimate of the mean is summated across PSUs.

Secondary Research Questions

In addition to the primary research questions about the content of pre-service teacher training programs and pre-service teachers' knowledge of the five essential components of early reading instruction, the *Study of Teacher Preparation in Early Reading Instruction* addressed three additional secondary research questions.

- Research Question 3a: Which characteristics of teacher training institutions and programs are associated with their focus on the essential components of early reading instruction?
- Research Question 3b: To what extent are teacher training programs' focus on the essential components of early reading instruction associated with pre-service teachers' knowledge about these components?
- Research Question 3c: To what extent is pre-service teachers' knowledge about the essential components of early reading instruction related to these pre-service teachers' feelings of preparedness to teach various aspects of beginning reading?

Data from both the Program Survey and the Knowledge Assessment were used to answer these questions. Final teacher weights were applied to these analyses so that the results are generalizable to the population of pre-service teachers in the nation. This appendix supplements information in Chapter 4 by presenting the analytic methods used in answering these questions.

Research Question 3a: Institutional Characteristics, Program Type, and Pre-service Teachers' Reports of Programmatic Focus in Their Teacher Training Programs

The HLM model presented at the beginning of this appendix allowed the study team to compute not only national estimates of program focus but also estimates of program focus for individual

pre-service teachers.[43] Based on the pre-service teacher–level estimates, the study team computed the national estimates of program focus for subpopulations defined by the characteristics of the institutions that agreed to participate in the study (i.e., sector and highest degree offered) and program type (i.e., early childhood education, elementary education, combined programs, and other program) and tested the differences among the subpopulations, while taking into account the sampling design of the study. To guard against inflated Type I errors (i.e., obtaining false findings due to chance) resulting from multiple pairwise comparisons of more than two subgroups, the study team first performed a global F test of each program focus measure to determine whether there was a significant overall difference among the subgroups and conducted pairwise comparisons only if the global test indicated a significant overall difference. No covariates were used in these analyses.

Research Question 3b: Pre-service Teachers' Scores on the Knowledge Assessment and Their Reports of Programmatic Focus in Pre-service Teacher Training Programs

The three-level HLM model is described below:

Level 1 (teacher level)

$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{Cert})_{ijk} + \pi_{2jk}(\text{Degree})_{ijk} + \pi_{3jk}(\text{Race/ethnicity})_{ijk} + \pi_{4jk}(\text{Field experience exposure})_{ijk} + e_{ijk}$,

Where

- Y_{ijk} is a measure of teacher knowledge about a given area of early reading instruction for teacher i in institution j in state k. In this study, teacher knowledge was measured as empirical Bayes estimate from Item Response Analysis using Rasch model.
- π_{0jk} is the average level of teacher knowledge about the given area of early reading instruction across all teachers in institution j in state k, which will be modeled as a random effect at level 2, indicating that it varies systematically across institutions.
- Cert (certification status), Degree (degree level), and Race/ethnicity (racial/ethnic background) are the set of background characteristics of teacher i in institution j in state k. They were centered at their respective grand means. Certification status indicates whether the pre-service teacher held any previous teaching certification (0 represents "not certified" and 1 represents "certified"). Degree level represents the degree a pre-service teacher was working toward (0 represents "undergraduate or post-baccalaureate" and 1 represents "graduate"). Race/ethnicity status is based on pre-service teachers' self-reported racial/ethnic background and was coded such that 0 represents White/Asian, which includes White and Asian groups, and 1 represents non-White/Asian, which includes all other racial/ethnic groups. π_{1jk} , π_{2jk}, and π_{3jk} represent the relationships between pre-service teachers' background characteristics and their knowledge about the given area of reading instruction in institution j in state k. They will be modeled as

[43] The pre-service teacher–level estimates were computed as the sum of level-2 intercept and residual, which could be obtained from the level-2 residual file produced as part of the HLM analysis.

fixed effects at both the institution level and the state level, assuming they are constant across different institutions and states.[44]

- π_{4jk} represents the relationship between pre-service teachers' self-reported program exposure to the given area of early reading instruction through field experience and their knowledge about the area in institution j in state k, adjusted for teacher background characteristics. It will be modeled as a fixed effect at both the institution level and state level.[45]
- e_{ijk} is a random error associated with teacher i in institution j in state k on teacher knowledge.

Level 2 (institution level)

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ (Coursework emphasis) + r_{0jk}
$\pi_{1jk} = \beta_{10k}$
$\pi_{2jk} = \beta_{20k}$
$\pi_{3jk} = \beta_{30k}$
$\pi_{4jk} = \beta_{40k}$,

Where

- β_{00k} is the average level of teacher knowledge about the given area of early reading instruction across all institutions in state k.
- β_{01k} represents the relationship between the institution-level coursework emphasis on the given area and teacher knowledge about the area across all institutions in state k, adjusted for teacher background characteristics and field experience exposure. It will be modeled as a random effect at the state level.[46]
- r_{0jk} is a random error associated with institution j in state k on the average level of teacher knowledge.
- β_{10k}, β_{20k}, and β_{30k} represent the average relationships between teacher characteristics and teacher knowledge across all institutions in state k.
- β_{40k} represents the average relationship between teacher-reported field experience exposure and teacher knowledge, adjusted for teacher background characteristics, across all institutions in state k.

Level 3 (state level)

$\beta_{00k} = \gamma_{000} + u_{00k}$
$\beta_{01k} = \gamma_{010} + u_{01k}$
$\beta_{10k} = \gamma_{100}$
$\beta_{20k} = \gamma_{200}$
$\beta_{30k} = \gamma_{300}$

[44] The relationships between teacher background characteristics and teacher-reported program emphasis may actually vary across different institutions. However, since they were not the emphasis of the study, they were modeled as fixed effects at both the institution level and the state level for simplicity.

[45] In the initial estimation of the model, π_{4jk} was modeled as a random effect at both institution level and state level, assuming it would vary across different institutions and states. However, the results showed that it did not vary systematically across institutions or states and therefore was specified as a fixed effect at both levels in the final model.

[46] The initial estimation of the model confirmed that β_{01k} varied across states and therefore was modeled as a random effect at the state level.

$\beta_{40k} = \gamma_{400}$,

Where

- γ_{000} is the average level of teacher knowledge about the given area of reading instruction across all states.
- u_{00k} is a random error associated with state k on the average level of teacher knowledge.
- γ_{010} represents the average relationship between the institution-level coursework emphasis and teacher knowledge across all states, adjusted for teacher background characteristics and field experience exposure.
- u_{01k} is a random error associated with state k on the relationship between coursework emphasis and teacher knowledge.
- γ_{100}, γ_{200}, and γ_{300} represent the average relationships between teacher characteristics and teacher knowledge across all states.
- γ_{400} represents the average relationship between teacher-reported field experience exposure and teacher knowledge, adjusted for teacher background characteristics, across all institutions and all states.

Prior to applying this model, the study team estimated a three-level unconditional model without any control variables or predictor variables. Such an unconditional model is often a good starting point in hierarchical data analysis because it provides useful information about the outcome variability on the different levels of the hierarchy. The variance components on each of the three levels of the unconditional model show that for all scales of early reading instruction, pre-service teachers' knowledge varies significantly on both the institution level and the state level, which warranted the use of the three-level model empirically.

Research Question 3c: Pre-service Teachers' Reports of Feelings of Preparedness to Teach the Five Essential Components of Early Reading and Their Knowledge Assessment Scores

The HLM model used to address Research Question 3b was adapted to address this research question. In specific, the model is shown below:

Level 1 (teacher level)

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{Cert})_{ijk} + \pi_{2jk}(\text{Degree})_{ijk} + \pi_{3jk}(\text{Race/ethnicity})_{ijk} + \pi_{4jk}(\text{Teacher Knowledge})_{ijk} + e_{ijk}$$

Level 2 (institution level)

$\pi_{0jk} = \beta_{00k} + r_{0jk}$
$\pi_{1jk} = \beta_{10k}$
$\pi_{2jk} = \beta_{20k}$
$\pi_{3jk} = \beta_{30k}$
$\pi_{4jk} = \beta_{40k}$

Level 3 (state level)

$\beta_{00k} = \gamma_{000} + u_{00k}$
$\beta_{10k} = \gamma_{100}$
$\beta_{20k} = \gamma_{200}$
$\beta_{30k} = \gamma_{300}$
$\beta_{40k} = \gamma_{400}$

In this model, the level-1 outcome, Y_{ijk}, was pre-service teachers' self-reported feelings of preparedness for teaching a particular area of early reading instruction, and the main predictor was pre-service teachers' knowledge about that area. All the other model specifications are identical to those used for Research Question 3b.

Again, an unconditional model without any control variables or predictor variables was estimated before applying the full model. Similar to the results for research question 3b, the results on variance components for this analysis also show significant variation on all three levels and therefore it was appropriate to use the three-level model to address this research question.

APPENDIX L.
ADMINISTRATION OF THE KNOWLEDGE ASSESSMENT TO EXPERTS AND NOVICES

The mandate for this study was to provide information on whether pre-service teachers are prepared to teach the essential components of reading instruction. The study team chose to address this question using the Knowledge Assessment described in Chapter 3. However, since this study represents the first large-scale use of this assessment, no benchmarks are available against which to judge the results. To provide context for the scores pre-service teachers obtained, we also administered the assessment to 34 reading experts (reading researchers and teacher educators) and 28 novices. Specifically, the additional reference groups were:[47]

- Reading Researchers: Professors at universities who are recognized for their contributions to reading research through publications in professional journals, positions on editorial boards for research publications, receipt of funding for large reading-related research projects, and positions on national panels such as the National Reading Panel (n = 15)[48]
- Reading Teacher Educators: Professors who teach courses about reading pedagogy in the pre-service teacher education programs in the largest of the colleges and universities in the sample (n = 19)
- Novices: Recent college graduates working in research at Optimal or AIR who do not work on reading-related projects and have no experience tutoring or teaching young children reading and no formal reading pedagogy courses (n = 28).

To identify potential reading researchers and reading teacher educators for participation, the study team first listed the 25 largest institutions that prepare teachers and determined full-time faculty members who taught pre-service reading "methods" courses or were actively involved in research related to early reading. The study team also listed other academics with specific prominence in early literacy research. In addition to websites of the 25 universities with the largest teacher preparation programs, the study team consulted sources such as published lists of individuals who develop the ETS early reading tests; early literacy researchers among the editorial boards of major peer-reviewed reading research journal (*Reading Research Quarterly*; *Scientific Study of Reading*); and listings of former members of the Center for the Improvement of Early Literacy Achievement (CIERA). From among these potential candidates, the study team selected researchers whose expertise was valued widely among research or consulting, organizations (e.g., RAND, AIR, National Research Council). Individuals who served on prominent committees considering early literacy, such as the National Academy's Research Council on the Prevention of Reading Difficulty in Young Children or the National Reading Panel, were also considered. The focus on the expertise of the researcher, not his or her institution, meant that the list of potential researchers contained individuals from sampled and non-sampled institutions. The study team excluded researchers who appeared to have potential conflicts of interest (e.g., they may have developed similar knowledge assessments). A review

[47] Unlike administration of the Knowledge Assessment to the pre-service teacher sample, these additional individuals took the assessment online.

[48] The study team excluded professors working at institutions from the main study sample.

of potential participants' published *curriculum vitae*, any posted information about courses these individuals teach at their home institutions, and available information about recent conference presentations (at AERA, the International Reading Association, or the National Reading Conference) resulted in their classification as reading researchers or reading teacher educators. The initial list, with connections to individual's websites or *curriculum vitae*, was submitted to the Department of Education for approval and final selection.

The recruitment process included sending each approved potential participant an email explaining the study and the process for administering the Knowledge Survey. Those who agreed to participate were sent a log-in number to a secure website and were allowed to take the assessment at any time within a two-week window. Each reading researcher or reading teacher educator who completed the assessment received an honorarium of $250.

The novice sample was selected based on relative work experience as a general policy/research analyst. All selected novices were employed by either AIR or Optimal at the time of the survey. Novices had no experience with reading instruction or with this study. Most novices were recent college graduates, and none had more than four years experience working within the field of policy research.

Table L-1 presents the results of the Knowledge Assessment for the experts, novices, and the pre-service teacher sample. Mean scores, mean percentages, median score, median percentages, and standard deviations are provided for the total assessment score, as well as the alphabetics, fluency, and meaning subscale scores. Results for the first two groups (reading researchers and teacher educators) are provided separately as well as combined into a group called "all experts."

Table L-1. Assessment scores, by each comparison group

	N	Mean Raw Score	Mean Percent Correct	Median Raw Score	Median Percent Correct	SD
All components (total score)						
All experts	34	42.21	79.6	43.00	81.1	5.68
Reading researchers	15	44.33	83.6	44.00	83.0	4.32
Teacher educators	19	40.53	76.5	42.00	79.2	6.16
Pre-service teachers	2187	30.21	57.0	31.00	58.5	5.92
Novices	28	25.11	47.4	25.00	47.2	4.24
Alphabetics						
All experts	34	14.35	79.7	15.00	83.3	2.24
Reading researchers	15	15.27	84.8	15.00	83.3	1.98
Teacher educators	19	13.63	75.7	14.00	77.8	2.22
Pre-service teachers	2187	9.53	53.0	10.00	55.6	2.64
Novices	28	8.61	47.8	8.00	44.4	2.33
Fluency						
All experts	34	10.06	83.8	11.00	91.7	1.67
Reading researchers	15	10.53	87.8	11.00	91.7	1.19
Teacher educators	19	9.68	80.7	10.00	83.3	1.92
Pre-service teachers	2187	7.31	60.9	7.00	58.3	2.06
Novices	28	5.65	47.0	5.00	41.7	1.87
Meaning						
All experts	34	17.79	77.4	18.00	78.3	2.99
Reading researchers	15	18.53	80.6	19.00	82.6	2.75
Teacher educators	19	17.21	74.8	17.00	73.9	3.12
Pre-service teachers	2187	13.36	58.1	14.00	60.9	3.04
Novices	28	10.86	47.2	11.00	47.8	2.66

NOTE: Data presented for "All experts" represents the sum of scores achieved by individuals designated as researchers and teacher educators.
SOURCE: *Study of Teacher Preparation in Early Reading Instruction*, 2007, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

As can be seen in Table L-1, the difference in total scores between expert subgroups (researchers compared with educators) was not statistically significant ($t = 2.03$, $df = 32$, $p > .05$).

Because of differences in how the samples of experts and novices and pre-service teachers were drawn, conducting statistical significance tests between these groups is not possible using the observed sample sizes (i.e., given that the expert sample was a simple sample and the pre-service teacher sample was a clustered multistage sample). As such, the study team calculated an adjusted sample size for the pre-service teacher sample to make the sample size comparable to a simple random sample. The effective sample size is equal to the actual sample size divided by the design effect (DEFF). The DEFF is equal to one plus the within-school intraclass correlation coefficient for the scale (ICC; see Tables I-3 to I-5) multiplied by the average cluster size minus one; $DEFF = 1+ICC*(n_{avg}-1)$. For the total score comparisons, an effective sample size of 540 was used for the pre-service teacher group.

As expected, there were significant differences in the total scores among the experts, pre-service teachers, and novices ($F(2,577) = 80.72$, $p < .01$). Follow-up contrasts showed that experts perform better than pre-service teachers ($F(1,550) = 136.07$, $p < .01$), and that pre-service teachers outperform novices ($F(1,544) = 28.54$, $p < .01$). Specifically, the expert group members, on average, answered 12 more questions correctly (22.6 percentage points) than pre-service teachers on the Knowledge Assessment. On the subscales, differences between the expert and pre-service teacher samples range from 26.7 percentage points on the alphabetics subscale to 19.3 percentage points on the meaning subscale.

Likewise, pre-service teachers did better than recent college graduates. On the Knowledge Assessment, pre-service teachers answered 5.1 more questions correctly (9.6 percentage points) than recent college graduates. Subscale differences ranged from 13.9 percentage points on the fluency subscale to 5.2 percentage points on the alphabetics subscale. To sum, pre-service teachers' scores on the assessment fall between the scores of novices and subject matter experts. These differences among the groups were significant for all three subscales: alphabetics ($F(2,599) = 57.56$, $p < .01$), fluency ($F(2,599) = 47.34$, $p < .01$), and meaning ($F(2,599) = 45.55$, $p < .01$).