EFFECTS OF A CURRICULUM-SPECIFIC STATE READING ASSESSMENT ON TEACHER PRINCIPLES AND METHODS

by

John C. White

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Education

Baltimore, Maryland
November 2022

© 2022 John Charles White
All Rights Reserved
Abstract

While U.S. literacy rates have improved over the last century, in recent decades progress in U.S. student reading levels have stagnated while other economically developed nations have progressed. Progress among adolescents has been particularly limited when compared with some improvements in younger U.S. readers in the last three decades. This problem is associated with in-school and out-of-school factors. Policymakers intended the advent of standards-based reforms to restore U.S. competitiveness in mathematics, reading, and science. While there is evidence of improvements in mathematics, there have not been corresponding improvements in reading. Furthermore, the standards-based tests that emerged have influenced teachers to emphasize discrete reading skills with no basis in evidence over helping students to discern the meaning of whole texts, which is fundamental to effective reading instruction for adolescents. In 2019, the state of Louisiana initiated a new Grade 7 reading test under the federal Innovative Assessment Design Authority (IADA) in which students respond to multi-select and constructed-response items directly related to texts they have read in the widely used curriculum ELA Guidebooks 2.0. This is a mixed-methods study of the implementation and outcomes of the IADA test. Implementation research questions examine the extent to which students participated in the test as intended and had necessary technology available to participate. Outcomes questions measure the extent to which IADA participation correlated with teacher methods or principles that differentiate teachers of students participating in IADA and other Louisiana teachers using ELA Guidebooks 2.0. Quantitative outcomes are dominant. To measure outcomes, the study collects data from an online survey administered to treated and non-treated teachers by
the Louisiana Department of Education in 2019-2020 and 2021-2022. Where there are significant differences in mean ranked responses among groups in a given method or principle in one year, the study indicates promising but incomplete evidence of a correlation. When there are significant differences in mean ranked response for both years, the study notes compelling evidence of a correlation. Qualitative data from 12 interviews of treated participants validate or mediate quantitative findings but do not establish findings unto themselves. While promising evidence of several differences in activities were present in one year, the two groups showed repeated significant differences in the influence of interim assessments provided outside the curriculum. Treated teachers were more likely to eschew such tests, relying instead on curriculum-embedded interim tests, which, qualitative evidence indicated, teachers trusted as an adequate basis for evaluating fitness for the IADA state test. This finding provides evidence of the potential for curriculum-aligned state tests to encourage teacher emphasis on comprehending the meaning of texts in a curriculum, rather than to emphasize fragmented skills extraneous to a curriculum or its texts.

*Keywords: standardized tests, curriculum, reading, ESSA, NCLB, Common Core*

Dissertation Advisor: Dr. David Steiner
Dedication

When I started this project in the spring of 2017, Katherine and I were months from being married, living in a New Orleans shotgun double with Ben the dog, commuting to work in schools across Louisiana and around the country. We are now parents to three children (and still Ben). We are inhabitants of new jobs and new (old) homes. A pandemic has come and gone. We have lived together in a sleepless state of incessant motion, a lifetime in five years, contained within the walls of home. What a blessing to have a partner like you. I am forever grateful for this phase of our lives together.
Acknowledgments

This project was made possible by an extraordinary array of people. A small set have helped me to propose, design, and implement this study. A larger set carried out Louisiana’s decade-long focus on high-quality instructional materials. This list of acknowledgments will sufficiently identify the former but not the latter, who are too many and whose contributions could themselves fill a book.

My dissertation advisor, David Steiner, is responsible for this project as much as I am. David is a friend, mentor, colleague, and routine source of admiration. His gift is not just his intellect, but also his belief that the thinking life is a blessing, and that a thinking person should wish that such blessings be bestowed upon others. I hope that this project has embodied that creed.

The committee members advising me each leant me unique talents, insights, and guidance. Ashley Berner’s brilliant writing and personal generosity have provided years of inspiration. Alanna Bjorklund-Young’s intellect and discipline provided me a standard for what it means to be a social scientist. Henry Smith, in his willingness to give students time for conversation and advice, has been a well of solutions whenever I presented him a problem.

My undergraduate thesis advisor at the University of Virginia, Charles Vandersee, passed on nearly 20 years ago. We had unfinished plans to publish articles together on the “isms” we were both helping our students to use as intellectual building blocks of inquiry and understanding. I would have liked Chuck to have read this study, even if it doesn’t qualify me for the English literature degree we’d so often discussed.

Julie Mikuta and the team at the Lynne and Charles Schusterman Foundation provided invaluable vision and support, both for this study and for Louisiana’s efforts to align its tests
CURRICULUM-SPECIFIC READING TEST

with high-quality instructional materials. Friends at the Baton Rouge Area Foundation, Bloomberg Philanthropies, the Hewlett Foundation, and the Walton Family Foundation similarly made early efforts to support Louisiana’s Innovative Assessment. Joanne Weiss was an instrumental advisor to many of the above and to me in the process.

The teams at Odell Education, NWEA, and the Center for Assessment were routinely helpful over many years, assisting me in understanding the assessment they had developed for Louisiana’s children.

Members of the Louisiana Board of Elementary and Secondary Education have continued to support the ongoing development of the Innovative Assessment, over many years and across different administrations.

I will resist trying to name the many colleagues at the Louisiana Department of Education who contributed to the state’s Innovative Assessment and to this study. An appropriate list would run into the hundreds. From finance to policy, from curriculum design to assessment administration, from field support services to data and analytics – the team at the Department embodied a rare blend of imagination and focus. It takes both to achieve what they achieved over the nine years we worked together.

In that time, I was invited into conversation with thousands of Louisiana teachers, school leaders, and district leaders. Most every good idea I ever had as state superintendent, and virtually every good idea our team had, came from these conversations. The grace I was shown by Louisiana’s educators in those years, and the gift of their insights, is something I carry with me.

Last, when each was newly born, Grace White, John Charles White, and Ella White got up before dawn with me to read and write, most days, for months on end. With age, each tired of
sitting calmly; at a certain point, it’s time to start rolling over and to babble. But I will cherish those first months of life, those quiet mornings together, and the gift of a baby smiling up at you as the sun rises.
# Table of Contents

*Abstract* ................................. ii  
*Dedication* .................................... iv  
*Acknowledgments* ............................... v  
*Table of Contents* ................................ viii  
*List of Tables* .................................. xi  
*List of Figures* ................................... xiii  
*Executive Summary* ............................. 1  
**Problem of Practice: Adolescent Reading in the U.S.** ............................................. 3
  Factors Associated with the Problem of Practice ....................................................... 3  
**Problem of Practice Situated in Louisiana Schools** ................................................... 7  
**Intervention: Innovative Assessment Design Authority Assessment** ............................... 8
  Interrupted Implementation ............................................................................................ 9  
  Evaluating Effects of IADA on Teaching Principles and Methods .................................... 11  
**Chapter 1: Literature Review** .................................................................................... 13  
**The Problem of Practice in Context** ......................................................................... 14
  Reading and College Readiness .................................................................................... 15
  Adolescent Reading ....................................................................................................... 16
  Louisiana Context ......................................................................................................... 20  
**Theoretical Framework** ............................................................................................. 21
  Human Capital .............................................................................................................. 24
  Social Capital ............................................................................................................... 25
**Macrosystem: Educational and Social Factors** ............................................................. 27
  Impact of Cognitive Skills on Economic Growth .......................................................... 27
  Relationships among Income Disparities, Social Networks, and Cognitive Skills ........... 30  
**Exosystem: Influence on Teaching and Parenting** ....................................................... 31
  Family Structure and Parenting .................................................................................... 31
  Influences on Teaching Approach ................................................................................ 33  
**Microsystem: Parenting Approach to Reading and Teaching Approach to Reading** ..... 50
  Parenting Approach ..................................................................................................... 50
  Teaching Approach ..................................................................................................... 53  
**Chapter 2: Needs Assessment** .................................................................................... 64  
**Research Questions and Methods** ............................................................................. 65
  Sample ......................................................................................................................... 66
<table>
<thead>
<tr>
<th>Chapter / Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>70</td>
</tr>
<tr>
<td>Discussion</td>
<td>72</td>
</tr>
<tr>
<td>Chapter 3: Intervention Literature</td>
<td>74</td>
</tr>
<tr>
<td>Implications of Prior Chapters for Potential Intervention Types</td>
<td>74</td>
</tr>
<tr>
<td>Review of Intervention Literature</td>
<td>75</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>76</td>
</tr>
<tr>
<td>Factors for Intervention</td>
<td>77</td>
</tr>
<tr>
<td>State Reading Tests</td>
<td>90</td>
</tr>
<tr>
<td>Chapter 4: Design and Methods</td>
<td>98</td>
</tr>
<tr>
<td>IADA Test Intervention</td>
<td>99</td>
</tr>
<tr>
<td>2019-2020 IADA Administration</td>
<td>102</td>
</tr>
<tr>
<td>2021-2022 IADA Administration</td>
<td>103</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>104</td>
</tr>
<tr>
<td>Research Design</td>
<td>109</td>
</tr>
<tr>
<td>Procedure</td>
<td>111</td>
</tr>
<tr>
<td>Measures and Instrumentation</td>
<td>119</td>
</tr>
<tr>
<td>Threats to Validity</td>
<td>126</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>127</td>
</tr>
<tr>
<td>Chapter 5: Results</td>
<td>133</td>
</tr>
<tr>
<td>Implementation</td>
<td>134</td>
</tr>
<tr>
<td>RQ 1: Student Participation in IADA</td>
<td>135</td>
</tr>
<tr>
<td>RQ2: Digital Device Access in IADA Schools</td>
<td>136</td>
</tr>
<tr>
<td>Outcomes</td>
<td>141</td>
</tr>
<tr>
<td>RQ3: Guidebook Usage</td>
<td>141</td>
</tr>
<tr>
<td>RQ4: Intentions in Teaching Texts</td>
<td>148</td>
</tr>
<tr>
<td>RQ5: Classroom Activities When Using Texts</td>
<td>152</td>
</tr>
<tr>
<td>RQ6: Factors Influencing Teaching Approach</td>
<td>165</td>
</tr>
<tr>
<td>Discussion</td>
<td>180</td>
</tr>
<tr>
<td>Results Synthesis</td>
<td>183</td>
</tr>
<tr>
<td>Implications</td>
<td>189</td>
</tr>
<tr>
<td>Limitations to Findings</td>
<td>190</td>
</tr>
<tr>
<td>References</td>
<td>194</td>
</tr>
</tbody>
</table>
CURRICULUM-SPECIFIC READING TEST

Appendix A: Needs Assessment Survey Form..........................................................221
Appendix B: Survey Form......................................................................................224
Appendix C: Interview Protocol............................................................................227
Appendix D: Interview Consent Form.................................................................229
Appendix E: Qualitative Codebook.....................................................................232
List of Tables

Table 1.1 Theoretical framework .................................................................22
Table 2.1 Weekly Hours Teachers Spent Selecting or Developing Curricular Materials........70
Table 2.2 Extent to Which Classroom Activities Were Influenced by State Tests................71
Table 4.1 IADA Grade 7 Test Form Differences..............................................100
Table 4.2 Outcomes Constructs and Related Research Questions..........................108
Table 4.3 Years of Teaching Experience, 2019-2020 and 2021-2022.......................115
Table 4.4 District Affiliation of IADA-treated Group, 2019-2020 and 2021-2022........117
Table 4.5 District Affiliations of Non-treated Groups, 2019-2020 and 2021-2022........117
Table 4.6 Process Evaluation Data Collection Matrix...........................................121
Table 4.7 Survey Items and Origins..................................................................122
Table 4.8 Outcomes Data Collection Matrix....................................................124
Table 4.9 Survey Item 7 Influence Factors.......................................................125
Table 5.1 Student Participation in 2019-2020 Grade 7 IADA....................................135
Table 5.2 Student Participation in 2020-2021 Grade 7 IADA..................................136
Table 5.3 Digital Device Access in 2019-2020 IADA-participating Schools...............137
Table 5.4 Digital Device Access in 2021-2022 IADA-participating Schools...............139
Table 5.5 IADA School Device Access, 2019-2020 and 2021-2022.......................139
Table 5.6 In a Typical Week, Frequency of Guidebook Instruction, 2019-2020 ...........142
Table 5.7 In a Typical Week, Frequency of Guidebook Instruction, 2021-2022 ..........145
Table 5.8 Percentage of Typical Lesson in Which Students use Guidebook Text, 2019-2020...144
Table 5.9 Percentage of Typical Lesson in Which Students use Guidebook Text, 2021-2022...144
Table 5.10 Preference to Teach Skills and Strategies versus Textual Meaning, 2019-2020......148
Table 5.11 Preference to Teach Skills and Strategies versus Textual Meaning, 2021-2022

Table 5.12 Teachers Endorsing Definitions of Understanding a Text, 2019-2020

Table 5.13 Teachers Endorsing Definitions of Understanding a Text, 2021-2022

Table 5.14 Percent of Time Students Engage in Grade-level versus Reading-level Texts in class, 2019-2020

Table 5.15 Percent of Time Students Engage in Grade-level versus Reading-level Texts in class, 2021-2022

Table 5.16 Frequency of Classroom Activities When Using Texts, 2019-2020

Table 5.17 Frequency of Classroom Activities When Using Texts, 2021-2022

Table 5.18 IADA-treated Teacher Classroom Activities Descriptive Statistics over Both Years

Table 5.19 Mean Influence of Factors over Lesson Plans, 2019-2020

Table 5.20 Mean Influence of Factors over Lesson Plans, 2019-2020

Table 5.21 Influence Factor Descriptive Statistics for Treated Teachers, 2019-2020 and 2021-2022
List of Figures

Figure 1.1 Theoretical framework demonstrating relationships among factors and levels........23

Figure 3.1 Theoretical framework demonstrating integrated model of teaching reading comprehension..........................................................................................................................77

Figure 5.1 Quantitative and qualitative results summary.................................................181
Executive Summary

The worst part of holding the memories is not the pain. It's the loneliness of it.

Memories need to be shared. (Lowry, 1993, p. 133)

Lowry’s *The Giver* (1993) tells the story of Jonas, a boy raised in an enclosed society in which knowledge of the past and of the outside world has been systematically eradicated in order to sustain pain-free lives of “sameness.” Appointed to the exclusive job of holding memories of authentic human experience on behalf of the community, Jonas is supernaturally endowed with repeated doses of knowledge and experience by his mentor, The Giver. Memories awaken in Jonas new levels of understanding, an instinct to question, and a capacity to distinguish right from wrong. Jonas ultimately breaks free of his home, to seek a world beyond, and to leave behind the knowledge he possesses.

*The Giver* is among several whole books taught as part of the ELA Guidebooks 2.0 curriculum developed by LearnZillion (n.d.) and the Louisiana Department of Education (LDOE). Organized into groups of texts that embody related knowledge—*The Giver* (1993) is grouped with other texts that relate to dystopia (LearnZillion, n.d.)—the curriculum builds a base of knowledge from one text that provides relevant background for comprehending other, related texts (LearnZillion, n.d.). The curriculum thus seeks a virtuous cycle in the reader: knowledge engenders comprehension, while the capacity to comprehend facilitates access to further knowledge (Pearson, 2013).

In this dissertation, I will address the problem of long-stagnant middle grade reading levels in the United States (U.S.) by evaluating Louisiana’s attempt to base its middle-grade reading tests on the content of a specific curriculum, ELA Guidebooks 2.0. I will assess in particular the effect that the new test has on how reading comprehension is taught in the state’s
classrooms. The study will evaluate on the degree to which teachers whose students experience
the curriculum-aligned test use the curriculum in class, help students to discern a text’s meaning
in class, use reading strategies in the service of this approach rather than as skills unto
themselves, and challenge students with texts and knowledge of appropriate and increasing
complexity.

This dissertation thus also relates to a pertinent tension for policymakers and educators,
between teaching and measuring specific skills valued in individuals and emphasizing commonly
held bodies of knowledge among a citizenry. As will be discussed in chapter 1, governments
regulate U.S. schools largely based on standards-based tests that measure the proficiency of
students in specific reading skills rather than their knowledge of texts, histories, or works of art.
Over the five years in which I have conducted this research, many U.S. universities have ended
the use of standardized skills tests as a factor in admissions decisions (del Rio, 2021, May 15);
many states and school boards have further sought to reinterpret the nation’s history (Fortin,
2021, November 8); and a global pandemic has exacerbated deficits in the skills, knowledge, and
well-being that schooling affords (McKinsey & Co., 2021). This evaluation, while not itself an
analysis of these political and social concerns, comes at a time of particular prominence for
debates among what a nation values in its humanities classrooms and what skills and knowledge
schools should be accountable for teaching.

Left with no shared memories, no beliefs forged as a people, community members who claim the
mantle of sameness in fact live in persistent isolation, committing regular acts of tyranny against
one another, cloaked in the comfort that to follow procedure, to simply do one’s appointed job, is
to be virtuous. Read as a parable, *The Giver* (1993) tells the reader that to find meaning in the
world requires knowledge of the world, and that a free society is one in which common knowledge is widely shared, via the written word and otherwise.

**Problem of Practice: Adolescent Reading in the U.S.**

Adolescence marks a transition for readers, from reading to learn the fundamentals of decoding words to reading in order to build knowledge (Chall, 1983). Reading to build knowledge is an act of comprehension, which is not simply deciphering the fixed meanings of multiple words but is an act of constructing a text’s meaning, from words within the text, from the structure of the text, and from sources of knowledge outside the text (Shanahan, 2005). On current measures of reading skill, adults in the U.S. trail those in peer nations (OECD, 2013). While 12% of American adults score at the highest level in reading on the internationally administered Survey of Adult Skills (OECD, 2013), top performing countries such as Japan and Finland register 22% of their population at this level. While 1 in 20 adults in Japan performs within the survey’s lowest 2 levels, in the U.S. 1 in 6 adults reads at the lowest levels (OECD, 2013). These results among adults are mirrored by outcomes among adolescents. According to the National Assessment of Education Progress (NAEP), U.S. performance in middle school and high school reading has stagnated over the last two decades (National Center for Education Statistics, 2020a).¹

**Factors Associated with the Problem of Practice**

---

¹ At the time of this dissertation’s completion, NCES had released 2022 nationwide, long-term NAEP outcomes for nine-year old children in reading. These were the first NAEP results released in the wake of the COVID-19 pandemic. Because these results are only for students younger than those whose classrooms are studied here, I refer throughout this chapter to NAEP results published in 2020 as the most recent, relevant statistics.
CURRICULUM-SPECIFIC READING TEST

To examine the large-scale problem of U.S reading levels, I use the model of an ecological system (Bronfenbrenner, 1994), in which interrelated manifestations of the problem occur at varying levels of scale and breadth. The macrosystem includes national and global factors influencing the problem; the exosystem includes factors more directly influencing teachers and parents; and the microsystem includes the actual interaction of parent or teacher and child. To create a full picture of U.S. adolescent reading, at each of these levels, I identify in chapter 1 the many factors associated with adolescents’ reading levels, using two broad categories of student-related constructs: skills that students possess, called human capital (Becker, 1963), and relationships they experience, called social capital (Coleman, 1988).

Income Inequality and Social Capital

In examining the role of social capital, the dissertation draws on literature that associates nations’ reading comprehension levels with three interrelated factors: national income inequality, the distribution of social capital with a nation, and national family formation trends. Nations’ average adolescent reading levels have a strong association with the extent of a nation’s family income disparities, which is notably high in the U.S. (Siddiqi, Kawachi, Berkman, Hertzman, & Subramanian, 2012). Income disparities also have an inverse association with social capital in a community or society (Kollmeyer, 2013; Robison, Siles, & Jin, 2011). Relatedly, the presence of single-parent households increases income inequality in a society (Kollmeyer, 2013). Amato (2005) estimates that the U.S. would achieve an annual reduction of more than 700,000 students 12-18 years old who repeat grade levels were the U.S. rate of single-parent households to return to 1960 levels.

With regard to reading, given that nuanced parental approaches to supporting early- and middle-grade reading skills have lasting effects on reading skill and motivation (Desjardin, 2003;
CURRICULUM-SPECIFIC READING TEST

Durik, Vida, & Eccles 2006; Myrberg & Rosen, 2009), having two parents in the home is associated with greater capacity for a nuanced parenting approach, greater resources with which to adopt the approach, or simply greater odds of a parental involvement (Martin et. al., 2007). Education policy and policymakers exist within this complex social context, neither fully limited by out-of-school realities nor able to ignore them.

Standards-based Reform

State-produced learning standards have historically possessed neither a vision for student capacities upon exiting high school nor a coherent sequencing of learned skills and knowledge necessary to achieve that end goal (NAE, 2009). Initiated in 2009, Common Core state standards (CCSS) established coherent bodies of interrelated skills and learning goals that could be shared across state lines (Coleman & Pimentel, 2012; Pearson, 2013). Years after their implementation, however, CCSS had also failed to generate its intended shifts in teaching approach across U.S. teachers of reading (Opfer, Kaufman & Thompson, 2016). Contrary to CSSS requirements for escalating text complexity from one grade level to the next, for example, most teachers continued the practice of selecting texts for students to read tailored to those students’ current reading levels (Opfer et al., 2016). Additionally, the majority of teachers of reading continued to use texts to teach students discrete, isolated skills, such as locating a text’s main idea, rather than assisting students with comprehending a whole text’s meaning, as also prescribed by CCSS (Opfer et al., 2016).

The No Child Left Behind Act of 2001 (NCLB) and its requirements for states to implement skills tests in Grades 3-8 were associated with positive and significant gains in mathematics achievement but also non-existent gains in reading achievement, after a decade of implementing the law (Dee & Jacob, 2011; Lee & Reeves, 2012). A broad body of literature
CURRICULUM-SPECIFIC READING TEST

indicates that standards-based assessments have significantly influenced the activities and content taught in U.S. middle schools (Au, 2007; Hirsch, 2016; Moon, Brighton, Jarvis, & Hall, 2007; Musoleno & White, 2010; Zoch, 2015). In particular, state testing regimes encouraged teachers to focus lessons on narrow, specific skills (Au, 2007; Kaufman et al., 2016; Hirsch, 2016). Hirsch (2016) attributes the failure of CCSS to achieve its intentions with respect to reading teaching methods to standards-based tests, which have continued to emphasize isolated, generic reading skills over students’ understanding of specific texts they have read.

Fisher and Frey (2008) suggest that strategies teachers provide students to help them comprehend texts—summarizing a passage, for example—have become “curricularized” and taught as skills (p. 16). Compelled by the influence of tests that ask students to perform discrete reading tasks such as summarizing or locating a main idea, these discrete tasks have become the organizing principle of teachers’ lessons (Fisher & Frey, 2008). The National Reading Panel (NRP; 2000) found that such strategies are only useful to students’ capacity to comprehend when used in the service of making meaning of a whole text; they are not evidence-based learning objectives unto themselves. The optimal environment for comprehension instruction, is instead a “naturalistic setting” (NRP, 2000, p. 4-6) in which the reader and teacher alike are engaged in trying to figure out what the text actually means, using reading comprehension strategies in combination with one another, as means to that end. In chapter 1, I will synthesize existing literature to establish the pillars of an evidence-based approach to teaching reading comprehension to the adolescent reader: one focused on a whole text’s meaning, using reading strategies in the service of discerning meaning, applying knowledge from outside the text to construct meaning, and increasing over time the complexity of both the text and the knowledge
required to make sense of it. I will further establish a disconnect between such an approach and the U.S. testing regime as currently conceived.

**Problem of Practice Situated in Louisiana Schools**

Louisiana’s approach to implementing CSSS, including curriculum, professional preparation, and ongoing support for teachers, has been uncommonly assertive among states (Doan, Kaufman, Woo, Prado Tuma, Dilberti & Lee, 2022; Kaufman, Cannon, Culbertson, Hannan, Hamilton, & Meyers, 2018; Opfer et al., 2016; Pondiscio, 2017). In 2014 the state commenced a rolling review of ELA and math curricula, placing each into tiers based on alignment with state standards and other factors (Doan et al., 2022; Kaufman et al., 2018; Pondiscio, 2017). The state then provided financial incentives for school systems acquiring curricula evaluated at the highest level (Doan et al., 2022; Pondiscio, 2017). When the state review process did not result in highly rated middle school ELA curricula, the state commenced a partnership with LearnZillion, a private publisher, to create ELA Guidebooks 2.0 that is now in wide use across the state (Will, 2017).

While 70% of teachers nationally selected text based on the reading level of students, reducing text complexity when readers struggle, only 47% of Louisiana teachers identified with this practice, choosing instead to select texts appropriate to students’ grade level, per CCSS guidance (Kaufman, Opfer & Thompson, 2016). Similarly, while 75% of ELA teachers in the nationwide sample organized lessons around teaching discrete, isolated literacy skills, only 49% of Louisiana teachers identified with the practice (Kaufman et al., 2016). Instead, 49% of Louisiana ELA teachers identified texts and organized lessons around helping students to discern the meaning of a text, while this was only true for 21% of teachers elsewhere (Kaufman et al., 2016).
Nevertheless, the percentage of Louisiana teachers Kaufman et al. (2016) found not to have adopted CCSS guidance in their principles and methods was still sizable. This was later validated by a needs assessment I conducted in 2018, in which a survey of Ouachita Parish, Louisiana, teachers found that 64.8% tended to organize lessons around particular reading skills rather than around the meaning of texts. Asked about the activities included in lesson plans, 93.3% reported including skills such as summarizing and locating the main idea of a text either often or always.

Similar to findings from other studies (Au, 2007; Musoleno & White, 2010), 98.1% of respondents to the needs assessment survey I created indicated that they often or always ask students to respond to questions similar to those on Louisiana’s state assessment when engaged in reading a text. Teachers reported being very significantly or significantly influenced by the content of state standards (98.1%) in choosing classroom activities and texts as well as generally high levels of influence (94.3%) from the state assessment. The needs assessment, which is described in chapter 2, pointed to a common theme: in spite of some successes in shifting principles and methods toward CCSS guidance, Louisiana teachers still labored under the influence of tests that may have been counteracting some of the direction sought by the standards and by state leadership.

**Intervention: Innovative Assessment Design Authority Assessment**

In 2018, the U.S. Department of Education (USDOE) granted the state of Louisiana permission to conduct an experimental middle grade reading test under the Innovative Assessment Design Authority (IADA) provision of the Every Student Succeeds Act (ESSA) of 2015, replacing for some districts the state’s conventional end-of-year reading test (Sentell, 2018). Louisiana’s IADA assessment, which launched in the 2019-2020 school year for Grade 7
students in participating schools, was a companion to ELA Guidebooks 2.0. Students completing the IADA test responded to passages from Guidebooks 2.0 texts they had already read in class or passages from texts with thematic connections to Guidebook 2.0 texts, rather than responding to texts with no relationship to books read during the school year (LDOE, 2019).

**IADA Test Form**

As part of the Grade 7 IADA assessment, students completed three assessment forms over the course of the school year. Students first answered a series of multi-select questions to demonstrate knowledge of texts read in the relevant Guidebooks 2.0 unit or to respond to texts demonstrative of related knowledge (LDOE, 2021). Students subsequently wrote an end-of-unit essay responding to a prompt requiring that they draw on textual knowledge from the relevant Guidebooks 2.0 unit (LDOE, 2021). To conclude the school year, the Grade 7 IADA test also included a summative essay in which students draw on knowledge gleaned from across the three units to respond to a prompt. Parents and school systems received three periodic reports annually on students’ knowledge of Guidebook 2.0 texts, application of that knowledge to other outside contexts, and synthesis of knowledge developed over the course of the year (LDOE, 2021). In chapter 4 I describe the test form and its administration in detail.

**Interrupted Implementation**

In the 2019-2020 school year, Grade 7 students in ten Louisiana school districts completed end-of-unit assessment forms in the October and February testing periods (C. Johnson, personal communication, July 27, 2019). However, the school year was subsequently interrupted due to the COVID-19 pandemic, and the test’s full administration was not completed. In the following school year, 2020-2021, Louisiana resumed conventional statewide testing (Sentell, 2021), but the IADA was not administered given the complexities of frequently
interrupted schooling during a second year of pandemic impacts (C. Johnson, personal communication, December 14, 2021). In the 2021-2022 school year, the Grade 7 IADA assessment was fully operational for seven participating school districts (C. Johnson, personal communication, December 14, 2021).

Just as the IADA test itself was disrupted due to the COVID-19 pandemic, this study has been complicated and extended by the disruption. I include in the study the interrupted IADA test administration and associated data collection undertaken winter 2020, even though neither the three-part test series nor the original plan for data collection was completed in 2020. I then include in the study the IADA implementation and associated data collection from 2021-2022, the IADA’s first complete test administration cycle.

The effects of the pandemic on schooling globally are already evident. By initial estimates, the problem of adolescent reading in U.S. schools itself has been greatly exacerbated (McKinsey & Co., 2021). By other estimates, the U.S. teacher workforce has been strained and has experienced abnormal levels of turnover (National Center for Education Statistics, 2022). This is not a study of the pandemic’s impacts; I treat the 2019-2020 and 2021-2022 IADA implementation and data collection cycles as part of one, long-term intervention that was conceived before the global pandemic. Nevertheless, the context of the pandemic is present throughout the study, and I attempt to point out potential limitations, implications, and opportunities related to the pandemic as they arise. In particular, when I discuss findings, I emphasize the pandemic context, both as a potential limitation to explaining the study’s results and as an additional reason for urgent improvements in the way U.S. schools teach reading and support teachers.
Evaluating Effects of IADA on Teaching Principles and Methods

This study evaluates both the implementation and outcomes of IADA in Louisiana. Two research questions focus on the reach and fidelity of the IADA in two school years.

- **RQ1**: Are all eligible Grade 7 students in IADA-participating school districts consistently completing all available administrations of the assessment?

- **RQ2**: Do schools with participating teachers and students demonstrate adequate levels of technology readiness for implementing the IADA assessment?

Outcomes research questions focus on teachers’ inclination toward using texts, their intentions when teaching texts, and the sources of influence over their decisions as to how they use texts.

- **RQ3**: Do Grade 7 ELA students taking a state reading test focused on pre-determined texts use those texts in class with greater frequency than do peer students?

- **RQ4**: To what extent are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more likely than peer teachers to design lessons to teach the meaning of those texts rather than discrete reading skills?

- **RQ5**: Are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more or less likely than peer teachers to have students discuss or write about the significance of textual features; access texts of appropriate grade-level complexity; or discuss or write about how texts relate to other texts or subject matters?

- **RQ6**: What factors do Grade 7 ELA teachers identify as influential in how they use texts in the classroom?
This is a mixed-methods study involving correlational analysis of quantitative data and exploratory analysis of qualitative data (Mertens, 2018). The study uses standard reports generated by the LDOE to evaluate student participation, administrative irregularities, and schools’ technology readiness levels in IADA implementation. In determining outcomes, quantitative results are dominant. The study uses data gathered from an LDOE survey that was administered annually in two years to evaluate teacher principles and behaviors. The study then uses interviews of teachers whose students complete the IADA test in order to validate quantitative outcomes and to explore possible mediating factors associated with quantitative outcomes. This qualitative aspect of the study is exploratory in nature and is meant to validate quantitative findings, to identify related factors, and to suggest further lines of inquiry. The resulting study both tells the story of a comprehensive, multi-year, state implementation and measures constructs related to a specific policy intervention that is replicable in other states, with implications for both federal and state policymakers.
Chapter 1

Literature Review

In the century and a half since the American Civil War, functional illiteracy, the inability to identify and assign meaning to commonplace words, has gone from being a regular condition in the United States (U.S.) to being virtually nonexistent (Snyder, 1993). While 20% of all Americans were illiterate as of the late 19th century, by the end of the 20th century the figure was below 1% (Snyder, 1993). While 16% of African Americans were illiterate as recently as 1930, by 1980, 1% were illiterate (Snyder, 1993). As the 20th century progressed, an increasing share of the American population attended public schools, and the typical number of years spent in school increased as well (Snyder, 1993). Public schools educated masses of Americans to read at a level necessary for daily work and life in their time.

The demands of the contemporary workplace, however, call for a more sophisticated set of literacy skills than previous generations of Americans needed to possess. According to Murnane, Sawhill, and Snow (2012),

Advanced literacy is a prerequisite to adult success in the twenty-first century. By advanced literacy we do not mean simply the ability to decode words or read a text, as necessary as these elementary skills are. Instead we mean the ability to use reading to gain access to the world of knowledge, to synthesize information from different sources, to evaluate arguments, and to learn totally new subjects. (p. 3)

Far beyond functional literacy for daily life, literacy skills appropriate to the contemporary age involve readers independently determining the meaning and significance of what they have read.

By this definition of reading proficiency, the U.S. trails its global peers (OECD, 2013). While 12% of American adults score at the highest level in reading on the internationally
CURRICULUM-SPECIFIC READING TEST

administered Survey of Adult Skills (OECD, 2013), top performing counties such as Japan and Finland register 22% of their population at this level. While 1 in 20 adults in Japan performs within the survey’s lowest 2 levels, in the U.S. 1 in 6 adults reads at the lowest levels (OECD, 2013). Thus, while the U.S. has succeeded in virtually eradicating fundamental illiteracy, the U.S. population does not possess advanced reading skills on a par with peer global economic leaders.

This problem runs counter to the progress the nation achieved in growing the foundational literacy skills and education levels of its people over the course of the 20th century, largely by way of its public schools (Snyder, 1993). It is remarkably distant from Horace Mann’s (1855) vision of an educated nation in which “no community should rest contented… while it is inferior to its own capabilities” (p. 14). And while it is a multi-dimensional phenomenon that touches on both social and educational factors (OECD, 2013), it is reasonable to consider what public policy responses might compel progress on a par with that made in past generations. This chapter explores multiple dimensions of this nationwide problem of practice with an eye toward understanding replicable, evidence-based options policy makers have for a large-scale response.

The Problem of Practice in Context

According to the National Assessment of Education Progress (NAEP), U.S. performance in middle school and high school reading has stagnated over the last two decades (National Center for Education Statistics, 2020a). U.S. performance on PISA in reading has in fact

2 Because the problem of practice is longstanding and widespread, this chapter proceeds without accounting specifically for projected impacts of the pandemic on reading levels. It is reasonable to suspect that the problem of practice has only been exacerbated by the pandemic, to extents not understood at this time. As noted in the executive summary, the dissertation will thus occasionally discuss potential impacts of the pandemic on reading levels, as well as potential
declined over the last two decades, as has the U.S. place in international rankings of PISA reading proficiency (OECD, 2015). This problem of practice related to American middle school and high school students contrasts with some demonstration of progress (National Center for Education Statistics, 2020b) made by U.S. Grade 4 students over this period of time. Average U.S. Grade 4 performance increased 7 scale points between 2000 and 2019, with African-American and Latino students making greater gains than the national average (National Center for Education Statistics, 2020b). Grade 8 scale scores by comparison showed no improvement (National Center for Education Statistics, 2020a). These desultory results in middle schools and high schools present a vexing problem of practice for educators and policy makers concerned with literacy skills of the American people.

**Reading and College Readiness**

The problem of stagnant adolescent literacy levels carries implications for education attainment, prosperity, and equality among U.S. residents. Between 1990 and 2013, employer demand for U.S. workers with postsecondary credentials increasingly outstripped the supply of workers with postsecondary credentials (Carnevale & Rose, 2014). In the same period, Americans holding associate and bachelor’s degrees experienced higher rates of wage increase than did those holding high school diplomas alone (Carnevale & Rose, 2014).

Studies have repeatedly demonstrated that prior academic performance is predictive of postsecondary credential completion. High school grade point average is associated with the

probability of a student completing associate and bachelor’s degrees (Porchea, Allen, Robbins, & Phelps, 2010; Radunzel & Noble, 2012). The rigor of the curriculum completed by high school students, as measured by the number of advanced courses a student completes, is associated with rates of college-degree completion (Adelman, 2006). Student performance on the ACT and SAT, college readiness and admissions examinations, is also associated with two- and four-year college persistence and completion (Lotkowski, Robbins, & Noeth, 2004; Porchea, Allen, Robbins, & Phelps, 2010; Radunzel & Noble, 2012).

Reading proficiency underlies each of these indicators of college readiness. Students who struggle to comprehend what they read have difficulty mastering high school content necessary for strong grade point averages (Joftus, 2002). Low-performing readers struggle in challenging courses that require reading complex texts and are often discouraged from pursuing challenging curricula (Au and Raphael, 2000). Postsecondary education necessitates that students not only be able to recall what they have read, but also that they draw inferences, recognize patterns and structures, and be able to form critical arguments in response (Kamil, 2012). As Murnane et al. (2012) write, “higher-level skills are now essential to young Americans who wish to explore fields as disparate as history, science, and mathematics; to succeed in postsecondary education, whether vocational or academic; to earn a decent living in the knowledge-based globalized labor market; and to participate in a democracy facing complex problems” (p. 3). A citizenry dependent on the education level of its adults for future prosperity has a great deal riding on the reading levels of its adolescents.

**Adolescent Reading**

The transition out of elementary school and into secondary grades marks a change in the school-based reading experience for students, as most students no longer need to be taught
foundational skills aimed at fostering the ability to decode a word and to read a series of words fluently. A commonly held “inoculation fallacy” (Snow & Moje, 2010, p. 1) pervades middle schools and high schools, holding that elementary grade students who demonstrate proficiency in these foundational reading skills require no further explicit or formal reading instruction in order to make sense of what they read as they age. Snow and Biancarosa (2003) summarize this mindset: “If the basics of ‘decoding’—or accurately sounding out the words on a page—are in place, shouldn’t comprehension simply follow?” (p. 6).

The disparity in progress made by U.S. Grade 4 students and U.S. Grade 8 students (NCES, 2016) provides some indication that this mindset has failed American adolescents. “Notwithstanding the laudable efforts…. to bring all children to similar levels of reading skill by end of 3rd grade,” Snow and Biancarosa (2003) write, “it would be foolhardy to think that these efforts, even if wildly successful, will by themselves eliminate reading failure or the achievement gap in the middle and high school years” (p. 6). The demands on middle school and high school readers are distinct from demands they experienced in early grades. Students in middle schools and high schools experience a curriculum, for example, divided into academic disciplines, each of which require specific knowledge and vocabulary in order to read proficiently (Shanahan & Shanahan, 2008). Students are further required to read across genres, to draw on a larger vocabulary and a larger body of knowledge, to read for a wider variety of purposes, to form critical opinions in response to what they have read, and to muster the motivation necessary to do much of this independently (Snow & Biancarosa, 2003). If the primary grades involve learning to read, the middle and high school grades necessarily involve “reading to learn” (Chall, 1983). Proficiency as an adolescent reader is thus defined by the student’s capacity to understand, integrate, and apply what has been read across varying disciplines and toward varying objectives.
Comprehension

At the heart of this shift in the orientation of a student’s reading life is the notion of reading comprehension. To comprehend a thing, according to Merriam-Webster (n.d.), is “to grasp the nature, significance, or meaning.” Durkin (1993) defines comprehension as “intentional thinking during which meaning is constructed through interactions between text and reader” (p. 5). Words signify ideas or things; to discern these meanings is, according to these simple definitions, to comprehend words.

It is notable that within the simple definition, Durkin (1993) depicts comprehension as an act of construction. Shanahan (2005) similarly describes comprehension as an act of assembly: “a form of active and dynamic thinking [that] includes interpreting information through the filter of one’s own knowledge and beliefs, using the author’s organizational plan to think about information (or imposing one’s own structure on the ideas), [and] inferring what the author does not tell explicitly as well as many other cognitive actions” (p. 1). Here the reader calls on an understanding of words on the page, of the manner in which a document is organized, and of related knowledge outside the text, to ascribe meaning to the text.3

As Murnane et al. (2012) note, this vision of what comprehending readers do eclipses the “simple view” that reading is a function of decoding skill and reading fluency skill. In this more complex view, adolescent students read not just to discern words on a page but to build and store knowledge.

---

3 The distinction between sources of knowledge within the text and knowledge from outside the text will inform a general framework for proficient reading comprehension among adolescents later in this chapter. To the extent that these discussions rely on students’ efforts to discern a text’s meaning, I hew to the notion of meaning in the denotative and connotative sense: what the words say and what they imply to an adolescent reader. I do not address the notion of reading for meaning in the universal or theoretical sense. Thus I do not address prominent critical and intellectual theories concerned with textual meaning. For a survey of literary theory, see Literary Theory by Terry Eagleton (University of Minnesota, 2008).
knowledge – of words, of how texts are structured, and of facts – initiating a virtuous cycle in which stored knowledge from what one has read informs what one is able to comprehend in the future (Chall, 1983). From this complex notion of comprehension as an act of constructing meaning from knowledge within and without a text emerges the possibility that meaning itself can grow in complexity and sophistication, as what the reader knows – her vocabulary, her familiarity with textual structures, and her background knowledge of the world – expands.

Vygotsky’s (1978) “zone of proximal development” (p. 86) provides a framework for considering how a reader grows in this capacity to construct meaning from a text. The zone of proximal development, which is “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined by the level of problem solving under adult guidance or in collaboration with more capable peers” (p. 86) represents a challenge of complexity—in individual words, in textual structure, or in contextual knowledge—that calls the student to assemble ever more nuanced meaning from a text. It is through interaction, with peers and with the teacher, that the student navigates this complexity, forming a sense of meaning that eclipses prior comprehension (Vygotsky, 1983). The act of making meaning, according to this framework, is not a comprehension skill that either exists or does not exist; it is an evolving capacity that may grow to further levels of complexity.

The emerging picture of adolescent reading is essential to this dissertation. With age comes a transition for readers, from reading to learn the fundamentals of decoding words one

---

4 I discuss teaching methods to stimulate growth in comprehension later in this chapter. I present Vygotsky here to establish the principle that comprehending meaning is not a fixed skill but a growing capacity that takes on greater levels of complexity given appropriate complexity and interaction.
CURRICULUM-SPECIFIC READING TEST

reads to reading in order to build knowledge (Chall, 1983). Reading to build knowledge is an act of comprehension, which is not simply deciphering the fixed meanings of multiple words but is an act of constructing a text’s meaning, from words within the text, from the structure of the text, and from sources of knowledge outside the text (Shanahan, 2005). That readers possess inevitably varying complexity in their vocabulary, their knowledge of textual structure, and their background content knowledge, and that texts vary in the complexity they present, implies that the complexity of meaning that readers construct from a text may vary, and that an individual’s capacity to make meaning of a text is a developing function, capable of evolving and growing toward greater levels of complexity itself.

**Louisiana Context**

I formerly served as state superintendent of education in Louisiana, where nearly 700,000 students, roughly 2% of the national student population, annually attend public schools (Louisiana Department of Education, 2017a). Between 2003 and 2019, average NAEP reading performance among Louisiana Grade 4 students increased from 205 scale points to 220 points (National Center for Education Statistics, n.d.). However, average Grade 8 reading scores showed no change over the course the same period (National Center for Education Statistics, n.d.). This pattern of relatively low reading performance in secondary schools continued into high school; in 2017, 36% of Louisiana high school graduates achieved ACT college benchmark standards in reading (ACT, 2017), indicating strong likelihood of successfully completing first-year community college or university coursework. This is troublesome not only for graduates not going on to attend college, but also for the 57% of graduates who enroll in college (Louisiana Department of Education, 2016), many of whom have not achieved reading proficiency levels predictive of completing postsecondary courses. Given previously discussed relationships
among reading proficiency, student success in high school courses, and college completion, it should not be surprising that only 27% of Louisiana residents ultimately complete a bachelor’s or associate degree, nearly 10% below the national average rate (Statistical Atlas, 2018). The problem of stagnant adolescent literacy levels in the U.S. manifests itself in troubling ways in Louisiana.

**Theoretical Framework**

Because the problem of stagnant adolescent reading levels is a nationwide and longstanding phenomenon, it should be examined via the broad lens of both in-school and out-of-school activity. It should also be examined at varying levels of granularity, from nationwide and global conditions to individual actions and interactions. Ecological systems theory (EST) provides a framework for examining child development through interconnected layers of society, allowing for observation not only of a problem from multiple vantage points, but also of relationships across those vantage points (Bronfenbrenner, 1994). The theoretical framework for this dissertation examines the problem of practice at three levels of an ecosystem: the macrosystem, which includes global and domestic trends, events, and conditions; the exosystem, which includes policies, trends, and events with direct impact on teachers, parents, and children; the microsystem, which includes the actions and interactions of parent and child or teacher and student. As detailed in Table 1.1, at each level of these three levels, I examine the problem through the lenses of human capital and social capital.\(^5\)  This theoretical framework allows for

\(^5\) While I emphasize schools’ role in fostering human capital and families’ role in generating social capital, these are related constructs and are not exclusively associated with particular settings or institutions. Schools, for example, may be sources of social capital; families may be sources of human capital. They are separated here for purposes of identifying the broadest array of factors associated with schools, families, and adolescent reading levels.
inquiry into the nature of the whole education system and the whole of U.S. society, just as it allows for incisive inquiry into classrooms and homes. It permits vantage points wide and narrow, and it examines the problem of adolescent reading levels from both social and educational perspectives.

*Table 1.1 Theoretical Framework*

<table>
<thead>
<tr>
<th>Construct</th>
<th>Macrosystem</th>
<th>Exosystem</th>
<th>Microsystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human capital</td>
<td>Education policies and investments seek increased skills and economic returns.</td>
<td>Education policies influence schools’ and teachers’ approaches to teaching reading.</td>
<td>Adolescent reading gains are associated with particular teaching approaches.</td>
</tr>
<tr>
<td>(Internal to schooling)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social capital</td>
<td>Strength of social network is associated with income disparity and cognitive skill, such as reading.</td>
<td>Changes in social networks, such as family formation and structure, influence parents’ approach to reading.</td>
<td>Adolescent reading gains are associated with particular parenting approaches.</td>
</tr>
<tr>
<td>(External to schooling)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The constructs reflected in the theoretical framework are dynamic parts of a system and are not isolated or contained. They exert influence on one another, across levels of the system. Figure 1.1 represents this dynamism, demonstrating a complex system with multiple, interactive factors.
Ecological Systems Theory

EST includes the concept of proximal processes, relations between the child and an individual caregiver, which constitutes a primary influence over child development (Bronfenbrenner & Morris, 2006). But Bronfenbrenner (1994) also notes that child development should be analyzed through interrelated contexts that operate at varying levels of proximity to the child. EST can thus serve as a schema through which to identify potential factors in a problem of practice and to establish relationships among those factors.6

As outlined in Table 1.1, adolescent reading can be analyzed through the lens of both human capital and social capital at each level of the ecological system. The macrosystem of adolescent reading skills in the U.S. includes, for example, broad trends in both education policies meant to build human capital and social conditions that can be evaluated by way of their

---

6 EST includes five levels of influence on child development. The theoretical framework for this literature review focuses on three levels but does not analyze the mesosystem, involving interactions among influencers on the problem, and the chronosystem, involving changes in influencing factors over time.
relation to social capital. The exosystem similarly includes in-school factors such as teacher skills, academic standards, curricula, and assessments, as well as out-of-school factors related to family structure and work life. The microsystem includes individual teachers’ approaches to reading and parents’ approaches to reading, with the term “approach” notably used to capture not only an individual’s methods but also the principles that compel the teacher to use such methods.

At every level of this ecosystem, this chapter makes the case that the problem of stagnant adolescent reading skills in the U.S. is related both to trends in the nation’s educational policies—the nation’s effort to increase human capital—and to trends in the nation’s social fabric, as measured by social capital.

**Human Capital**

Human capital is the economic value of skills and knowledge possessed by human beings (Becker, 1962). Because investment in skills generates a return of human capital, investment in the formal education system can be viewed as an input seeking an economic output. Governments, schools, and taxpayers make investments of resources in order to produce skilled individuals, who in turn generate economic returns (Becker, 1962).

Viewing the education system through the lens of human capital investment, economists can evaluate education policies based on their comparative returns. Having a teacher regularly associated with significant gains on standardized tests is associated with lifelong economic returns for students (Chetty, Friedman, & Rockoff, 2014). Chetty et al. (2014) assert that a policy of replacing teachers performing in the bottom 5 percentiles of test score returns with teachers performing at mean levels would yield lifelong income gains for impacted students of

---

7 As this is a synthesis of relevant literature rather than an in-depth analysis of particular sources, the discussions of Becker and Coleman in this section are intentionally brief.
$250,000. Public investment in formal education within a child’s earliest years is another example of a policy returning lasting economic benefits to participants (Heckman, 2010).

As discussed earlier, U.S. education policies have changed significantly over the course of the last 25 years, with the advent of state standards for student learning, state standardized tests, and rating systems for schools and educators. At each layer of the ecosystem, this chapter will outline the relationship between policies that have sought to increase the nation’s human capital, including the reading proficiency of its adolescents, and the problem of stagnant adolescent reading performance.

**Social Capital**

Social capital is the value to individuals generated by the relationships, interpersonal networks, and affiliations they maintain (Coleman, 1988). While individuals may be self-interested actors, it does not follow that they seek only economic benefit for themselves (Coleman, 1988). Instead individuals can be motivated to enhance the assets of those with which they are related or associated. Changes in the value attached to these relationships, attributable to changes in larger social networks or institutions, are changes in the social capital possessed by or available to individuals (Coleman, 1988). The concept of social capital can thus serve as a basis for understanding the effects of changes in social relationships, including effects on reading skills, as discussed later in this chapter.

Coleman (1988) established the presence of social capital through a study of high school dropout rates among students attending public schools, Catholic schools, and non-Catholic private schools. Controlling for socioeconomic status, academic history, family structure, and religious affiliation, Coleman (1988) found that students attending Catholic schools were significantly less likely to drop out of high school than were similar peers in public schools and
non-Catholic private schools. Coleman (1988) similarly found that regular attendance at religious services, Catholic or otherwise, predicted a reduction in school dropout rates. These findings Coleman (1988) attributes to the social capital added to the lives of students by religious institutions and Catholic schools.

Putnam (2000) asserts that social capital has not only direct effects on individuals but also external effects on individuals not immediately associated with the given social network. Strong bonds among neighbors, for example, may be associated with lower rates of crime, benefiting anyone who visits the neighborhood. While the strength of social networks does not necessarily reduce economic inequality and may exacerbate it (Helliwell & Putnam, 2004), the density and strength of social networks at the nationwide level correlates with a broad range of social welfare and public health indicators (Putnam, 2000). The level of trust within and across a nation’s social networks is thus a factor related to nationwide social outcomes. This chapter will associate broad social trends that contribute to a nation’s social capital, such as changes in family structure and formation, with resulting impacts on children and young adults, especially the presence and approach of parents, that may affect reading levels.

This chapter’s emerging examination of education policies on one hand and social trends on the other would appear to set up an ideological conflict. Former New York City schools chancellor Joel Klein (2009; I was deputy chancellor under Klein at the time), summarized and embodied this conflict in an opinion article: “America will never fix education until it first fixes poverty—or so the argument goes. In fact, the skeptics of urban schools have got the diagnosis exactly backward. The truth is that America will never fix poverty until it fixes its urban schools.” Klein (2009) described education as a contributing factor to social conditions. But the prevailing ideology, he asserted, was the reverse: one’s social conditions determine one’s human
capital. By intertwining a social capital strand and a human capital strand at each level of an ecosystem, this chapter will paint a less absolutist picture. Our nation’s efforts to develop human capital can be evaluated not for their capacity to repair fully the negative effects of social trends, but for their successes in expanding skills and opportunities in spite of those social trends. It is that relationship to which the chapter now turns.

**Macrosystem: Educational and Social Factors**

The macrosystem of adolescent reading performance in the U.S. includes relationships among educational factors, social factors, and economic factors. This section discusses research on the relationship between cognitive skills and economic growth. It also discusses the relationship between skills and the distribution of family income within nations. And it discusses the relationships both skills and income distribution have with the strength of social networks in a given nation, including but not limited to family networks. The picture painted of the macrosystem is not one of linear or exclusive relationships between specific factors and the problem of practice. Instead, this section paints a complex picture of interconnected factors underlying societal trends in reading levels. According to this rendering of the macrosystem, social, economic, and educational factors influence one another as part of an interconnected system, a troubled one in this case, implying salient interventions may exist both within schools and without.

**Cognitive Skills and Economic Conditions**

Cognitive skill is a form of human capital that includes knowledge and intellectual capacities measurable by standardized tests (Hanushek, 2015). International tests of cognitive skills in math, reading, and science provide a basis on which the cognitive skills of nationwide populations can be compared (Barro, 1999; Hanushek, 2016). By associating longitudinal gains
by countries on the PISA scale with longitudinal change in GDP, Hanushek and Woessmann (2010) posit that increases in cognitive skills are associated with nationwide economic gains. The researchers create a model to show hypothetical economic effects of cognitive improvement across nations. According to the model, the U.S, which falls slightly below the mean in both economic gains and cognitive improvement between 1963 and 2003 (Hanushek & Woessmann, 2010), would see an increase of $40 billion in GDP with a 25-point PISA increase.\(^8\) Relatedly, researchers associate socioeconomic factors, including family income and parental education attainment, with cognitive skill levels in children and education attainment (Bailey & Dynarski, 2011; Coleman, 1966). Children raised in low-income households are more likely than peers to demonstrate low cognitive skills, including reading levels, on standardized assessments (Bailey & Dynarski, 2011; Coleman, 1966). Economic conditions are associated with cognitive skills across national populations and among individual children.

**Cognitive Skills and Economic Conditions in the U.S.**

Hanushek, Peterson, and Woessmann (2014) compare the PISA performance of U.S. students with parents of various educational backgrounds to the performance of international students whose parents have the same educational backgrounds. At each level of parental education, the rate of U.S. students performing at proficient levels falls short of international averages (Hanushek et al., 2014). According to this analysis, U.S. schools rank no higher in educating the children of advantaged families than they do in educating the children of

---

\(^8\) In a less frequently cited paper, Komatsu and Rappleye (2017) refute the internal validity of Hanushek and Woessmann’s (2010) claims to a significant relationship between cognitive skills and GDP. Researchers present correlational evidence of an inverse relationship between nations’ 2006 PISA performance and students’ views on the economic value of education, as reflected on surveys. It may be the case, they suggest, that a prevalent view among students that the value of education lies in its individual economic returns stunts students’ long-term commitment to learning, especially when economic returns are not realized.
disadvantaged families, when compared with other nations. Petrilli and Wright (2016) plot the internationally ranked PISA performance of nations’ wealthy students against the internationally ranked PISA performance of nations’ poorest students. The analysis shows a correlation in the ranked performance of nations in educating their poorest children with ranked performance in educating their wealthiest. Ranked performance in educating the poor correlates with ranked performance in educating the wealthy, and the U.S. correlation ranks among the highest of all participating nations (Petrilli & Wright, 2016). Both studies (Hanushek & Woessmann, 2010; Petrilli & Wright, 2016) find that U.S. ranked performance on international tests, when compared against peer nations, is similar for wealthy students and poor students.

OECD (2012) compares the average PISA performance of students within home income deciles, compared across OECD nations. The analysis demonstrates the relative performance of nations’ low-income and wealthy populations, just as it demonstrates the comparative performance within nations of their poor and wealthy groups. While the wealthiest decile group of U.S. students ranks 31 of 39 in comparative, average PISA scale score, the least wealthy decile group of U.S. students ranks 18 of 39. Lowest-decile U.S. students outperform on average comparable decile students in 15 of the 30 countries whose top-decile students outperform the comparable U.S. group. U.S. comparative outcomes in educating its very poorest students exceed comparative outcomes in educating its very wealthiest. This analysis adds to evidence that comparatively low U.S. PISA performance is not associated with comparatively low performance among low-income U.S. students.

In an analysis of the impact of income inequality on reading skills, Siddiqi et al. (2012) examine 24 countries participating in the 2000 administration of the PISA. Independent variables in their analysis are the Gini coefficient to measure income distribution; per capita
GDP; education spending as percentage of GDP; school-level opinion of material and human resource constraints, as measured by surveys; and student-level factors, such as socioeconomic status (SES) and gender. Adolescent reading levels, while associated with GDP, have a stronger association with the extent of nationwide income disparities (Siddiqi et al., 2012). Greater average income disparities among earners had a downward effect on nationwide reading levels after the researchers controlled for social factors (Siddiqi et al., 2012). The researchers further find that income disparity and its effect on reading were particularly high in the U.S. (Siddiqi et al., 2012), which had the highest income inequality among national populations included in the sample. This correlation corresponds with other findings (OECD, 2015) of a relationship between U.S. income disparity and lagging cognitive performance among U.S. students. The finding is notable for understanding stagnant adolescent literacy levels: while there is evidence that the U.S. is not a comparatively low-performing nation in developing cognitive skills among its poorest students, the magnitude of income disparities within the U.S. as a whole is associated with the nation’s educational challenges, including its desultory reading levels.

**Relationships among Income Disparities, Social Networks, and Cognitive Skills**

Income disparities have an inverse association with social capital in a society (Kollmeyer, 2013; Robison, Siles, & Jin, 2011). Kollmeyer (2013) notes a strong relationship between family structure and income inequality, finding that the presence of single-parent households increases income inequality in a society. Kollmeyer (23013) further identifies the U.S. as an outlier nation in both its high presence of single-parent households and its magnitude of income disparity. Robison et al. (2011) similarly examine a large set of social capital indicators

---

9 A discussion of the effects of single-parent households on parental approach and cognitive skills will follow later in this chapter.
measured on the U.S. census, finding an association between the lack of social capital indicators and large income inequality. Similar to Kollmeyer (2013), Robison et al. (2011) find that the presence of single-parent households is particularly highly associated with income inequality in a given society. The relationship between social capital and income inequality is particularly profound because of other research indicating a steady decline of social networks in the U.S. (McPherson, Smith-Lovin, & Brashears, 2006; Putnam, 2000) and an increased presence of single-parent households in the U.S. (Amato, 2005). If income inequality is associated with lower nationwide cognitive skills, and if social capital is associated with income equality, then changes to social capital, including to family structure, could plausibly affect nationwide cognitive skills.

**Exosystem: Influence on Teaching and Parenting**

Factors in the exosystem do not directly impact the learner; rather they involve events or relationships that affect conditions in which the learner lives or learns (Bronfenbrenner, 1994). In particular, the exosystem involves the influence of particular social conditions on parenting and the influence of particular policies on teaching.

**Family Structure and Parenting**

As previously established, income inequality is associated with cognitive skills across nationwide populations (OECD, 2015; Siddiqi et al., 2012). Income inequality is also associated with weak social networks within a society (Kollmeyer, 2013; Robison et al., 2011). In the U.S., the comparatively large percentage of households headed by a single parent contributes to income inequality (Kollmeyer, 2013; Robison et al., 2011).

Adolescent literacy levels have seen no significant improvement in the U.S. in the last 30 years (National Center for Education Statistics, 2016; OECD, 2015). In that time, the structure
of the American family has undergone extraordinary change. The percentage of 35-year-old women in the U.S. who are married and living with children declined by 35% between 1970 and 2010 (Haskins, 2015). The percentage of children living in two-parent homes dropped from 85% in 1970 to 69% in 2010 as well (Haskins, 2015). Family structure is associated with academic and behavioral performance in school (Amato, 2005; Ginther & Pollack, 2004; Ribar, 2015).

Ginther and Pollack (2004) find that being raised by two parents biologically related to the child has a positive effect on grades and in-school behavior relative not only to performance in children raised by single parents but also to performance in children raised by one biological parent and one step-parent. Amato (2005) estimates that the nation would achieve an annual reduction of more than 700,000 students 12-18 years old who repeat grade levels were the U.S. rate of single-parent households to return to 1960 levels, impacting 2.6% of the total student population in the adolescent age group. Demuth and Brown (2004) note higher levels of adolescent delinquency not only in single-parent households generally but also when comparing the effects of single-parent households headed by men to those headed by women.

The relationship between family formation and adolescent development or cognitive skills having been established, this chapter will later examine the discrete social functions of parenting such as time allocation and parenting approach impacted by family structure (Ribar, 2015). There are additional exosystemic factors, including influences from outside the home, that relate to parenting approach. In one example, Presser and Ward (2011) note that 89% of American workers report having worked a nonstandard work schedule prior to turning 39 years of age, including night shifts and irregular daytime hours. More than 70% of Americans report having worked a night shift by the age of 39 (Presser & Ward, 2011). There is evidence that working nonstandard schedules contributes to parenting practices less sensitive to the needs of
CURRICULUM-SPECIFIC READING TEST

children, including less encouragement and less questioning when parenting young children (Grzywacz, Daniel, Tucker, Walls, & Leerkes, 2011). Hsueh and Yoshikawa (2007) find similarly find modest associations between nonstandard schedules and specific behaviors in the home, such as less sensitive parenting techniques, teacher-reported behavioral infractions, and irregular meals at home for children. For the purposes of this literature synthesis, these studies illustrate that family formation, while associated with parenting approach, is not an exclusive influence, and that society and the economy present myriad complex factors that influence parental approach from outside the home.

Influences on Teaching Approach

As salaries and benefits comprise 80% of public-school spending (National Center for Education Statistics, 2015), states and school systems are heavily invested in the capacities of teachers to achieve human capital returns. This section will thus examine a number of factors that relate to the manner or effectiveness with which adolescent students are taught to read. Rather than examining the actual approach teachers take in the classroom—teaching approach will be discussed in the microsystem section of this chapter—this section will outline prevailing influences on teachers’ approaches, starting with evidence that the effects of individual teachers vary widely to begin with.

Teacher Effectiveness and Characteristics

There is wide variation in the returns of student skill associated with different teachers (Chetty et al., 2014; Goldhaber, 2016; Hanushek, 2011; Staiger & Rockoff, 2010). Using regression analyses, Hanushek (2011) ranks the effects of teachers on standardized test scores, controlling for social factors outside of the teacher’s control. Hanushek (2011) asserts that a teacher in the 84th percentile has the effect of a standard deviation of at least .2 above mean
teacher effects, which the researcher translates to an increase of $400,000 in lifelong earnings for impacted students. Chetty et al. (2014) create a model demonstrating that replacing a teacher in one year who is performing in the bottom 5 percentiles of test score returns with a teacher performing at mean levels yields lifelong income gains for impacted students of $250,000. Goldhaber (2016) finds not only income gains associated with assignment of one teacher performing at the 85th percentile, but also an increase of nearly a full percentage point in the odds of attending college at age 20.

Some variation in the effects of teachers is attributable to years of experience (Staiger & Rockoff, 2011). In one analysis, student gains in classes taught by a first-year teacher trailed those of teachers taught by experienced teachers by .06 to .08 standard deviations (Staiger & Rockoff, 2011). But the variation quickly diminishes as teachers gain experience; students taught by second-year teachers lag only by .01 to .04 standard deviations (Staiger & Rockoff, 2011). The cost of teacher attrition in terms of returned skills, according to these findings, is greatest when experienced teachers are immediately replaced by novices.

Research also indicates that having a same-race teacher may boost long-term education attainment (Gershenson, Lindsay, Hart, & Papageorge, 2017). Gershenson et al. (2017) find that assignment of African-American teachers to African-American, male students for at least one year in Grade 3 through Grade 5 is associated with a reduction in high school dropout rates of such students by 39%. This finding adds to research documenting a higher likelihood of teachers perceiving students as behaviorally problematic when the teacher and student do not share racial backgrounds (Dee, 2004). Research has also documented positive effects on reading skills of same-race teachers for Black students in the U.S. (Egalite, Kisida, & Winters, 2015).

**Influences on Teachers**
A teacher is not simply an individual operating in isolation of other factors, however. A teacher is part of a system, one that prescribes standards for learning, requires tests that measure student learning, certifies that teachers have requisite skills, and equips teachers with particular tools, such as curricula. This chapter now turns to the influence of these systemic factors on teacher effectiveness and teacher approach.

Teacher Preparation. States regulate the preparation requirements and qualifications of teachers in U.S. public schools. Darling-Hammond (2000) finds that the extent to which teachers are certified in the content area in which they teach is predictive of statewide student achievement on standardized tests, before and after accounting for student demographic factors. Darling-Hammond (2000) further finds that the effect of teacher qualifications holds true across states and exceeds the effects of class size reduction or spending increases. In some cases, the effect on student learning of increasing the percentage of educators qualified to teach the subject in question is greater than the effect of changes in poverty status among students (Darling-Hammond, 2000). The evidence and case Darling-Hammond presents is, broadly speaking, one of faith in the capacity of teaching colleges and the rules of state credentialing systems to prepare teachers adequately.

These findings are contradicted, however, by other evidence that teacher preparation programs have widely mixed effects (Goldhaber & Cowan, 2014; Henry et al., 2014). Neither colleges of education nor other programs that states permit to certify teachers demonstrate consistently positive or negative effects (Goldhaber & Cowan, 2014; Henry et al., 2014). Goldhaber and Cowan (2014) further find that greater variation exists in the effectiveness of teachers within given programs than across the general cohort of programs. Even when there are distinctly positive effects of individual programs, these effects are often mitigated by higher
levels of attrition among program graduates of those programs, muting their positive effects (Goldhaber & Cowan, 2014). Neither a given type of teacher preparation institution, nor the credentials these institutions grant aspiring teachers, can be consistently associated with effectiveness in the classroom. By focusing on the effects of individual teacher preparation programs rather than on certification as a general condition teachers either experience or not, Goldhaber and Cowan (2014) and Henry et al. (2014) illustrate that a given school or school district cannot count on any given college of education or other preparation provider to reliably deliver fully prepared teacher candidates.

In a comprehensive study of the curricula, perceptions, and effectiveness of teacher preparation programs within colleges of education, Levine (2006) asserts that U.S. educator preparation programs are generally incoherent in their sense of purpose and course of study, resulting from the lack of a commonly acknowledged empirical basis on which to evaluate their effectiveness: “We don’t know what, where, how, or when teacher education is most effective. This means the education our teachers receive today is determined more by ideology and personal predilection than the needs of our children” (p. 19). Steiner and Rozen (2004) analyze the syllabi of core coursework in 16 nationally representative college of education preparation programs, focusing on 165 courses across the institutions. In 45 courses termed “general foundations,” they find a “preponderance of works that embrace a constructivist and/or progressive standpoint” (Steiner, 2005), while only two syllabi mention the noted scholar of core knowledge domains, E.D. Hirsch. In 59 courses dedicated to the teaching of reading, the authors find not a single mention of several scholars prominently featured in the National

10 As noted in the executive summary, the centrality of Hirsch’s work to the role of building knowledge in teaching reading comprehension will be discussed throughout this dissertation.
Reading Panel (2000) review of significant reading intervention findings. College of education curricula, Steiner (2005) writes, are inevitably a function of judgments made by faculty prescribing a world view for prospective educators. Too often, Levine (2006) writes, those judgments are anchored in no particular view as to what skills and knowledge are essential for the aspiring teacher: “The fundamental weakness in the teacher education curriculum is the lack of agreement about what it should produce. The result is that the teacher education curriculum is governed by a philosophy of ‘let 100 flowers bloom’” (p. 35).

This general lack of purposeful program design extends beyond the core curriculum and into the applied dimension of teacher preparation programs. Levine (2006) finds that alumni of teacher preparation programs overwhelmingly rank the applied aspect of their preparation experiences, during which candidates practice in a classroom setting, as most important to their success as full-time teachers. Yet such experiences, even for full-time undergraduates, typically do not exceed one semester of study (Levine, 2006). In its conclusive report, discussed in depth later in this chapter, the National Reading Panel (2000) finds the most promising classroom methods for increasing reading comprehension necessitate teaching reading multiple comprehension strategies in combination with one another, an intricate act of lesson design and implementation. Yet Steiner and Rosen (2004) find that of 59 foundational reading course syllabi reviewed, requirements to present a lesson or pass an assessment demonstrating mastery of applied reading strategies were rare. The National Council on Teacher Quality (2020) finds that the percentage of preparation programs incorporating established reading methods such as teaching awareness of words’ sounds to young children into their curricula has increased in recent years, the percentage of evidence-based reading preparation programs remains relatively low, at only 51 percent of programs. Similarly, Grisham et al. (2014) find that there are wide
gaps between the knowledge and skill college administrators and professors prioritize in syllabi and what undergraduate candidates perceive they are being taught. These differences of perception between college faculty and teaching candidates primarily regard the skills necessary to implement curricular and instructional knowledge in the classroom; candidates express less confidence that they are being taught skills needed to apply the pedagogical knowledge they have ostensibly gained. The literature thus paints a compelling and troubling picture of the role preparation programs may play in the problem of stagnant literacy skills; preparation is at once essential to effective application of effective practice and, in its current form, largely unmoored from fidelity to those very practices.

Standards-based Reform. Public school districts in the U.S. are overseen by thousands of elected local school boards. Meyer and Rowan (1977, p. 217) refer to school districts as “loosely coupled systems” that, much as they may conform to norms in how they behave, are largely independent actors. The seminal report A Nation at Risk (1983), issued by a panel of noted American educators commissioned during by the administration of President Ronald Reagan, began to change this notion of local autonomy, bringing together a new coalition of political actors dissatisfied with nationwide educational outcomes and willing to override the nation’s tradition of local authority (Mehta, 2013). In response, states largely stiffened requirements for high school graduation, developed more ambitious hurdles for teacher certification, and began to develop learning standards to determine what subject matter and skills should be taught and learned at each grade level in reading and mathematics (Mehta, 2013).

____________________

11 This description of the research and policy making efforts surrounding national and state standards and assessments is a very brief overview of two decades of intensive studies, reports, and legislation. The papers cited each offer detailed versions of the history.

12 I use the term learning standards as distinct from achievement standards or performance standards, which represent quantifiably verified levels of performance demonstrated on tests.
Rejecting both the traditional liberal call for greater levels of funding and the traditional conservative call for eradicating the federal role in education, President George H.W. Bush and President Bill Clinton successively envisioned new federal requirements for state-developed systems of learning standards and performance measurement (McGuinn, 2006). Congress passed the Improving America’s Schools Act of 1994, requiring states to develop learning standards and a regime of occasional tests that measured students’ performance on those standards (Hauptli & Cohen-Vogel, 2013; Mehta, 2013). The No Child Left Behind Act of 2001 (NCLB) added the requirement that states use these tests to rate school performance and intervene according to the magnitude and duration of a given school’s failure to meet standards (Hauptli & Cohen-Vogel, 2013; Mehta, 2013).13 Summarizing the sweeping nature of these changes not long after their development, Guthrie and Springer (2004) write, “A centuries-long American tradition of state plenary authority and local operating discretion is now giving way to a pressing national uniformity of federally imposed accountability requirements” (p. 7). By requiring states to lead the development of standards, Congress had successfully imposed a national system of standards and measurements on America’s schools.

The movement toward uniformity, however, did not end with NCLB. Convinced that NCLB was not manifesting gains on a national scale, governors and state education leaders

Learning standards are descriptions of the actual knowledge and skills to be taught and learned at given grade levels. As will be discussed at length later in this chapter, ELA learning standards have tended to emphasize skills more than knowledge and have not required students and teachers to emphasize any specific content. For a description of the distinctions among types of standards see NAE (2009).

13 Congress also commissioned the creation of a National Reading Panel during this period, whose report (National Reading Panel, 2000) included a catalogue and evaluation of research conducted on literacy education and a set of recommendations for local and state approaches to teaching reading. The findings of this report are central to understanding adolescent literacy approaches in the U.S. and will be discussed later in this chapter.
developed Common Core learning standards (CCSS) in literacy and mathematics that, adopted across state lines, sought to eradicate disparities in what students were taught from one state to the next (Lavenia, Cohen-Vogel, & Lang, 2015). In 2009, Congress tied state adoption of these or similar standards to a new pool of competitive federal funding called Race to the Top (RTTT), made available to states and school systems through the American Reinvestment and Recovery Act (Cassidy, Ortlieb, & Grote-Garcia, 2016; Lavenia et al., 2015). RTTT also included funds available to consortia of states seeking to develop and implement cooperative, shared tests to measure proficiency against shared literacy and math achievement standards (Gerwertz, 2016). Forty-four states adopted CCSS in the RTTT period (www.corestandards.org, 2018). While states experienced significant political backlash against CCSS and related testing consortia from teacher unions opposed to accountability provisions and conservative activists opposed to national standardization, the vast majority of states maintain a foundation of CCSS in their policies, even if the states have modified the substance or naming of their standards (Gewertz, 2016). This political environment and the resulting shifts in standards policy will be discussed later in this section in the context of Louisiana’s CCSS implementation experience and strategy.

Influence of Specific Standards-based Reforms. A Nation at Risk catalyzed a political movement centered on greater ambition for what American students learn and a lack of trust that local institutions were capable of delivering on that ambition (Mehta, 2013). Learning standards and tests to measure student performance became that movement’s vehicles. While there is an intuitive logic to standardizing expectations in order to increase transparency, motivation, and productivity, the mechanics of how such a system influences teaching approach are complex. This chapter now turns to this topic before discussing the approach to standards-based reform taken in Louisiana.
**Learning Standards.** By leaving the substance of ELA and math learning standards largely to the states, neither the Improving America’s Schools Act nor NCLB articulated a national vision for curricular or pedagogical change (Hauptli & Cohen-Vogel, 2013). State learning standards created during that period were often developed through local consensus on best practice, via political negotiation in state boards of education, state education departments, and state legislatures (National Academy of Education, 2009). The resulting lists of skills, strategies, and knowledge typically possessed neither a summative vision for student skills upon exiting high school nor a clear progression of learning goals in the years prior (National Academy of Education, 2009).

State ELA learning standards had additional and unique flaws. According to a report from the American Federation of Teachers (AFT; 2008), only 31% of all Grade 6-8 ELA learning standards met the simple standard of being both easily understood and unique to a specific grade level. More than 40% of all state ELA learning standards were verbatim copies of a standard in a different grade level (AFT, 2008). Nine states created no ELA learning standards for Kindergarten, Grade 1, or Grade 2 (AFT, 2008). And, critically, ELA standards typically made no reference to specific texts for students to read (AFT, 2008) and no explicit connection to standards or texts addressed in other subjects, such as science and social studies (NAE, 2009). State ELA learning standards thus frequently offered teachers, whose job it was to foster students’ capacities to comprehend ever more complex texts, no sense of how one grade level should differ from the next and no guidance on what specific readings or topics would best foster reading comprehension in the student.

---

14 For an egregious instance of the vagueness that often typified ELA learning standards, note one state’s Grade 4 example: “Demonstrate the understanding that the purposes of literary works include personal satisfaction and lifelong literature appreciation” (AFT, 2008)
Initiated by the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO) in 2009, CCSS were learning standards in ELA and math that could be shared by any and all states and could rectify design flaws of the past (Coleman & Pimentel, 2012; Pearson, 2013). CCSS in ELA emphasized three “key shifts” (CCSSI, 2018) in their design and in their implications for teaching practices: a continua of increased text complexity across grade levels, connections across academic disciplines and types of text to develop a foundation of knowledge, and close reading to identify specific evidence supporting arguments (Coleman & Pimentel, 2012; Pearson, 2013). As one example of how states communicated these shifts in practice, LDOE (n.d.) created a “crosswalk” to demonstrate changes in specific standards. CCSS “shifts” (CCSSI, 2018) are evident in the crosswalk (LDOE, n.d.). What was the vague instruction to Grade 7 teachers and students, for example, to “demonstrate understanding of grade-appropriate texts using a variety of strategies” translated under CCSS an instruction to “cite several pieces of textual evidence to support analysis of what the text says explicitly as well as inferences drawn from the text,” a more specific, verifiable skill for teacher and student. The crosswalk (LDOE, n.d.) further illustrates that, while prior Louisiana Grade 7 ELA standards bore no mention of text complexity or of reading across subject areas, CCSS included instruction to “read and comprehend literary nonfiction in the grade 6-8 text complexity band proficiently.” As the LDOE crosswalk illustrates, the shifts were evident in the language of CCSS and distinguished the standards from prior state-specific versions.

Absent in the CCSS instruction to read certain genre and to insist on texts of appropriate complexity, however, is any stipulation to read certain texts or to read about specific topics (Chambers, 2013). CCSS instead includes only an appendix that lists texts demonstrative of appropriate grade-level text complexity and a range of genre, periods, and authors (Chambers,
2013). Hirsch (2016, p. 106) notes that in the absence of direction as to specific texts or essential knowledge in the actual curriculum, the CCSS authors’ guidance carries no weight for school district decision makers, whose attention is given to state standards-based tests. Teachers “do what they are told to do by their administrators, under the ever-present threat of reading tests that do not actually test the content that is being taught” (p. 36). This distinction between the heavy influence of CCSS over skills tested by states and limited influence over knowledge and texts taught by schools and teachers is an essential theme to this study and will be discussed further later in this chapter.

Seeking to understand whether CSSS have effectuated a change in teaching methods, Opfer et al. (2016) conduct the American Teacher Panel (ATP), a survey of 1,705 teachers in states that adopted CCSS. They find that teachers of ELA, in an overt contradiction to explicit language in CCSS, largely continue the practice of selecting texts for students to read based on the current reading level of students rather than on the appropriate level of text complexity given grade-level learning standards (Opfer et al., 2016). ELA teachers further report a predominant intention to teach reading as a series of discrete, isolated skills, such as locating a text’s main idea, rather than centering instruction on assisting students with determining a text’s meaning, as also prescribed by CCSS (Opfer et al., 2016). This is a critical finding, both for its implication that CCSS, on their own, are limited in their impact on teaching methods, and for its implication that ELA teachers do not predominantly approach reading as an act of “intentional thinking during which meaning is constructed through interactions between text and reader” (Durkin, 1993, p. 5). Instead, teachers appear to be focused on teaching reading as an assortment or amalgam of discrete skills, a phenomenon that will explored further as this chapter proceeds.
Standards-based assessments. Kane, Marinell, Thal, and Staiger (2016) find that ELA teachers assign significantly more writing in CCSS states than was assigned prior to the adoption of CCSS. But Kane et al. (2016) notably attribute this not to the learning standards’ focus on writing in response to texts but to the writing-intensive requirements of CCSS-aligned ELA state assessments. This finding builds on a broad body of literature indicating that standards-based assessments significantly influence the activities and content taught in U.S. middle schools (Au, 2007; Moon, Brighton, Jarvis, & Hall, 2007; Musoleno & White, 2010; Zoch, 2015). Au (2007) and Musoleno and White (2010) find in particular that standards-based tests contribute to middle schools limiting the range subject matter taught by focusing largely on ELA and math, to the detriment of social studies, science, and the arts. Au (2007) and Zoch (2015) similarly attribute to standardized tests the tendency to parse reading instruction into increments centered on isolated skills. Au finally notes a change in pedagogy that results from the influence of tests, toward teacher-centric direct-instruction and away from dialogic models in which peers interact in order to apply skills or discuss meaning.

These studies imply that teachers take direction from standards-based tests, translating the requirements of tests into skills they teach students. Skills become “curricularized” (Fisher & Frey, 2008, p. 16) because the test requires that students demonstrate said skill. Among the three shifts embodied by CCSS in ELA—toward complexity, use of evidence, and broad knowledge—one could expect a generic state test to reasonably measure students’ skills at comprehending texts of certain complexity, locating certain evidence. But the standards required no specific texts or knowledge, minimizing the possibility that states’ standards-based tests would require students to demonstrate specific knowledge, of texts or otherwise (Hirsch, 2016; Steiner, Magee & Jensen, 2018). This knowledge- and text-agnostic feature of the state tests that
CCSS spawned, in Louisiana and across the country, will be a prominent theme throughout this study.

**Curriculum.** If standards unto themselves are ineffectual in influencing teaching practice, and if standards-based tests may influence teaching approach in ways that run counter to the aspirations of reading learning standards, a question remains as to how education systems can ensure the ambitions of their learning standards translate into approaches adopted by their teachers. Steiner’s (2017) comments on curriculum evidence a potential response:

>[A]ttention to the academic content – *what* we teach – is only now beginning to emerge as a serious lever in education policy. For starters, with the exception of California and a few districts across the country, we have no clear idea what is being taught in America’s classrooms. This would strike other countries as ludicrous: As Common Core, Inc., discovered in 2009, while top-performing countries differ in many facets of education policy, they all share a commitment to high-quality curricular content. (p. 1)

Emerging research indicates that use of specific mathematics curricular materials boosts student learning and skills above achievement using other curricular materials. A math teacher ranked at the 20th percentile in quantitative effect on student learning, for example, demonstrates a positive effect of .13 when using specific, prescribed lessons (Jackson & Makarin, 2017). Students using particular math curricula aligned with CCSS saw a positive effect on standards-aligned assessments, versus a negative effect for students using non-aligned curricula (Kane, 2016). Whitehurst (2009) writes that, relative to evaluations of policy interventions such as starting charter schools, using specific teacher recruitment initiatives, and offering pre-kindergarten, “[c]urriculum effects are large compared to most policy levers.”
Yet there is evidence that, even in states that have adopted CCSS, ELA teachers are not making use of curriculum based on external guidance or research. Opfer et al. (2016) find that ELA teachers are significantly less likely than teachers of mathematics to change instructional materials, such as textbooks, as a result of changes in state standards. Both Kane at al. (2016) and Opfer et al. (2016) find that a large majority of ELA teachers continue to develop their own instructional materials, such as activities, handouts, and assessments. As previously noted, CCSS called on ELA teachers to make three essential shifts, toward appropriate text complexity, use of evidence, and broad knowledge in the student. These shifts ostensibly have implications for the curricular content to be taught in classrooms (Hirsch, 2016; Steiner et al., 2018). But evidence indicates a breakdown between that aspiration and changes in the instructional materials teachers used.

This breakdown may have been attributable to lack of information and leadership. While CCSS included guidance to publishers as to how instructional materials should be designed in order to embody the standards and their related shifts (Coleman & Pimentel, 2012), as of 2015, five years after CCSS were released and broadly adopted, only 11% of textbooks reviewed by the consumer advocacy organization EdReports adequately embodied CCSS (Collette, 2015). Chingos and Whitehurst (2012) note that states typically neither monitor nor report the curricular selections made by school systems. Thus, for as much as the federal government and states have asserted control over local school systems via standards-based reforms, the substance of what is published and taught largely elude federal and state influence. This lack of central government influence over curricular materials is not exclusively a function of the decentralized nature of the U.S. education system (NAE, 2009). Instead, NAE (2009) writes, “[N]ational control is not required for coherence. Coherence leads to effective outcomes if it is achieved at whatever level
of governance has authority over policy instruments” (p. 5). Louisiana, the professional context for this dissertation, provides one example of a state seeking coherence between learning standards and curriculum.

**Standards-based reform in Louisiana.** The state of Louisiana adopted CCSS in ELA and math in 2010 (Loewus, 2016). In 2016, the state completed a review of ELA and math standards that resulted in modestly adjusted, state-specific standards derived from CCSS (Loewus, 2016). Similarly, Louisiana was affiliated with the Partnership for Access to Readiness for College and Careers (PARCC), a consortium of states designing and implementing CCSS-aligned standardized tests in math and ELA (Gewertz, 2016). In 2015, the state reinstituted its state-specific assessment, but this assessment heavily borrows content from PARCC (Gewertz, 2016). Finally, in 2017, as part of the state’s plan to comply with the federal Every Student Succeeds Act (ESSA), Louisiana passed policy requiring performance at the second highest of five performance levels on its state test, called “mastery,” in order for students to be declared fully proficient (Sentell, 2017). This step brought Louisiana into line with definitions of student proficiency used by PARCC and by NAEP.

Much as Louisiana has sought comparability with other states in its academic policy instruments, its approach to curriculum, professional preparation, and ongoing support has been uncommonly assertive among states (Doan et al., 2022; Kaufman et al., 2018; Opfer et al., 2016; Pondiscio, 2017). In 2014 the state commenced a rolling review of ELA and math curricula,\(^\text{15}\) placing each into tiers based on alignment with state standards, inclusion of formative assessment series, and intervention instruments for struggling readers. More information can be found at https://www.louisianabelieves.com/academics/ONLINE-INSTRUCTIONAL-MATERIALS-REVIEWS.

\(^{15}\) In the time since, this review process has been extended to science and social studies curricula, formative assessment series, and intervention instruments for struggling readers. More information can be found at https://www.louisianabelieves.com/academics/ONLINE-INSTRUCTIONAL-MATERIALS-REVIEWS.
assessment, and other factors (Doan et al., 2022; Kaufman et al., 2018; Pondiscio, 2017). The state then provided financial incentives for school systems acquiring curricula evaluated at the highest level, including free online access to curricula, statewide bulk printing contracts to limit the expense of hard-copy texts, and free access to curriculum-specific professional development for teachers (Pondiscio, 2017). When the review process did not result in highly rated middle school ELA curricula, the state commenced a partnership with LearnZillion, a private publisher, to create a standards-aligned curriculum entitled ELA Guidebooks 2.0 that is now in wide use across the state (Will, 2017). ELA Guidebooks 2.0 is organized around knowledge domains, foundations of thematically related information that are embodied in individual units of the curriculum. Each knowledge domain is reflected in the respective unit’s books, shorter texts, and other media.16

Louisiana uses highly rated curricula as the basis for an extensive program of professional support to teachers. In 2013, the state created the Louisiana Teacher Leader initiative, which provides roughly 7,000 teachers, representative of every school system in the state, with quarterly professional development specific to highly rated curricula (Kaufman et al., 2018). To conduct trainings at this scale, the state develops relationships with vendors who have demonstrated proficiency in implementing intensive professional learning specific to curricula (Kaufman et al., 2018). These vendors in turn contract directly with school systems, which have been encouraged to work with highly rated vendors through the state’s requirements that federal dollars dedicated to school improvement be used to acquire top curriculum and to contract with effective vendors (Kaufman et al., 2018).

16 A discussion of the role of knowledge in reading comprehension, largely centered on the work of E.D. Hirsch, is included later in this chapter.
Kaufman et al. (2016) use results from the American Teacher Panel (Opfer et al., 2016) to analyze teacher perceptions and behaviors related to state standards, instructional materials, and teaching approach. While 70 percent of teachers nationally select text based on the reading level of students, reducing text complexity when readers struggle, only 47 percent of Louisiana teachers identified with this practice, choosing instead to select texts appropriate to students’ grade level, per CCSS guidance (Kaufman et al., 2016). Similarly, as mentioned previously, while 75 percent of ELA teachers in the nationwide sample organized lessons around teaching discrete, isolated literacy skills, only 49 percent of Louisiana teachers identified with the practice (Kaufman et al., 2016). Instead, 49 percent of Louisiana ELA teachers identified texts and organized lessons around helping students to discern the meaning of a text, while this was only true for 21 percent of teachers elsewhere (Kaufman et al., 2016). Teachers in Louisiana were also significantly more likely to focus students on drawing inferences from texts and on analyzing the structure of texts to determine how parts relate to the meaning of the whole text (Kaufman et al., 2016). Kaufman et al. write

>C ompared with other teachers nationally, Louisiana teachers use some CCSS-aligned curricula at a higher rate than other teachers, demonstrate a better understanding of their CCSS-aligned standards, and report undertaking more instructional activities that align with their standards. Our interviews with Louisiana state officials suggest that [the Louisiana Department of Education’s] work to create a coherent environment for instruction is one likely reason for the differences we observed between Louisiana teachers and their peers in other states. (p. 15)

The implication of this connection between a coherent system of instructional materials and supports and teacher understanding of standards is notable. In the absence of “a coherent
environment for instruction” (Kaufman et al., 2016, p. 15) that includes curriculum and professional development aligned to learning standards, standard assessments and accountability systems may influence ELA teacher approach in an outsized and counterproductive manner, focusing teachers on isolated skills reflected in test questions (Au, 2007; Musoleno & White, 2010). As discussed later in this chapter, such a skills-based approach to teaching adolescent reading contradicts the stated intent of CCSS to emphasize building bodies of knowledge rather than skills alone (Coleman & Pimentel, 2012; Opfer et al., 2016).

Microsystem: Parenting Approach to Reading and Teaching Approach to Reading

Microsystemic factors associated with the problem of practice extend the application of human capital and social capital constructs to proximal relationships (Bronfenbrenner, 1994) with the learner. Macro trends in social capital and income distribution are associated with changes in family structure that influence how parents approach reading in the home. Similarly, national and state education policies seeking increased skills among students and graduates influence the incentives and approaches classroom teachers. This section of the chapter addresses the problem of adolescent reading skills at the proximal level, between parent and child or between teacher and child.

Parenting Approach

The relationship between parental education level and their children’s cognitive skills, including reading skills, is well documented and has been discussed in this chapter (Coleman, 1966; Davis-Kean, 2005; Durik et al., 2006). Desjardins (2003) finds that parental education

Doan et al. (2022) find that while Louisiana ELA teachers are more likely than peers in other states to use standards-aligned curricular materials, they express no greater confidence than other teachers that the curricula are sufficient to prepare students for state ELA tests. This finding will be particularly relevant in chapter 5 discussion of results.
level is associated with a child’s literacy skills, lasting through adolescence and into adulthood. The effect of parental education is mediated by the parent’s encouragement of further education in the child—highly educated parents tend to encourage more education—which is the specific behavior to which Desjardins (2003) attributes effects. Durik et al. (2006) find that parental education level is associated with reading efficacy in elementary school. As students age, their self-conception as a reader, formed in the early years, has a propelling or depressing effect on reading skills (Durik et al., 2006). Students who believe they are good readers, and who have high grades at younger ages, tend to gain reading skills at increasingly high rates as they progress through adolescence (Durik et al., 2006). The opposite is true for struggling readers, who fall further behind as they age (Durik et al., 2006). Deslandes, Royer, Turcotte, and Bertrand (1997) similarly find that parental approach and style are associated with adolescent reading skills, concluding that an approach combining affective support directly to adolescents, parental communication with teachers, and a blend of control and autonomy in guiding adolescents, is related to positive change in reading skills. While some of these studies attribute the effects of parents on adolescent reading to the parents’ level of education, all conclude that an encouraging, motivating parental approach to reading has a positive association with adolescent reading skills.

In a study of representative samples from 27 countries, Evans, Kelley, Sikora, and Treiman (2010) find that children raised with many books in their homes can expect to achieve 3.2 years of education over their lifetimes more than can children raised with no books. Evans et al. (2010) find that a child raised with 500 books in the home, for example, is 36% more likely to graduate from high school and 19% more likely to graduate from college than is a child raised without books. These effects are strongest among children and families with the lowest levels of income and parental education, again demonstrating that the effect of parent choices and home
environment on reading, while associated with parental education levels, are related also to parental choices and home environments themselves. Myrberg and Rosen (2009) find that parental education level contributes to literacy in adolescents precisely because it is associated with the number of books to which parents expose their children, and because highly educated parents have the knowledge necessary to make use of books for educational purposes in the home (Myrberg & Rosen, 2009). It is highly educated parents’ capacities to use books for educational purposes that differentiates their effects on readers, the researchers find, rather than the number of books alone (Myrberg & Rosen, 2009).

The implication of this research on parental approach to reading is profound: it is both a specific method and an underlying understanding of why that method matters that have lasting effects on the reading skills of children, through to adolescence. Though it is not dependent on parental education level—effects of books in the home have been shown to be positive where parents lack education—this effective parental approach to reading is associated with parental education levels and thus to social trends that distribute educational and economic effects unequally, especially in the highly stratified U.S. (OECD, 2015; Siddiqi et al., 2012). Family structure is associated both with large-scale social trends and with parenting approach; two parents may have more time for a nuanced parenting approach, greater resources with which to adopt the approach, or simply greater odds of including a parent in the approach (Martin et. al., 2007). There thus may be a problematic cycle associated with the problem of stagnant adolescent literacy levels: parents with lower educational levels have fewer financial resources and are less likely to reside in two-parent homes, making it less likely they will be able to provide voluminous libraries or to engage in approaches that foster strong reading skills in their children.
Teaching Approach

Previous sections of this chapter identified adolescent reading as a distinct activity from reading in the primary grades. Adolescent readers are required to read across disciplines that require knowledge of specific vocabulary and content (Kamil, Bowman, Dole, Kral, Salinger, & Torgesen, 2008; Murnane et al., 2012; Snow & Biancarosa, 2003). They are required to comprehend the meaning of texts rather than simply to decode the words on a page (Chall, 1983; Kamil et al., 2008; Murnane, et al. 2012; Snow & Biancarosa, 2003). Comprehension is an act of construction, in which readers actively form meaning through integrating knowledge of words in a text, the structure of a text, and related content (Shanahan, 2005). An individual’s capacity to make meaning grows through a zone of proximal development (Vygotsky, 1978) as these sources of knowledge—words, structure, facts—grow more complex. This discussion of teaching adolescent reading will draw on literature to define both methods and principles that compel teachers to use certain methods—that foster students’ making meaning of texts. This section will thus complete this chapter’s line of logic on education policy: from policy makers’ aspirations to develop human capital, through to the intentions and impacts of standards-based reform, down to the classrooms in which adolescents are taught to read.

National Reading Panel (2000)

Commissioned by an act of Congress, the NRP was a cohort of researchers and literacy experts charged with providing a definitive inventory of effective practice in teaching reading. The panel’s ultimate report (NRP, 2000) entailed separate sections on alphabetics, fluency, comprehension, teacher education, and emerging technologies. Because this chapter focuses on reading skills of adolescents, centered around understanding and making meaning of texts, it will focus on NRP findings related to teaching text comprehension. This brief synthesis of that
section of the report will provide the basis for later exploring two particular elements of teaching approach: strategies and knowledge.

From 203 studies the NRP reviewed, the NRP identified 16 teaching practices associated with growth in student reading comprehension (NRP, 2000). Of these 16 practices, 8 were backed by repeated and clear scientific evidence of effectiveness. These practices included cooperative learning among students, use of graphic organizers to capture a text’s meaning, and regular summarization of a text (NRP, 2000). Grouped together, these individual methods were referred to as “reading strategies,” discrete techniques that could be implemented in order to assist a student in comprehending a text.

However, the panel found that teaching students to comprehend by teaching students to use reading strategies—as with use of graphic organizers or summarizing a passage—is most effective when these strategies are used in combination with one another and are used in the context of teaching to make meaning out of a whole text:

Multiple-strategy instruction that is flexible as to which strategies are used and when they are taught over the course of a reading session provides a natural basis for teachers and readers to interact over texts. The research literature developed from the study of isolated strategies to their use in combination to the preparation of teachers to teach them in interaction over texts with readers in naturalistic settings. The Panel regards this development as the most important finding of the Panel’s review because it moves from the laboratory to the classroom and prepares teachers to teach strategies in ways that are effective and natural. (NRP, 2000, p. 4-6)

This is a conclusion of critical importance to the problem of stagnant adolescent reading performance. While teaching students to use isolated comprehension strategies has a basis of
supportive evidence, it is the use of strategies in combination with one another while reading a text to discern its meaning that has the greatest basis in evidence (Kamil et al., 2008; NRP, 2000). The optimal environment for comprehension instruction, is a “naturalistic setting” (NRP, 2000, p. 4-6) in which the reader and teacher alike are engaged in trying to figure out what the text actually means, using reading comprehension strategies as means to that end. Two practical implications emerge of this conclusion. First, reading strategies are not themselves isolated skills that should be taught outside efforts to comprehend a whole text. Second, even teachers skilled at using strategies must know what it is necessary to help students discern the meaning of a text. These two concepts—isolated skills and strategies versus making meaning of a whole text—will be explored for the remainder of this chapter.

**Reading Comprehension Strategies and Reading Skills**

Fisher and Frey (2008) advocate particular caution with respect to teaching reading strategies in isolation of one another or in the absence of attempts to discern a whole text’s meaning. “We can’t find any evidence,” they write, “for the effectiveness of teaching one strategy at a time, especially with pieces of text that require readers to use a variety of strategies to successfully negotiate meaning” (p. 16). Too often, write Fisher and Frey (2008), strategies themselves are “curricularized” (p. 16), becoming the organizing principle of a lesson rather than serving as tools for helping students discern textual meaning. This warning echoes concerns discussed previously that the structure of standard tests may influence teachers to teach individual, discrete skills, such as summary or pinpointing a text’s main idea, rather than to organize lessons around discerning the meaning of texts (Au, 2007; Kaufman et al., 2016; Hirsch, 2016). “We also know that some teaching strategies are ineffective,” write Steiner et al (2018). “One common example is drilling students on isolated comprehension skills such as
“‘finding the main idea’ instead of taking an integrated approach to teaching reading comprehension in the context of a specific content area.” The literature is clear on this point: strategies that involve isolating and teaching comprehension skills is a false and unproductive pursuit, prevalent though it is. What NRP (2000) called a “naturalistic setting” and what Steiner et al. (2018) refer to as an “integrated approach” comprise an alternate vision that is anchored in a student’s effort to construct meaning from a whole text. CCSS ostensibly aspired to foster an integrated approach by guiding teachers and curriculum publishers toward three shifts: use of evidence in analyzing texts, appropriate levels of text complexity, and the fostering of content knowledge across disciplines (Coleman & Pimentel, 2012). Building on these themes, this chapter turns to an examination of elements that might make for an integrated approach to reading comprehension.

Close Reading

A premise of CCSS in ELA is that comprehending a text necessitates the ability to derive meaning from the text itself, above all other sources (Coleman & Pimentel, 2012; Pearson, 2013; Shanahan, 2013; Snow & O’Connor, 2016). CCSS thus include specific instruction to teachers to focus reading lessons “on what lies within the four corners of the text” (Coleman & Pimentel, 2012, p. 4). This call to engage in careful examination of the words and phrases for indicators they offer as to a text’s larger meaning, called “close reading” (Coleman & Pimentel, 2012, p. 4), ostensibly offers an antidote to oft-practiced strategies, such as priming the reader’s knowledge of relevant background facts prior to reading, that ease the burdens of struggling readers but take readers’ focus and time away from actual texts (Pearson, 2013; Shanahan, 2013; Snow & O’Connor, 2016). By insisting on the closest examination of the text itself, CCSS recall Durkin’s (1993) essential definition of comprehension: “intentional thinking during which
meaning is constructed through interactions between text and reader” (p. 5).

The relationship between reader and text is a reciprocal one, however. Shanahan (2005) defines comprehension as “a form of active and dynamic thinking [that] includes interpreting information through the filter of one’s own knowledge and beliefs, using the author’s organizational plan to think about information (or imposing one’s own structure on the ideas), [and] inferring what the author does not tell explicitly as well as many other cognitive actions” (p. 1). Comprehension, the capacity to formulate a text’s meaning, is thus a function not only of the microscopic view of what the text contains, but also of the macroscopic view of its genre and structure, and of the ever-evolving knowledge base the reader possesses prior to and during the reading experience. Comprehension, Shanahan’s definition tells us, is a function not only of the combined connotations and denotations of words, but also of the reader’s capacity to synthesize connotation and denotation with the structural and worldly context within which the text exists.

This balance between comprehension as a function of inquiry “within the four corners of the text” and comprehension as a function of knowledge outside the text is in line with the National Reading Panel’s (2000) call to put reading strategies in the service of a whole text’s meaning. As discussed, efforts to build background knowledge prior to reading can distract from actually reading texts (Pearson, 2013; Shanahan, 2013; Snow & O’Connor, 2016). But the emphasis on close reading has its risks to developing comprehension ability as well. Notably, by insisting that meaning be derived from the text alone restricts the students’ capacity to read in order to gain knowledge, as gaining knowledge necessitates the interplay of new and old knowledge (Pearson, 2013; Snow & O’Connor, 2016). Similarly, the CCSS focus on textual evidence as the exclusive basis for argumentation may risk a “lethal mutation” (Snow & O’Connor, 2016, p. 6) of this habit in students, “curricularizing” (Fisher & Frey, 2008) the
discrete skill of locating evidence to the detriment of the integrated intellectual operations a comprehending reader must complete. The CCSS focus on close reading may contribute profoundly to reading teachers’ incentives to focus instruction on texts, just as it may inhibit, perhaps unintentionally, their incentives to foster broader, contextual knowledge and inquiry necessary to glean knowledge and make full meaning of what one reads.

Knowledge

It is important to distinguish between background knowledge, which is the base of understandings a reader possesses prior to engaging a text, and the knowledge gleaned from reading the text itself. The two form a virtuous cycle as it relates to a reader’s capacity to make meaning of texts; possessing relevant background knowledge enables a reader to comprehend new text, just as the newly gleaned knowledge becomes part of the background the reader brings to reading future texts (Pearson, 2013).

The mind makes meaning of received information, such as language on a page, by briefly holding the information in its short-term memory (Willingham, 2006). By drawing on related information stored for greater periods of time in its long-term memory, the mind associates new knowledge with previously held knowledge. This previously held knowledge includes schema, or ways of organizing knowledge (Willingham, 2006). The mind consequently can make quick decisions regarding newly received information based on the new information’s relationship with long-stored information. New information that is easily related to prior knowledge is more easily connected to prior knowledge and more easily retained (Willingham, 2006). Studies tying reading comprehension capacity to previously held knowledge can thus be explained by a basic

---

18 There is not yet definitive evidence of the preponderance of teaching students to locate evidence as a discrete skill in the years after CCSS adoption, but frequent visits to classrooms will indicate to most observers its prominence as a prioritized skill in ELA classrooms.
psychological concept: learning new things is made easier by previously knowing other related things.

A strong base of evidence backs the notion that possessing relevant background knowledge assists readers in comprehending. In a seminal study, Recht and Leslie (1988) measure students’ capacities to retell and summarize a written description of an inning of baseball action, finding significantly greater capacity to comprehend that description based on prior knowledge of baseball. Salmeron, Kintsch, and Canas (2006) find superior comprehension of online text in readers who progress through series of web pages displaying related content, as opposed to readers who repeatedly select disconnected web pages to read and thus read thematically unrelated content. Cromley and Azevedo (2007) similarly find that middle school students explicitly taught background knowledge related to texts comprehend those texts with much greater proficiency. Reznitskaya, Anderson, and Kuo (2007) find that reading comprehension improves when students read the text after having been taught an argumentation schema, a way of organizing information they read into an argument. These studies indicate that the presence of previously held knowledge related to a given text—topical facts, related online texts, or structural frameworks for making sense of the new knowledge—increases the likelihood of comprehending the text.

Hirsch notes that for a member of society to comprehend what she is generally expected to read, she must possess a body of knowledge on commonly read subjects (Hirsch, 2003). Hirsch writes that “the need for making inferences by activating already existing domain knowledge” (2003, p. 18) makes it imperative that teachers build knowledge in students based on specific and coherent domains. Prior knowledge of relevant subject matter allows the development of further knowledge, increasing likelihood and ease of comprehension.
For Hirsch (2003), a curriculum that fosters bodies of knowledge addresses deleterious educational effects of social inequities:

A poor [curriculum] adversely affects low-income students more than middle-income students who are less dependent on the school in gaining knowledge. By contrast, a good program is inherently compensatory because it has a bigger effect on low-income than middle-income students. This is because low-income students have more to learn—and in an effective program they begin to catch up. (p. 28)

James Paul Gee (2008) similarly notes that students are equipped with varying cognitive schema and varying levels of background knowledge. Explicitly equipping students with new knowledge and schema for organizing knowledge is an essential condition of a more equitable education system (Gee, 2008). However, neither CCSS nor the standardized tests states typically use to measure reading proficiency specify a body of knowledge or knowledge of specific texts (Hirsch, 2016; Willingham, 2017). While this practice may originate in federal and state desires to stay out of what have traditionally been local decisions regarding curricular content (Mehta, 2013), the effect is partially to test students’ knowledge on a random basis, with students who have had greater exposure to a wider world more likely to possess background knowledge on tested passages (Willingham, 2017). Related to the problem of practice, students experience these assessments on an annual basis, year after year. Disadvantaged students may experience a compounding effect of an education in reading that equips them with an incoherent base of knowledge even as the need for knowledge as a prerequisite to reading comprehension increases with age.

**Teaching Textual Meaning**

The emerging picture of teaching students to make meaning from a text is one of balance
between deriving meaning from the text’s words and applying tools exogenous to the text that readers must possess to comprehend what they read. Close reading is an important call to focus reading instruction on the text and the knowledge it signals, but an exclusive focus on the words within a text as the only source of inquiry may limit the external context necessary for the reader to discern a full sense of meaning (Hirsch, 2016; Pearson, 2013; Shanahan, 2013; Snow & O’Connor, 2016). Reading strategies may be necessary to achieving a close reading of the text, but they are not, unto themselves or on their own, equivalent to the ability to comprehend a text (Cervetti, Barber, Dorph, Pearson, & Goldschmidt, 2012; NRP, 2000; Fisher & Frey, 2008).

Background knowledge is essential to comprehension, but simply teaching background knowledge in lieu of having students read may distract from the important effort readers put forth in order to gain knowledge from a text (Hirsch, 2003; Snow & O’Connor, 2016).

The notion of a zone of a proximal development (Vygotsky, 1978) can be seen as a call to growing complexity in the reader across the range of knowledge that comprehending a text requires: a text’s words, its structure, and related background knowledge. As conventionally understood, complexity in CCSS has been taken to be a function exclusively of the vocabulary and understanding of textual structure required for a reader to comprehend (Hirsch, 2016). But in Hirsch’s vision of a whole-group, coherent curriculum, the ability to comprehend complex texts is chiefly a function of increased complexity in the knowledge a student possesses: “As topics become familiar, student ability to master rarer words and longer sentences (the chief elements of complexity) increases” (p. 115).

Given this understanding of complexity, the learning process ideally develops unrealized capacity in the reader to make meaning of texts, via escalating complexity in text’s words, textual structures, and base of knowledge. Facilitating this process of engaging to address
increased complexity across these dimensions of a text is the role of the teacher. The integrated reading classroom is one in which teachers equip readers with strategies and knowledge that readers use “with automacity, applying them as they read” (Fisher & Frey, 2008, p.16) to make meaning of increasingly complex materials. The goal of reading comprehension instruction is thus the independent and integrated application of knowledge—from within the text, from its structure, and from beyond the text—in order to construct textual meaning. This integrated approach will serve as a framework for synthesizing intervention literature, and ultimately identifying salient interventions, in chapter 3.

Conclusion

The problem of stagnant adolescent reading in the U.S. in a multi-dimensional problem manifested at varying levels of an ecological system. Nations seek returns of human capital when they invest in education and make education policy. For the last 30 years, standards-based reforms have centered improvement efforts in the U.S. around learning standards for student skills and minimum performance expectations on standards-based assessments. It is plausible that this system of measurement and accountability has not positively contributed to adolescent reading comprehension skills in the same way it has for younger readers. While adolescent readers require increasing bodies of knowledge and the ability to use reading strategies independently and in combination with one another, U.S. students may not experience curricula focused on building knowledge or teaching approaches that integrate strategies with attempts to make meaning of whole texts. This may be due in particular to inchoate curricula in colleges of education or in the inability of teacher preparation programs to assist teachers in transferring learned practices to the classroom setting. It may also result of the inability of states and school systems to provide teachers with coherent and aligned curricula and professional support that
embody the aspirations of CCSS, which sought to influence teachers and curriculum publishers toward an integrated approach to reading.

At the same time, income inequality is particularly pronounced in the U.S. and is associated with stagnation in nationwide cognitive skills, including reading skills. Income inequality is further associated with weak social cohesion. In the U.S., the prevalence of single-parent households is associated with both income inequality and weakening social networks. An effective parenting approach builds knowledge of books, engages in dialogue about books, and develops a sense of self-efficacy in reading. However, parenting approach to reading may be hindered by disruptions to family structure.

The problem of stagnant adolescent reading performance in the U.S. is unlikely to be resolved by seeing either of these two diagnoses, one focused on schools and the other on homes, as an exclusive contributor to the problem. Neither the home nor the school has complete influence over the reading skills of adolescents. Instead, policy makers can identify those ways in which opportunities for adolescent readers can be improved, recognizing the inevitable and inextricable relationship and tension between the social and the educational. This dissertation now turns to an initial assessment of the problem of stagnant adolescent reading levels in the specific context of Louisiana public schools.
Chapter 2

Needs Assessment

As discussed in the previous chapter, the problem of stagnant U.S. adolescent reading levels is a multi-dimensional problem associated with in-school and out-of-school factors. This chapter presents evidence from an initial needs assessment related to the problem of reading skills among high school and middle school students in the context of Louisiana public schools. Based on factors detailed in chapter 1, the chapter establishes research questions seeking further understanding of the problem as it manifests in Louisiana schools, as well as a plan for gathering data necessary to answer these questions. The chapter concludes with an analysis of the data and a discussion of the results, adding further substance and context to the problem of practice and factors identified in chapter 1.

Louisiana is a notable setting for understanding in-school factors related to adolescent reading levels. As noted in chapter 1, the state of Louisiana has played an uncommonly assertive role in matters of curriculum, conducting online reviews of published ELA curricula, and even partnering with a vendor to design a middle school ELA curriculum for use statewide (Pondiscio, 2017). Relatedly, the shift toward teaching approaches espoused by CCSS, including using texts appropriate to grade level and planning instruction to make meaning of texts rather than to foster discrete reading skills, is more commonly understood by teachers in Louisiana than by teachers in other states that adopted CCSS (Opfer et al., 2016).

In the time since Opfer et al. (2016) evaluated changes to teaching approach in response to CCSS, Louisiana has further adopted a plan to comply with the federal Every Student Succeeds Act (Sentell, 2016), new standards for educator preparation programs (Sentell, 2017), and a new Grade 7 reading test under the Innovative Assessment Design Authority (IADA) of
the Every Student Succeeds Act (ESSA; Sentell, 2018). These policies imply not only that Louisiana has taken its implementation of new standards further than had been the case in 2016, but also that the state is exploring a variety of strategies for improving student learning, creating options for studying interventions that will be discussed later in this dissertation.

**Research Questions and Methods**

This needs assessment used an online survey emailed to teachers from the school system in which they are employed. This section outlines the research questions underlying the survey, the sampling method, procedures for obtaining consent, and details of the instrument itself.

**Questions**

The needs assessment answers three research questions necessary to describe curricular choices, teaching approaches, and influencers of Louisiana middle school ELA teachers.

- **RQ1:** What are the curricula that Louisiana middle school ELA teachers use in their classrooms?
- **RQ2:** What activities do Louisiana middle school ELA teachers in middle schools use to help students improve as readers?
- **RQ3:** What factors influence Louisiana middle school ELA teachers’ decisions about the activities and curricula they use in the classroom?

The first question regards the sets of books, handouts, and other published materials ELA teachers use in their classrooms. There is evidence that ELA teachers across multiple states have not made significant changes in the curricular materials they use, even as standards changed in most states (Opfer et al., 2016). There is also evidence that ELA teachers tend to create their own curricular materials more frequently than do teachers of mathematics (Kane et al., 2016; Opfer et al., 2016). Louisiana has been noted as a state whose leadership has promoted use of
particular curricula (Opfer et al., 2016; Pondiscio, 2017). Thus, it is important to understand the names and types of textbooks and other materials teachers are using as part of describing teaching approach and its influences in Louisiana ELA classrooms.

The second question regards teaching approach, the methods teachers employ and the principles that guide their choices. There is evidence that ELA teachers across the U.S. teach reading through discrete activities such as summarizing or locating the main idea within a passage (Au, 2007; Opfer et al., 2016; Zoch, 2015). However, approaches that help students make meaning of full texts, rather than focusing on discrete skills, have positive associations with reading comprehension among adolescents. These approaches include dialogic teaching (Applebee et al., 2003), building background knowledge (Cromley & Azevedo, 2007; Hirsch, 2003; Salmeron, Kintsch, & Canas, 2006; Willingham, 2006), and teaching specific schema for organizing knowledge students receive through reading (Reznitskaya et al., 2007).

The third question regards policies, sources of guidance, and other sources of information that inform ELA teachers’ approaches to teaching reading. Au (2007) and Musoleno and White (2010) find that annual state reading tests contribute to teachers teaching reading as a series of discrete operations. Kane et al. (2016) find that the requirements of new literacy tests associated with CCSS contribute to teachers requiring more writing in the classroom. Opfer et al. (2016) find that changes in state standards do not lead to changes in approach among ELA teachers. However, Opfer et al. (2016) find that awareness of CCSS-prescribed approaches, such as increasing student exposure to grade-level texts, varies widely across states.

Sample

The survey focuses on one school system, Ouachita Parish public schools, a participant in the Louisiana Department of Education’s implementation of the IADA pilot. In Ouachita Parish,
58.4% of students qualify for federally subsidized lunch service, compared with 66.9% statewide. The parish’s student body is 1.4% limited English proficient, compared with 3.5% statewide. With 18,920 students, Ouachita is the state’s 13th largest school system out of 69 school districts. There are 200 teachers actively teaching Grade 4-8 ELA courses in Ouachita Parish public schools (R. Lamury, personal communication, December 10, 2018). These teachers teach in 27 elementary and middle schools. The survey uses a voluntary sample of Grade 4-8 ELA teachers in Ouachita Parish. Respondents not actively teaching ELA or reading courses in Ouachita Parish schools were removed from the sample.

**Recruitment and Consent**

I first held a conference call November 2018 with Superintendent Don Coker of schools in Ouachita Parish to outline the objectives and procedures for the survey. So as to avoid coerced participation, I confirmed that Ouachita Parish had no relationship to any school within the state-run Recovery School District, the state-run Special School District, the state-run Office of Juvenile Justice, or type 2 charter schools, which are authorized by the state Board of Elementary and Secondary Education (BESE). In a subsequent letter, I also asked Superintendent Coker to distribute a link to the online study form to members of the sample. The Ouachita Parish school district distributed the online survey November 2018. The survey remained open for four weeks. Two weeks later I provided a follow-up email to Superintendent Coker, providing response rates to date.

**Instrumentation**

The online survey measured three constructs: use of curriculum, teaching approach, and influence on teachers. All items on the survey were modified items developed by the RAND Corporation for inclusion in the American Teacher Panel (ATP; Opfer et al., 2016). Opfer et al.
derived ATP items from items developed by Shanahan and Duffett (2013) to measure approach to teaching reading. Shanahan and Duffett piloted all items. Opfer et al. (2016) also piloted all items in 2015 and adjusted wording when the meaning of a question was confusing or unclear (J. Kaufman, personal communication, March 7, 2018). There has not been a study of the reliability or validity of the ATP survey (J. Kaufman, personal communication, March 7, 2018). However, Shanahan and Duffett (2013) conducted their study via both online and paper-based modes, finding consistent results on every item and no effect of the mode or other externalities.

**Survey Construct: Use of Curriculum**

This construct involved printed or online materials, including textbooks, handouts, tasks, novels, and other texts that ELA teachers use in their classrooms for instructional purposes. The construct also involved the entity developing these materials—a publisher or the teacher herself, for example—as well as the frequency with which the materials are used. Item 1 on the needs assessment survey form (see Appendix A) used a nominal scale to measure a variable, curriculum sources. Curriculum sources included the actual developers or selectors of curricular materials, including school districts or teachers themselves. Item 1 also used a ratio scale to measure a second variable, frequency of using curricular source. This variable was represented by the number of days per week in which respondents use each curriculum source. Item 2 was a multiple-choice item that uses a ratio scale to identify the variable of time spent creating or identifying curriculum. This variable involved the number of hours teachers individually spend creating or identifying curricular materials in a typical week.

**Survey Construct: Teaching Approach**

This construct involved the activities teachers use to develop adolescent students’ reading skills. Item 3 in Appendix A was a multiple-choice item that uses a dichotomous scale to
measure the variable of classroom activity orientation. This variable was represented by statements indicating whether the teacher plans activities starting with specific texts in mind or starting with discrete activities in mind. Item 4 used a nominal scale to measure the variable of classroom reading activities. These were specific tasks in which students engage during ELA class. Item 4 also used a five-point ratio scale to measure the variable of activity focus, which the item represented in terms of the percentage of time in a typical ELA lesson consumed by each activity listed on the nominal scale.

**Survey Construct: Influence on Teachers**

This construct involved the policies, guidance, experiences, and sources of information that inform the curriculum and the activities ELA teachers employed. Item 5 uses a four-point Likert scale to measure the variable of influence magnitude. This variable was represented by statements indicating the extent to which a given source of influence does or does not inform curricular materials Louisiana ELA teachers use. Item 6 used the same Likert scale to measure level of influence over reading activities.

**Data Collection**

To administer the survey and to collect data, I used the online service surveymonkey.com. Using the platform’s “analyze results” tool, I exported the results after the survey window closed into the application Statistical Package for the Social Sciences (SPSS). No names or other personally identifiable information were included in the SPSS file. Instead, I assigned every respondent an identification number based on the respondent’s place in the order of response.
Results

Of 200 eligible teachers, 115 (57.5%) responded to the survey. I removed 10 respondents from the sample because they responded “no” when asked whether they taught ELA or reading in the 2017-2018 school year. I then conducted an analysis of distribution frequencies for each item on the survey.

The share of the resulting sample attributable to any given school ranged between 1% and 7.6%. Teachers with one to three years of experience made up 6.7% of the respondents; the survey did not include first-year teachers. Teachers with 4-14 years of experience made up 47.6% of the sample. Teachers with 15 or more years of experience comprised 45.7%.

Teachers reported regular use of two particular curricula. LearnZillion’s ELA Guidebooks 2.0 is used at least twice per week by 61.9% of teachers in the sample. Only 18.1% never used ELA Guidebooks 2.0, a response potentially attributable to a misalignment between the grade levels sampled and the focus grade levels for that product. Renaissance’s Accelerated Reader was the second most frequently used curriculum product, with 57.2% reporting use at least twice per week; 29.5% never used Accelerated Reader.

However, teachers also reported great frequency in using materials they selected or developed themselves. Of teachers surveyed, 80.9% reported using self-generated materials at least weekly. Similarly, teachers reported spending considerable amounts of time generating their own materials. Nearly 75% reported spending at least three hours per week generating materials.

Table 2.1

*Weekly Hours Teachers Spent Selecting or Developing Curricular Materials*
Given a choice between stating that they organize lessons around particular reading skills or around knowledge of particular texts, 64.8% responded that they organize lessons around particular reading skills. Asked about the activities included in lesson plans, 93.3% reported including skills such as summarizing and locating the main idea of a text either often or always. This contrasted with lower percentages of teachers who reported teaching knowledge relevant to the text often or always (76.2%) and developing persuasive arguments about texts (55.2%).

Teachers reported strong alignment with activities meant to reflect tasks called for on state reading tests; 98.1% of respondents indicated that they often or always ask students to respond to questions similar to those on Louisiana’s state assessment when engaged in reading a text. Teachers reported being very significantly or significantly influenced by the content of state standards (98.1%) in choosing classroom activities and texts. They reported similarly high levels of influence (94.3%) from the state assessment.

Table 2.2
Curriculum-Specific Reading Test

Extent to Which Classroom Activities Were Influenced by State Tests

<table>
<thead>
<tr>
<th>Extent</th>
<th>Respondents</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at All</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>Somewhat</td>
<td>5</td>
<td>4.8</td>
</tr>
<tr>
<td>Significantly</td>
<td>22</td>
<td>21.0</td>
</tr>
<tr>
<td>Very Significantly</td>
<td>77</td>
<td>73.3</td>
</tr>
</tbody>
</table>

This profound influence of standards and tests contrasted with teachers’ assessment of influence from preparation experiences. Of all respondents, 53.3% reported that pre-service preparation experiences had no influence at all on the activities and texts they select. In-service professional development showed modestly greater influence than pre-service preparation; 51.4% stated that it influenced activities and texts significantly or very significantly. But pre-service preparation’s influence remained well behind the reported influence of standards and assessments.

Discussion

The needs assessment validated the presence of factors identified in the literature review. Reading teachers in middle grades reported an orientation toward organizing lessons around skills rather than around the meaning of texts. The activities they chose for the classroom reflected this; teachers surveyed, for example, were much more likely to ask students questions meant to replicate state reading test items than they were to teach background knowledge or to engage students in forming persuasive arguments. While curricular choices may have influenced activities and text selections, teachers continued to incorporate materials of their own choosing and development into classroom lessons. The greatest reported sources of influence over classroom activities were not curriculum but tests and standards.
The picture painted is a stark one. Standards and tests heavily influence how teachers approach teaching reading. Teachers frequently focus on skills such as summarizing and finding the main idea, mirroring the format and substance of state reading tests. In spite of research pointing toward an approach that integrates reading text closely for meaning and building knowledge to apply to reading, teachers in Ouachita Parish, Louisiana were parsing reading comprehension instruction into discrete skills, a pedagogical approach with no basis in evidence.
Chapter 3

Intervention Literature Review

This chapter builds on previous chapters’ conclusions, examining literature related to the range of policy interventions undertaken to address adolescent reading challenges and the particular in-school factors identified in chapter 1 and further defined in chapter 2. Based on a review of intervention literature, the chapter identifies a factor of profound importance related to U.S. adolescent reading levels and proposes to study this factor via an intervention, as described in the subsequent chapter on research design and methodology. This chapter concludes with a brief discussion of that proposed intervention and its relationship to the context of Louisiana public schools.

Implications of Prior Chapters for Potential Intervention Types

Both the literature review in chapter 1 and the needs assessment in chapter 2 provide evidence that teachers of adolescent reading have, as Fisher and Frey (2008) write, “curricularized” (p. 16) the teaching of comprehension strategies, making strategies such as summarizing a text into discrete skills that are perceived as the ends of teaching reading comprehension rather than the means. In such an approach, isolated skills and activities such as summarizing texts or locating texts’ main ideas become the organizing principles of lessons. Fisher and Frey (2008) write, “We can’t find any evidence for the effectiveness of teaching one strategy at a time, especially with pieces of text that require readers to use a variety of strategies to successfully negotiate meaning” (p. 16). Hirsch (2016) is even more pointed, calling this teaching approach a “skills delusion” (p. 37).

If it is true that teachers if adolescent readers have been misled into adopting an approach without basis in evidence, it is reasonable to understand what has so powerfully influenced the
principles and methods of so many educators. Though the needs assessment indicates teachers value their in-service and pre-service preparation less than other influences, there is evidence that preparation and training experiences may positively inform approach to teaching reading (National Reading Panel, 2000). Similarly, the needs assessment validated the influence of standards-based tests over teaching approach, as found by Au (2007), Kane (2016), Moon et al. (2007), and Musoleno and White (2010). The needs assessment further validated the influence, albeit less pronounced, of curriculum over teaching approach. What emerges is a picture of influence over teaching approach at a systemic scale. Preparation, standards and standards-based assessment, curriculum, each offer some possibility for guiding teachers toward an evidence-based approach to teaching not only because of evidence of their influence, but also because of their ubiquity. This chapter now moves toward an examination of literature evaluating interventions in those three areas of influence.

**Review of Intervention Literature**

The theoretical framework for the initial review of literature in chapter 1 adopted EST (Bronfenbrenner, 1994) in order to view adolescent reading at multiple levels of proximity to the child. The framework employed the concepts of human capital (Becker, 1962) and social capital (Coleman, 1988) to distinguish and examine the roles of in-school and out-of-school factors at three levels of an ecological system. The intervention literature considered in this chapter addresses in-school factors at two levels of the ecological system: the teacher-student interaction, or the microsystem, and the sources of influence over how teachers teach, or the exosystem. The theoretical framework, which will guide this review of interventions, will be the integrated model for teaching reading comprehension established in chapter 1. This chapter now turns to a
brief summary of the integrated model, followed by a synthesis of relevant intervention literature.

**Theoretical Framework**

Chapter 1 included evidence supportive of what I hence refer to an integrated model for teaching reading comprehension. Three elements are essential to this model: teaching texts to discern meaning, rather than to learn specific reading skills (Cervetti, Barber, Dorph, Pearson, & Goldschmidt, 2012; NRP, 2000; Fisher & Frey, 2008); applying knowledge of both words within a text and background beyond the text in order to discern meaning (Hirsch, 2016; Pearson, 2013; Shanahan, 2013; Snow & O’Connor, 2016); and increasing the complexity of knowledge required to comprehend texts in order to enhance the student’s capacity to read (Hirsch, 2016; NRP, 2000; Shanahan, 2005). Figure 3.1 depicts this framework of the integrated model, which will serve as an ideal in this chapter, informing the identification of relevant intervention literature and the selection of promising interventions for study.
Figure 3.1 Theoretical framework demonstrating integrated model of teaching reading comprehension.

Factors for Intervention

As discussed, the literature review and needs assessment in chapters 1 and 2 provide evidence that three in-school factors—preparation, standards-based tests, and curricula—significantly influence the principles and methods adopted by adolescent reading teachers. This section will thus synthesize literature related to interventions undertaken in these three areas, with a particular focus on empirical studies of effects related to adolescent reading. For each of the three factors, the literature synthesis will use the framework of the integrated model for teaching reading, described above, as its guiding ideal, assessing interventions in curriculum,
assessment, and preparation, not just for their potential to improve reading skills, but also for their capacity to effect the elements of the integrated model. Finally, in order to assess an intervention’s relevance to the large-scale problem of U.S. adolescent reading performance, the synthesis will consider the scale at which the intervention is enacted, in light of the nationwide scale of the problem.

**Teacher Preparation**

The NRP (2000) report states that “[t]he area within comprehension strategy instruction that currently seems to have the most potential for moving the field along is teacher preparation” (p. 4-120). However, in a seminal report on the quality of pre-service education preparation programs in America, Arthur Levine (2006) writes, “We don’t know what, where, how, or when teacher education is most effective. This means the education our teachers receive today is determined more by ideology and personal predilection than the needs of our children” (p. 19). This contradiction between the centrality of preparation to the mission of adolescent reading comprehension and the inchoate design of preparation curricula gives cause to consider interventions addressing this factor. What follows is a synthesis of studies of two distinct types of interventions that fit within the preparation category. The first, methods preparation, relates to interventions in which teachers are trained to implement a model or technique associated with the principles of the integrated teaching model. The second, preparation portals, relates to recruitment and training programs themselves and their effect on students’ reading skills.

**Methods Preparation.** The NRP report (2000) captures both the technical sophistication required of teachers who use strategies to help students comprehend a text’s meaning:

Successful teachers of reading comprehension must respond flexibly and opportunistically to students’ needs for instructive feedback as they read. To be able to do
this, teachers themselves must have a firm grasp not only of the strategies that they are
teaching the children but also of instructional strategies that they can employ to achieve
their goal. Many teachers find this type of teaching a challenge, most likely because they
have not been prepared to do such teaching. (p. 4-119)

Teachers must understand not just what students must do in order to read effectively, but also
what instructional methods they must use in order to transfer understanding (NRP, 2000). As
one example of how teachers can be effectively trained to employ specific methods, multiple
studies indicate that it is beneficial for teachers to be trained to use dialogue and questioning to
help students discern the meaning of texts (Applebee, Langer, Nystrand, & Gamaoran, 2003;
conduct a seminal study of the reciprocal teaching method in which teachers guide students
through a sequence of interactions with texts: summarizing, questioning, clarifying, and
predicting. Having trained teachers to implement the reciprocal model, the researchers divide 24
struggling readers into four groups of six, one group of which experiences full reciprocal
teaching on a set of selected passages. Another group receives no treatment, and the other two
receive modified treatments. On pre- and post-tests, the reciprocal teaching sample demonstrates
significant, positive effects relative to peers, with no drop-off after an eight-week maintenance
period following the post-test (Palinscar & Brown, 1994).

Reznitskaya, Anderson, and Kuo (2007) conduct an experimental study in which two of
six middle school classes engage in collaborative reasoning discussions without explicit
instruction on argumentation technique. Two other classes also participate in the discussions,
but these classes also experience explicit instruction in argumentation as part of the treatment.
Two classes experience neither treatment. Having ensured that the groups were comparable in
terms of pre-existing reading proficiency, the researchers find that students explicitly taught the principles of argumentation recalled and applied the principles in interviews with greater proficiency than did students in other groups. When taught an explicit schema for making arguments, students improve their reading skills through dialogue regarding texts (Reznitskaya et al., 2007).

Sample size is a notable limitation to the aforementioned findings. In each case, samples were limited to several schools or classrooms, a particularly notable limitation given the nationwide scale of the problem of practice. Applebee, Langer, Nystrand, and Gamaoran (2003) study dialogic teaching methods across a larger sample of 1,111 students across 64 schools in 5 states. Observers collected evidence of dialogic practice, extended curricular discussion, and rigorous academic demands in each class using the CLASS observation instrument. Students also completed three written tasks over the course of the year. Using hierarchical linear models to create expected scores on written tasks, the researchers calculated the effect of dialogic teaching on student learning, concluding that the combination of higher academic expectations and dialogic pedagogy contributed to significant reading comprehension gains irrespective of prior reading skill in the student. This study provides further evidence that teacher preparation on specific methods, as a category of interventions, has a base of positive evidence, even as the scale of the interventions may be limited by their interaction-intensive natures. This chapter now turns to the effects of the credentialing programs that prepare teachers for service, distinct from the more focused preparation interventions associated with one specific method.

**Preparation programs.** Participation in credentialing programs, including colleges of education and in-service programs such as National Board Certification, can be evaluated as an intervention with effects on students’ reading skills (Goldhaber & Cowan, 2014; Henry et al.,
Repeated studies indicate that pre-service preparation programs, such as colleges of education or alternative certification programs, have mixed effects on reading comprehension (Goldhaber & Cowan, 2014; Henry et al., 2014). Even when there are positive initial effects of individual programs, these effects may be mitigated by other factors, as with the high levels of attrition among graduates of certain programs (Goldhaber & Cowan, 2014).

Henry et al. (2014) compare the value added to student test scores by North Carolina teachers with three or fewer years of experience. The researchers group teachers by pre-service preparation program type: traditional colleges of education; Teach For America (TFA), the national recruiter of recent non-education college graduates; and other preparation programs for individuals already holding bachelor’s degrees. The researchers use five years of test score data across grades 3-11 and four academic subjects to calculate the value added to student test scores by individual teachers, finding that alternatively certified teachers generally underperform traditionally certified teachers. TFA teachers significantly outperform traditionally prepared teachers, but, notably for these purposes, not in middle school reading, a subject in which there was no significant difference (Henry et al., 2014). This finding corresponds with a Texas-based study of the effectiveness of TFA teachers in the first two years of teaching (Turner, Goodman, Adachi, Brite, & Decker, 2013), which finds no significantly unique effects of TFA teachers on middle school reading in the early years. This finding differs from the study’s conclusions regarding later-stage TFA teachers, who significantly outperform experienced teachers prepared through other programs in middle school reading gains (Turner et al., 2013).

This mixed picture in terms of the relationship between middle school reading gains and pre-service preparation programs extends to in-service preparation programs, or professional development programs. Cowan and Goldhaber (2016) find significant, positive effects on
mathematics scores of veteran teachers prepared through the nationally recognized National Board Certification Teacher (NBCT) initiative. NBCT teachers of reading demonstrate no unique effects on student learning (Cowan & Goldhaber, 2016). Fisher and Dickenson (2005) find similarly insignificant effects of NBCT teachers in middle school reading performance in a statewide study in South Carolina. Interestingly, Cowan and Goldhaber (2016) find no distinction in effects on middle school reading scores between the general population of teachers and those teachers who initially failed NBCT assessments but later qualified for NBCT. NBCT status itself thus appears to have little impact on middle school reading effects, even as NBCT teachers demonstrate modest, positive contributions to reading skills among adolescents. Therein is a limitation to the notion of preparation program as distinct interventions. NBCT teachers and TFA alumni may be associated with positive results (Cowan & Goldhaber, 2016; Fisher & Dickenson, 2005; Henry et al., 2005), but these findings offer no clarity as to what aspect of the program itself—recruitment, selection, or preparation—contributes to those outcomes.

**Curriculum and Reading Intervention Programs**

This chapter now turns toward a focus on the instruments teachers use to teach reading comprehension, examining the texts students are assigned, how texts are sequenced, and how they are used by teachers. The inquiry will distinguish between two types of tools involving assigned texts. The first, curriculum, refers to the broad array of instructional materials teachers use regularly with all of their students, including texts, activities, tasks, lessons, units, and sequencing guidance. The second, reading interventions, addresses instruments that are specifically developed for use with struggling adolescent readers, and that often are used in a setting other than the whole class. The review will continue to focus on middle school and high
school grade levels, discussing effects on younger students when helpful to illustrate larger points or trends.

**Curriculum.** Steiner’s (2017) comments on the American approach to developing and selecting curriculum provide context for an examination of research surrounding this curriculum in the U.S.

[A]ttention to the academic content – *what* we teach – is only now beginning to emerge as a serious lever in education policy. For starters, with the exception of California and a few districts across the country, we have no clear idea what is being taught in America’s classrooms. This would strike other countries as ludicrous: As Common Core, Inc., discovered in 2009, while top-performing countries differ in many facets of education policy, they all share a commitment to high-quality curricular content. (p.1)

Steiner’s (2017) statement indicates a difficult truth for those examining what is known about curricula and their effectiveness in U.S. schools: knowledge of taught content borders on non-existent, and knowledge of what content is working is rarely brought to bear on public policy.

Mathematics curricula demonstrate mixed effects on middle school and high school mathematics skills. Using a series of online lessons, Jackson and Makarin (2017) assign groups of middle school math teachers to one of three conditions: free access to online lessons, free access to online lessons with active support of the lessons, and no use of online lessons, the last being a control group. The researchers find that gains made on math assessments in the full treatment class are equivalent to the effect of a teacher moving from average to the 80th percentile in terms of effectiveness, or the effect of reducing class size by 15%. The researchers further find that the online lessons had greater effect in classrooms led by teachers achieving lesser gains with students in prior years. The researchers attribute this to a division of labor that
CURRICULUM-SPECIFIC READING TEST

takes place when published materials enhance the productivity of a given teacher; curricular materials, they assert, can enhance the quantifiable effectiveness of the teacher (Jackson & Makarin, 2017).

In a mixed-methods study, Kane et al. (2016) use stratified random sampling to identify schools representative of the general population in five states. The researchers survey math and reading teachers in Grades 4 and 5, conducting interviews regarding curriculum choices and steps taken to implement curriculum. The researchers then associate performance on state tests with practices demonstrated in surveys and interviews. In mathematics, results showed that greater value added to student skills on state tests was associated with greater exposure to professional development days, inclusion of testing data in personnel evaluations, and observation feedback with feedback tied directly to state instructional standards. Students using particular math curricula aligned with CCSS saw a positive effect ($SD = .1$) in student performance on standards-aligned assessments, versus a negative effect ($SD = .15$) for students using non-aligned curricula.

In a subsequent study, Blazar et al. (2019) use textbook adoption data and state test score data in six states to measure the effect of Grade 4 and 5 mathematics curricula on student learning. Unlike the earlier study, Blazar et al. (2019) find no significant effects across curricular products generally. Blazar et al. (2019) find modest variations ($SD = .02-.04$) when students are exposed to a given curriculum regularly versus when they’re exposed less frequently. But the study provides large-scale evidence, subsequent to the full implementation

---

19 Kane and Staiger are members of both this research team and the research team conducting the aforementioned 2016 study.
of CCSS, that curriculum adoption, on its own and without regard for differentiating conditions or implementations, is unlikely to contribute to significant effects.

Switching to certain reading curricula in grades K-2 has an established history as an effective intervention (Borman, Dowling, & Schneck, 2008; Smith, Ross, Faulks, Casey, Shapiro, & Johnson, 1993). Borman et al. (2008) find that adopting Open Court Reading is associated with gains of 10 percentile points when the performance of treated students is compared with non-treatment students. Smith et al. (1993) find gains of 19 percentile points for students treated with the Success for All curriculum. However, Kane et al. (2016), whose focus is on students in Grade 4 and 5, find that no individual curriculum or implementation strategy in literacy was associated with significant gains in reading skills, contrasting with significant effects of certain mathematics curricula found in the same study. In a randomized control trial measuring the effectiveness of a particular middle school literacy textbook series, Berry, Eddy, Fleischer, Asgarian, and Malek (2007), similarly conclude that the textbook choice itself contributes no significant effect. But Berry et al. (2007) also conclude that the intensity and fidelity of implementation, as measured by the percentage of teaching time in which the textbook was in use, was associated with significant results, and that among teachers using the treatment textbook regularly and consistently, students demonstrated significant effects. This finding is consistent with conclusions in a meta-analysis of curricular interventions conducted by the Best Evidence Encyclopedia (BEE) at Johns Hopkins University, which finds that in upper elementary grades, the effect of combined curricular and preparation approaches on reading skills yielded greater gains than did curriculum or preparation on their own, even though such combined interventions are rare (Slavin, Lake, Chambers, Cheung, & Davis, 2009). As the NRP
report (2000) surmised, teachers must be technically prepared to implement specific reading curriculum, yet such preparation rarely occurs.

More directly related to the problem of U.S. adolescent reading performance, the BEE found limited evidence of effectiveness for any middle school or high school reading curriculum.

No studies of secondary reading curricula met the criteria for this review. This is surprising in light of the widespread use of such programs in middle and high schools throughout North America. It is not the case that the inclusion standards applied in the present review excluded many studies. Despite an extensive search, only 14 studies of reading curricula were located. (Slavin, Cheung, Groff, & Lake, 2008, p. 292)

As noted in chapter 1, Louisiana has developed a middle-grades reading and writing curriculum called ELA Guidebooks 2.0, which is designed using the principle of knowledge domains (Hirsch, 2005) that group texts around common ideas, allowing students to build coherent bodies of knowledge as they read (Pondiscio, 2017). In this way, the curriculum’s design is similar to the organizing principle of the well-known, multi-subject Core Knowledge curriculum founded by Hirsch (Pondiscio, 2017), making research on Core Knowledge particularly important to understanding potential interventions in the Louisiana context. In addition to a number of studies of Core Knowledge conducted by the Core Knowledge Foundation or by individual school districts, CSRQ/AIR (2006) identified three studies that met the meta-analysis’s standards for evidence of effectiveness. Datnow, Borman, Stringfield, Overman, and Castellano (2003) compare the effects of whole-school Core Knowledge implementation on Grade 4 students when to effects on non-treated Grade 4 students in similar schools, yielding no differences in effects on the general population but significant, positive effects among students classified as English learners. In a quasi-experimental study, Taylor (2000) finds similarly positive effects of Core
Knowledge curriculum in reading, mathematics, and social studies across grades 3, 4, and 5. And Stringfield, Datnow, Borman, and Rachuba (2000) find large, positive effects of the Core Knowledge curriculum in language, social studies, and science, but not in reading and math. Notably, Stringfield et al. (2000) also find that effects vary with the fidelity and intensity of curriculum implementation. The effect of moderate to high rates of implementation on student outcomes is significant explaining most of the curriculum’s positive effects on learning.

This finding related to combined curriculum and implementation interventions and effects is a recurring theme. While earlier studies indicated that curriculum adoptions on their own in middle grades yield significant, positive results in some instances (Kane et al., 2016; Makarin & Jackson, 2017), a large-scale study of mathematics curricula across six states indicated that curricular adoptions on their own were not a cause of variation in student performance (Balzar, 2019). In reading, effects of curricular materials in Grades 3 and above are often associated with effects of implementation rather than with traits of the curricular materials exclusively (Berry et al., 2007; Slavin et al., 2008; Stringfield et al, 2000). This trend in the literature may be instructive to identifying high-potential interventions, which could encourage not only selection of certain curricular materials, but also complementary implementation supports.

Relatedly, there is evidence that implementation of curriculum among U.S. ELA schools and teachers is haphazard. Opfer et al. (2016) find that reading teachers are significantly less likely than teachers of mathematics to change instructional materials when state standards change. Both Kane at al. (2016) and Opfer et al. (2016) found that a large majority of reading teachers continue to develop their own instructional materials. And, as discussed in chapter 1, Opfer et al. (2016) find that aspects of CCSS implementation related to instructional materials,
CURRICULUM-SPECIFIC READING TEST

including the imperative to teach texts reflecting complexity appropriate to a student’s age, are largely ignored by, or unknown to, ELA teachers. Chingos and Whitehurst (2012) note that even states, which are responsible for establishing learning standards, typically neither monitor nor report the curricular selections made by school systems. The failure of reading teachers to make ostensibly more productive curricular choices, and the failure of those curricula to demonstrate positive effects at the middle school and high school level, could be a function of an absent system of support, guidance, and implementation, a notion this chapter will revisit when discussing the role of assessment.20

Reading Intervention Programs. Distinct from curriculum, reading intervention programs are intended for use with struggling readers, often in settings apart from the traditional classroom. To evaluate the effects of these tools, this literature review will synthesize two studies of a popular middle grade intervention product, Read 180. Read 180 is a package of online content coupled with a whole-group and small-group instructional model meant to improve reading skills of struggling adolescent readers (Meisch et al., 2011). Published by Scholastic, Read 180 is typically adopted on a schoolwide basis, becoming the ELA curriculum for designated students. In a Read 180 class, students spend time both online engaged with material and content customized to the student’s reading skill deficiencies and with a teacher receiving direct instruction on reading strategies. Meisch et al. (2011) examine five years of state test score data, conducting an experiment in which students in treatment schools participate in a program.

20 It is perhaps evident, but worth noting, that all three intervention types in this chapter are part of a scaled, nationwide system of related components and do not wholly exist at a given moment in time or entirely unto themselves. As discussed in chapter 1, standards-based reforms gave birth to tests, curricular changes, and other interventions that have evolved over time. In considering an intervention to address a national problem of practice, this dissertation seeks an evolution of standards-based reform rather than a distinct turn in a new direction or a novel intervention unrelated to reforms of the last 30 years.
in Read 180 for three years. Results show significant effects in the second year of the treatment related specifically to reading comprehension skill. The effects on comprehension are not evident after either one year or three years. On a schoolwide basis, schools experiencing the treatment show no significant effects in performance relative to schools not receiving the treatment in any year, a finding the researchers speculate is attributable to low teacher participation in professional development and high levels of teacher turnover over the period studied (Meisch et al., 2011).

Kim, Sampson, Fitzgerald, and Hartry (2010) also examine the effects of Read 180 on reading efficiency, oral reading fluency, and reading comprehension. The researchers establish a control group of 10 classrooms participating in standard district after-school activities and a treatment group of 10 classrooms participating in Read 180. The classrooms span Grades 4, 5, and 6. Children participating in both the control group and the treatment complete a number of assessments to measure baseline skills in specific areas. After 23 weeks of the treatment, the researchers conduct post-tests, as well as surveys of treated students to identify elements of the treatment experience. The study finds no impact of the treatment on reading comprehension or reading efficiency (Kim et al., 2010). Unlike Meisch et al. (2011), Kim et al. (2010) find no flaw in implementation to which results may be attributable. Kim et al. do note that the 60-minute after-school period in which they applied the treatment forced the researchers to condense Read 180 into a shorter block than the model typically necessitates, eradicating direct instruction on vocabulary that may be essential to increasing the skills of struggling readers. The researchers conclude, “Ultimately, findings from this study suggest that it may be unreasonable to expect one reading intervention—even a comprehensive, mixed-method approach to literacy instruction—to address all areas of reading weakness simultaneously” (p. 24).
State Reading Tests

This section of the chapter examines the effects of state literacy assessments required under federal and state law. In this literature review, up to this point, state assessments have served as means by which data related to intervention effects were gathered rather than an intervention unto themselves. This section of the inquiry will treat state assessments as the interventions themselves, first examining studies of the effect of state tests on teaching approach and next addressing the effects of state-based and other commonly recognized assessment series on student learning.

Influence of State Tests on Teaching Methods

Moon et al. (2007) study the effects of standardized testing on curriculum selection and classroom pedagogy in middle schools. In the first phase of the study, the researchers administer a 99-item survey to teachers and administrators regarding attitudes and behaviors. The authors also conduct a principal component analysis to validate the intended scales of the survey, keeping 69 of the original 99 items in the survey for inclusion in the study. The researchers seek a 1% stratified random sample of teachers nationwide for completion of the survey, gathering data by mailing the survey to members of the intended sample. In the second phase, the researchers conduct 21 focus groups of between three and five participants each, involving purposive samples of both students and teachers. The discussions were recorded and then transcribed. The researchers also conduct classroom observations and triangulate data collected from the transcripts with field notes, coding and sorting groups of data in order to determine patterns across the data. Data from both phases of the study confirm that standardized tests administered by states influence the curricular and instructional decisions made by teachers, also a finding of the needs assessment detailed in chapter 2. The effect is most pronounced in
teachers serving predominantly low-income students. In particular, teachers associate
standardized testing with prescribed curricula and intense preparation activities specific to the
test, with less focus on long-term projects (Moon et al., 2007). This finding supports the
conclusion of the NAE (2009) discussed in chapter 1, which states that standards and standards-
based tests have had the effect of incentivizing educators to focus on specific goals more than the
effect of manifesting evidence-based teaching methods. Tests that measure skills may
incentivize the teaching of skills reflected on the test, even though there is no basis in evidence
for reading comprehension being a function of discrete comprehension skills (Fisher & Frey,
2008).

These findings are similar to conclusions Au (2007) reaches in a study of the impact of
state standardized tests on curricular and pedagogical choices. Au conducts a quantitative
metasynthesis of 49 qualitative studies involving 845 educators across 38 districts and all grade
levels. In the first phase, the researcher develops a template for coding based on initial reading.
In the second phase, the researcher codes data gathered from all sources and creates new
categories for coding as the need arises throughout the reading. Having coded the textual data,
the researcher seeks patterns, anomalies, and high-frequency pairs and triplets of code
combinations. The narrowing of curricular content knowledge is the strongest association in the
study. Tests were also associated with fragmentation of knowledge within the curriculum,
parsing content into discrete pieces rather than teaching the whole substance of texts or bodies of
knowledge (Au, 2007). These findings are consistent with the influence of tests noted both in
chapter 2 and by Kane et al. (2016), who noted a significant shift across five states towards
greater volume of writing in ELA classrooms given the influence of writing-intensive state tests.
In this case, schools minimize teaching social studies and science in order to spend more time on
reading and mathematics (Au, 2007). But, as established in chapter 1, students in middle school
and high school read to build knowledge that will inform future reading (Chall, 1983; Hirsch,
2005; Willingham, 2006). If standards-based tests diminish the impetus to read across the
curriculum, they may not only neuter the cross-curricular intentions of CCSS, but also
undermine the base of background knowledge needed for a student to comprehend what she
reads.

**Effects of State Tests on Student Learning**

Though NCLB included no grant funding or program provision related to adolescent
reading, in the way its Reading First program provided resources and support for primary grade
reading improvements, the law’s requirement for annual state reading assessment in middle
grades marked a greater federal emphasis on adolescent reading than had the federal statutes that
preceded it (Hauptli & Cohen-Vogel, 2013). Dee and Jacob (2011) measure the ultimate effect
of NCLB’s testing and evaluation provisions by assessing changes in state NAEP math and
reading scores prior to and after NCLB implementation. The study used a comparative
interrupted time series approach, which contrasts changes in NAEP achievement trends among
students subject to NCLB with the same change in NAEP achievement trends for a comparison
set of students who were not as affected by the NCLB testing and accountability regime because
their states already had similar testing and accountability regimes in place prior to the onset of
NCLB. Dee and Jacob (2011) found that NCLB was associated with significant, positive
changes in both Grade 4 mathematics and Grade 8 mathematics. However, they find no change
associated with Grade 4 reading performance.²¹

---

²¹ ⁸th grade reading had only been assessed twice by NAEP prior to NCLB and thus was not
included in this assessment.
Lee and Reeves (2012) take a similar approach to evaluating the impact of NCLB, using a hierarchical linear modeling latent variable regression to account for pre-NCLB distinctions among states and to evaluate post-NCLB progress in light of these distinctions. The researchers find that while state effects in Grade 4 and 8 reading vary widely, the majority of states either maintain pre-NCLB trends or see slumps in reading trends. This finding contrasts with mathematics, in which, as Dee and Jacob (2011) conclude, gains accelerated after the passage of NCLB. Lee and Reeves (2012) also find that reading skills gaps among historically disadvantaged peers and their colleagues widened in post-NCLB, while such gaps narrowed in mathematics. As with curricular interventions, strategies associated with advances in math achievement fail to advance adolescent reading skills.

**Testing Series Other Than State Tests**

The research cited in this literature review paints a poor picture of the impact of NCLB-era state reading assessments and accountability systems on adolescent reading skills, in contrast to the generally positive effects of NCLB-era state math tests (Dee & Jacob, 2011; Lee & Reeves, 2010). It is worth considering whether the lack of reading comprehension progress under NCLB is associated with the design or elements of reading tests commonly used in the U.S. No widely used assessment series in the U.S. other than the AP series developed by the College Board, for example, provides a reading list of suggested texts (Steiner, 2018). This is a remarkable finding given the evidence base (NRP, 2000) that supports making meaning of whole texts as a fundamental and necessary approach to improving reading comprehension. Relatedly, ELA teachers in Louisiana, who use standards-aligned curricular materials at the highest rates in the nation, report only average faith that their curricula prepare students for the substance of state ELA tests (Doan et al., 2022). If tests are, as NAE (2009) asserts, motivators of teacher actions,
presumably the lack of reading lists, or the lack of connection between curriculum and test content generally, creates less incentive for teachers to focus on building knowledge of specific texts. Furthermore, no commonly used assessment series other than AP requires long-form essays that allow students to demonstrate knowledge across different texts or fields of study (Steiner, 2018). If comprehension can be demonstrated at increasingly complex levels, requiring increasingly sophisticated integration of the reader’s knowledge (Shanahan, 2005), minimizing students’ opportunity to write in response to texts would also seem to diminish a teacher’s incentive to demand such sophistication.

As one example of a nationwide testing series that emphasizes specific texts, the Cambridge International A-level assessment series administered in the United Kingdom requires students to demonstrate knowledge of eight specific texts across different genres (Cambridge International Examinations, 2017). Students draft analytic papers in response to passages and prompts based on those particular texts, which are pre-disclosed on a reading list. The International Baccalaureate (IB) program used widely internationally and in some U.S. school systems includes a multi-disciplinary curriculum with an assessment series that is requires students to demonstrate knowledge of what has been taught in the curriculum (Conley, McGaughy, Davis-Molin, Farkas, & Fukuda, 2014). Using multiple regression models, Conley et al. (2014) find that students completing the IB program are significantly more likely than peers to demonstrate college-ready skills when taking course placement examinations after college admissions. The researchers further find that students completing at least four IB courses are significantly more likely to persist through college to completing a baccalaureate degree (Conley et al., 2014). The integration of curricular content into one test at the end of a semester while in secondary school, students report in the study, prepares them particularly well for the challenge
of integrating broad sets of information into one task at the conclusion of college courses (Conley et al., 2014). Thus, the researchers speculate, the design of the IB assessment series contributes to an intellectual sophistication needed for post-secondary success. This is an important conclusion given the connection of reading deficits to post-secondary educational outcomes in the U.S (Au & Raphael, 2000; Conley et al., 2014; Jofus, 2002; Kamil, 2012). But Conley et al. (2014) also offer a counterfactual to the failure of the NCLB testing regime to demonstrate reading effects (Dee & Jacob, 2011; Lee and Reeves, 2012). By requiring students to integrate knowledge of texts in a demonstration of intellectual sophistication, IB may offer an example of an assessment series that creates incentives for teachers to focus on assisting students to make meaning of whole texts and to build coherent bodies of knowledge.

**Conclusion**

The problem of stagnant adolescent reading proficiency in the US. is vexing both for its persistence and for its enormous scale (NCES, 2016; OECD, 2015). As its theoretical framework, this synthesis of intervention literature adopts an integrated model for teaching reading, in which students apply knowledge from outside a text to discern the meaning of a given text, increasing over time the complexity of what they know and the text they read. This framework calls for a review of intervention literature that examines effects on both teacher practice and on the problem of reading comprehension performance itself.

The synthesis has established a pattern of failure among most interventions to achieve both positive reading outcomes and significant scale among middle school and high school students. Preparation programs vary widely in terms of quality, and the effectiveness of candidates within programs varies widely (Goldhaber & Cowan, 2014). Preparing teachers for specific pedagogical methods shows promising results, but often at limited scale (Palinscar &
CURRICULUM-SPECIFIC READING TEST

Brown, 1984; Reznitskaya et al., 2007; Reznitskaya et al., 2012), raising questions about the viability of such strategies at a scale commensurate with the problem of practice. Reading intervention programs such as Read 180 operate across entire school systems but show no positive effects on reading skills (Kim et al., 2011; Meisch et al., 2010). Though there is evidence that reading curriculum can be an effective intervention when coupled with aligned systems of preparation, ELA curricular interventions on their own show minimal positive effects beyond primary grades (Berry et al., 2007; Kane et al., 2016), while mathematics curricular interventions have shown significant, positive effects (Jackson & Makarin, 2017; Kane et al., 2016). Most concerning of all given its scale, the enactment of a nationwide testing and accountability regime via NCLB shows positive results in mathematics and no positive effects in reading (Dee & Jacob, 2011; Lee and Reeves, 2012).

State tests required under NCLB are associated with schools narrowing the subject matters to which students are exposed (Au, 2007). Teachers also may “curricularize” (Fisher & Frey, 2008) reading strategies as a consequence of tests, teaching discrete skills to students rather than asking them to discuss and discern the meaning of texts (Musoleno & White, 2012). The IB program, which aligns curriculum with embedded assessments also positive effects on college readiness, college completion, and both academic and non-academic skills related to college success (Conley et al., 2014). IB’s assessment design, however, is distinct from those of reading tests in all 50 states in that it requires prior exposure to specific texts and bodies of knowledge, and in that it measures students’ understanding of those elements rather than treating reading as a set of distinct skills that can be taught and evaluated irrespective of a student’s prior knowledge of texts and of other relevant context (Conley et al., 2014). This design distinction between the
unsuccessful NCLB middle grades testing intervention in reading and IB assessments provides cause for further inquiry and study.

The next chapter will outline an intervention involving a new design for a state reading test, one that shares essential traits with the IB model. Rather than measuring reading skills via texts undisclosed prior to the test, the test will ask students to respond to prompts based on pre-disclosed texts and knowledge domains with which students are already familiar. Administered three times annually in short sessions at the conclusion of curricular units, the test will also create incentives for teachers to focus on building students’ base of knowledge by integrating social studies content and knowledge of literary texts on the same on the assessment.

The LDOE was approved by the USDOE to pilot such an assessment under the IADA program (Sentell, 2018). Middle schools in 10 Louisiana school systems (LEAs) participated in the assessment program during the 2019-2020 school year, and 8 districts sustained participation in 2021-2022.

Given factors identified throughout this dissertation, the intervention’s logic model states that the new assessment may create different incentives and direction for teachers in the treated group from incentives experienced by teachers whose students take the traditional state standardized test. The model hypothesizes that teachers will have greater incentive to adopt an integrated model for teaching reading, rather than organizing lessons around teaching discrete comprehension skills. Treated teachers, the model states, will use assigned texts more frequently in class, and will assist students in using texts to strike a balance between close reading of the text itself with application of background knowledge from beyond the text.
Chapter 4

Methods and Procedure

As discussed in the conclusion of chapter 3, the state of Louisiana was granted permission by USDOE to conduct an experimental middle grades reading test from the 2019-2020 school year through the 2023-2024 school year, permitting the state to administer an alternative to the conventional state testing series during that period (Sentell, 2018). USDOE granted the authority to administer an experimental test under the IADA provision of the Every Student Succeeds Act (ESSA) of 2015, which allows states to implement promising testing models on an experimental basis (Sentell, 2018). As detailed later in this chapter, the school year in which the Louisiana IADA test was first administered, 2019-2020, was interrupted due to the COVID-19 pandemic, and the test’s full administration was not completed. In the following school year, 2020-2021, Louisiana resumed conventional statewide testing (Sentell, 2021), but the IADA was not administered given the complexities of frequently interrupted schooling during a second year of pandemic impacts (C. Johnson, personal communication, December 14, 2021). In the 2021-2022 school year, the Grade 7 IADA assessment was fully operational for participating school districts, two years after its originally scheduled implementation (C. Johnson, personal communication, December 14, 2021).22

This chapter will describe a study of the implementation of the Louisiana IADA test and of the intervention’s effects on teacher principles and methods. Just as the intervention itself was disrupted due to the COVID-19 pandemic, this study has been complicated and extended by the disruption. I include in the study the interrupted 2019-2020 IADA test administration and

22 In March 2020, unrelated to either the pandemic or this study, I vacated the role of state superintendent of education in Louisiana. This has no evident effect on the study’s design and implementation. I mention it here to provide appropriate context for the reader.
associated data collection undertaken; I also include in the study the IADA implementation and associated data collection from 2021-2022, the IADA’s first, complete test administration cycle. As described below, the resulting study provides a picture of implementation and outcomes at different points in time, throughout the COVID-19 disruption to schooling and testing.

**IADA Test Intervention**

Developed in partnership with the Center for Assessment, Johns Hopkins Institute for Education Policy, the assessment publisher NWEA, and Odell Education, Louisiana’s IADA assessment is a companion to ELA Guidebooks 2.0, a literacy curriculum for students in Grade 3 through Grade 8 developed by the Louisiana Department of Education (LDOE) and the publisher LearnZillion (LDOE, 2019). While the assessment involves both multi-select and written-response items, it involves few “cold reads” in which students are asked to explain or analyze previously undisclosed passages with which they are not ostensibly familiar, a prominent feature of conventional state ELA tests (LDOE, 2019). Students completing the IADA test instead typically respond to “hot reads,” passages from Guidebooks 2.0 texts they have already read in class, or “warm reads,” passages from texts with thematic connections to Guidebook 2.0 texts (LDOE, 2019).

ELA Guidebooks 2.0 contains five segments per grade level, called units, each organized around a specific theme, or domain of knowledge, and each including a whole book called an “anchor text” (LearnZillion, n.d.). As part of the IADA assessment, students in participating school systems completed three assessment forms over the course of the year, at pre-determined

---

23 The dissertation advisor to this project, Dr. David Steiner, and two committee members, Dr. Ashley Berner and Dr. Alanna Bjorklund-Young, are faculty at the Johns Hopkins Institute of Education Policy.
times that correspond with a school’s or school district’s completion of the appropriate unit (LDOE, 2019; LDOE, 2021). For each of the first two test administration periods, conducted in October and February, school systems chose one Guidebooks unit from two options. For the final test administration, conducted in May, all students took the same form based on the same Guidebooks unit, and completed an essay that asks students to draw on knowledge of texts read over the course of the year (LDOE, 2019). Table 4.1 details specific text options from which districts chose, as well as differences in the test forms across the three administration periods.

Table 4.1

**IADA Grade 7 Test Form Differences**

<table>
<thead>
<tr>
<th>Time period</th>
<th>Anchor-text/unit selections</th>
<th>Form description</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td><em>The Giver</em> (Lowry, 1993); <em>Written in Bone</em> (Walker, 2009)</td>
<td>Multi-select and written response to text; essay response to Guidebook text</td>
</tr>
<tr>
<td>February</td>
<td>Memoir unit selections; <em>A Christmas Carol</em> (Dickens, 2015)</td>
<td>Multi-select and written response to text; essay response to related text</td>
</tr>
<tr>
<td>May</td>
<td><em>Behind the Scenes</em> (Keckley, 2012)</td>
<td>Multi-select and written response to text; essay response to related text; and essay recalling yearlong knowledge</td>
</tr>
</tbody>
</table>

The *LEAP ELA Guidebook Grade 7 Test Administration Manual* (LDOE, 2021) details procedural and technical requirements for implementing the IADA test at a school site. The test is taken via an electronic device and does not allow for hand-written responses, though students

---

25 Henceforth, when referring to Louisiana’s IADA test, I wil be referring to the Grade 7 test, which was the first developed, piloted, and made fully operational. This study addresses Grade 7 only. The state and its partners have developed pilot tests for other grade levels aliged to other curricula; I do not study these tests here.
are permitted to use a pencil and paper to outline notes as they complete the assessment. Each of the three test periods includes three sub-sections students complete (LDOE, 2021). Students first answer a series of multi-select questions to demonstrate knowledge of texts read in the relevant Guidebooks 2.0 unit or to respond to texts demonstrative of related knowledge; this sub-section may take up to one hour. Students subsequently write an end-of-unit essay responding to a prompt requiring that they draw on textual knowledge from the relevant Guidebooks 2.0 unit; this section may also take up to one hour. A shorter, third section asks students to respond to a text unrelated to Guidebooks 2.0 in order to demonstrate reading comprehension levels; this section may take up to 20 minutes and may be completed on a different day from the other two sections. To conclude the school year, the IADA test also includes a summative essay in which students draw on knowledge gleaned from across the three units to respond to a prompt.

The LDOE reports to school districts and to parents results of each unit-based test for each participating student. Reports to parents and school systems include a summary score on a scale of 9, reported alongside average performance at state, district, and school levels. School districts also receive sub-scores that detail performance in three skill domains: the student’s knowledge of the Guidebook unit texts (as demonstrated responding to multi-select items), the student’s application of knowledge learned in the unit to other tasks and contexts (as demonstrated in unit-based essays), and the student’s synthesis of knowledge learned throughout the year (as demonstrated in a summative essay). Not only the test design, but also the reporting format, underscore the centrality of reading to build knowledge and responding to texts by activating learned knowledge.
CURRICULUM-SPECIFIC READING TEST

2019-2020 IADA Administration

In the 2019-2020 school year, ten Louisiana school districts—Assumption Parish, Central Community, Claiborne Parish, Evangeline Parish, Grant Parish, Lincoln Parish, Monroe City, Ouachita Parish, Rapides Parish, and St. John the Baptist Parish—and one charter school management organization. Redesign Schools used the IADA assessment for all Grade 7 students in lieu of the traditional Louisiana Education Assessment Program (LEAP) ELA and social studies assessments (C. Johnson, personal communication, July 27, 2019). Participating students completed end-of-unit assessment forms in the October and February testing periods. Grade 7 teachers in participating districts participated in one-day statewide professional development sessions related to each end-of-unit test form, its format, and the principles underlying its design (C. Johnson, personal communication, July 27, 2019).

A regulation passed by the state Board of Elementary and Secondary Education (BESE) in June 2019 stipulated that the scores of 2019-2020 Grade 7 students completing the IADA assessment could count toward a district’s school performance rating under the state’s accountability system, but only if the state determined that average IADA scores in a school or district was equal to or greater than 2018-2019 LEAP performance in those districts and subjects (BESE, 2019). If aggregated scores on the IADA assessment were lower than 2018-2019 LEAP performance, the policy stipulated, the 2018-2019 scores would hold over, effectively freezing a school or district’s summative rating in ELA and social studies proficiency (BESE, 2019). Thus,

---

26 The LDOE termed the 2019-2020 administration a “pilot,” while the 2021-2022 administration was termed “operational.” I will not treat these labels as indicative of any broad, notable differences in either the assessment format or the conditions under which the tests were administered. Generally speaking, the tests administrations were similar in content and condition. Where there are differences, I will establish those on an idiosyncratic basis.
administrators and teachers potentially implemented the Grade 7 IADA test with the knowledge that resulting scores could boost school ratings but could not diminish school ratings.

As discussed in the introduction to this chapter, the state of Louisiana suspended state tests and related professional development activities as of March 2020 due to the COVID-19 pandemic (LDOE, 2020). The 2019-2020 IADA intervention, therefore, was only partially implemented; students completed two unit-based assessments but not the third, final assessment window. Similarly, the state cancelled its annual Teacher Leader Summit and all entailed professional development sessions (LDOE, 2020). September 2020 the LDOE decided to suspend IADA implementation for the 2020-2021 school year, when it became apparent that in-person schooling and guaranteed state testing implementations were both unreliable assumptions given the ongoing pandemic (C. Johnson, personal communication, March 10, 2022).

2021-2022 IADA Administration

In the 2021-2022 school year, eight Louisiana school districts—Central Community, Claiborne Parish, Evangeline Parish, Grant Parish, Lincoln Parish, Monroe City, Ouachita Parish, and Rapides Parish—and Redesign Schools used the IADA assessment for all Grade 7 students in lieu of the traditional LEAP ELA test (C. Johnson, personal communication, October 21, 2021). Assumption Parish and St. John the Baptist Parish did not participate in the second year of Grade 7 IADA testing. Unlike in 2019-2020, the state also offered no program of professional development specific to the IADA test offered for teachers (C. Johnson, personal communication, December 14, 2021). Findings potentially related to this difference in implementation will be discussed in chapter 5.

As noted previously, state testing was suspended in the 2019-2020 school year. In the 2020-2021 year, BESE required that conventional state tests be administered but directed LDOE
to calculate only simulated school ratings that had no consequences in policy (Sentell, 2021, December 3).\textsuperscript{27} The latest, active, consequential school rating thus remained the rating awarded in 2019, prior to the pandemic. In September 2022, BESE voted to extend its past policy for IADA-participating districts, meaning that 2022 school and district ratings would be based in part on the higher of either mean ELA test results from 2019 or mean 2022 IADA results.

Another change in the IADA implementation between 2019-2020 and 2021-2022 was an expansion of the program to other grade levels and one additional curriculum in the second year (C. Johnson, personal communication, October 21, 2021). While this study measures only Grade 7 implementation and effects, school districts not participating in the Grade 7 IADA test may have participated in similar pilot assessments LDOE is conducting in Grades 4, 5, 6, and 8 (C. Johnson, personal communication, October 21, 2021). This additional layer of complexity in the 2021-2022 implementation requires considerations, discussed later in this chapter, of whether to include teachers in these piloting districts in the 2021-2022 study.

\textbf{Purpose of the Study}

Chapter 1 and chapter 2 of this dissertation established a broad failure of how standards-based policy reforms were designed or implemented to achieve pedagogical improvements in teaching reading comprehension. While standards and tests have vast influence over teachers’ principles and methods, that influence frequently manifests as teaching discrete reading skills, \textsuperscript{27}

\textsuperscript{27} As noted in the cited article, 2021 test results indicated broad declines in Louisiana student skills, including in reading skills. Cited sources in the article attributed the declines to limited in-person schooling. This study does not seek to validate that claim, but it is worth noting complications associated with evaluating teacher principles and methods in a period so disruptive to schooling as we know it. While those disruptions would appear to make the problem of adolescent reading skills all the more profound, factors related to mindsets and behaviors of teachers in this time would seem all the more complex.
rather than on a method that emphasizes the meaning of texts, discerned via knowledge of words, textual structure, and background knowledge.

Chapter 3 established the challenges to large-scale reading comprehension interventions. A number of interventions in which teachers were trained to implement specific pedagogical approaches had positive and significant effects (Palinscar & Brown, 1994; Reznitskaya et al., 2007; Reznitskaya et al., 2012), but most were conducted at limited scale. Pre-service preparation programs, such as colleges of education or alternative route certification models, demonstrate widely varied effects (Cowan & Goldhaber, 2016; Levine, 2006). While some primary grade reading curricula showed similarly promising effects (Slavin et al., 2009), middle school reading curricula and interventions on their own have shown limited evidence of positive and significant effects (Blazar et al., 2019; Slavin et al., 2009). Even repeated studies of nationwide testing and accountability regimes under NCLB concluded that the requirements had positive and significant effects in mathematics and no significant effects in reading (Dee & Jacob, 2011; Lee & Reeves, 2012).

Chapter 3 did identify, however, evidence that specific tests themselves may have significant, positive effects on student skills and attainment. Conley et al. (2014) find that the IB program, in which assessments are aligned with a cross-disciplinary curriculum, has significant and positive effects on participants’ persistence and course grades in post-secondary education. Conley et al. (2014) attribute this to the imperative for IB participants to integrate knowledge of varying subjects and texts into written products, a frequent stumbling block for undergraduate students. The implication of this for U.S. reading assessment may be profound: a state testing series that calls on students to demonstrate and integrate knowledge learned from specific texts,
rather than to demonstrate generic skills in response to random texts, may have lasting effects on their intellectual capacities.

Given the heavy influence a state test’s content has over the methods of teachers, given the failure of both NCLB and CCSS to translate standards and tests into large scale shifts in the teaching of reading comprehension, and given the demonstrated effects of IB on long-term student skills and attainment, there would seem to be value in studying the implementation and effects of a text- or curriculum-specific state test in the U.S. Given in particular the system-wide scale of state testing systems, a new test design that may engender evidence-based teaching approaches could offer policymakers a new tool for addressing the vexing challenge of reading comprehension in the U.S. Informing such policymaking considerations is an important purpose of this study.

The Louisiana IADA implementation is only in a nascent phase, and there is scant literature on the effects of wholesale changes to state testing series. Thus, this study reflects a potentially important starting point for researching the implementation and effects of this novel shift in public policy. This study does not establish causality, and it does not provide a summative evaluation of the intervention. Instead it makes evident promising constructs and questions for further study as the initiative matures.

Research Questions

The objective of this study is to establish the impact of a state standardized test featuring pre-disclosed texts and thematic bodies of knowledge on Grade 7 ELA teachers’ principles and methods. The study establishes the fidelity of the intervention’s implementation in two different school years, acknowledging pandemic-related disruptions, and will compare the principles and methods of teachers who use ELA Guidebooks 2.0 curricula and whose students are taking the
IADA assessment with those of teachers who use ELA Guidebooks 2.0 but do not have students taking the IADA assessment. Evidence that the IADA test contributes to teachers adopting principles or methods associated with an integrated approach to teaching reading may inform policymakers seeking ways to improve reading outcomes among U.S. middle school and high school students.

**Process Evaluation Questions**

One process evaluation question relates to the reach (Baranowski & Stables, 2000) of the assessment to the population of eligible students whose teachers participate in the intervention and study:

- **RQ1:** Are Grade 7 students in IADA-participating school districts consistently completing all available administrations of the assessment?

A second question relates to a contextual factor (Baranowski & Stables, 2000), technology availability, necessary to understanding the adequacy of educator and student experience across the sample:

- **RQ2:** Do schools with participating teachers and students demonstrate adequate levels of technology readiness for implementing the IADA assessment?

These questions create a framework for determining whether the assessment has been delivered as designed, whether there are inconsistencies in implementation, and whether certain contextual factors are associated with inconsistencies in implementation.

**Outcomes Evaluation Questions**

The outcomes measured in this study reflect constructs related to the integrated model for teaching reading outlined in chapter 3: text usage, purpose in teaching texts, classroom activities
CURRICULUM-SPECIFIC READING TEST

when using texts, and influences over activities teachers choose. The four outcomes-related research questions are listed below.

- **RQ3**: Do Grade 7 ELA students taking a state reading test focused on pre-determined texts use those texts in class with greater frequency than do peer students?
- **RQ4**: To what extent are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more likely than peer teachers to design lessons to teach the meaning of those texts rather than discrete reading skills?
- **RQ5**: Are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more or less likely than peer teachers to have students discuss or write about the significance of textual features and structure; access texts of appropriate grade-level complexity; or discuss or write about how texts relate to other texts or other subject matters?
- **RQ6**: What factors do Grade 7 ELA teachers identify as influential in how they use texts and plan activities in the classroom?

The research questions varyingly integrated model constructs through the lens of a teacher’s actual classroom methods versus the principles that guide the teacher’s decisions. Table 4.2 presents the constructs each research question measures, noting whether the question seeks data related to the teacher’s method or the teacher’s principles.

Table 4.2

*Outcomes Constructs and Related Research Questions*
### Outcomes Evaluation Question (abbreviated)

| RQ3: Do treatment students use texts in class more frequently than other students? | Text usage | Method |
| RQ4: Do treatment teachers intend to teach a text’s meaning over skills, to a greater degree than other teachers? | Purpose of teaching text | Principle |
| RQ5: Are treated teachers more likely than others to use integrated activities? | In-class reading activities | Method |

### Research Design

Evaluations have three essential “branches” that depict their respective purposes: methods, which evaluates how an intervention was designed and conducted; valuing, which evaluates an intervention’s worth; and use, which evaluates the intervention’s utility for specific audiences (Alkin & Christie, 2013). This study is grounded in the use tradition of evaluation. Its goal is to determine the nature and importance of a policy experiment to future U.S. education policy, as well as to shape and inform future lines of study. This section of the chapter discusses how the study is designed so that it can achieve this purpose.

### Logic Model

The key independent variable for this study will be the teachers experiencing the ongoing administration of the assessment itself. Dependent variables are indicators of teacher principles and methods. The study’s logic model (Figure 4.1) outlines the complex assembly of resources...
and activities necessary to execute the theory of treatment (Wholey, Hatry, & Newcomer, 2010, p. 55). Implementing the IADA assessment model necessitated investment of state, district, commercial, and philanthropically donated resources. Context for the treatment thus includes an assumption of ongoing financial support for the intervention. Context also involves relevant external factors, such as federal and BESE policy related to assessment and accountability systems, which to date permit IADA implementation.

Figure 4.1. Logic model

**Inputs**
- Staff: state department and school district coordinators
- Vendors: publisher, designer, technical advisor, academic advisor
- Funding: philanthropically raised funds to support design, development, and administration; public funding to support training and implementation
- Time: design process and test administration throughout the year
- Hardware: devices to administer assessments

**Activities**
- Develop IADA forms
- Implement IADA forms
- Train teachers and district officials on assessment form and curriculum support
- Teach curricular units associated with assessments (3x annually, roughly 1.5 months each)
- Administer assessment forms (3x annually, roughly two hours each)
- Report results of IADA test three times annually to school systems and families.

**Immediate Outcome**
- Treated teachers are more likely than peers using same curriculum to report they use grade-level texts frequently in class; teach texts to assist students making meaning of texts; and teach in the naturalistic method, with its balances among strategies, close reading, and knowledge outside the text.

**Participants**
- Grade 7 ELA teachers of students in Louisiana school systems that use ELA Guidebooks 2.0 curriculum an administer IADA test

**Assumptions**
- District participation equals willing teacher participation
- Assessment development team delivers product
- Funding commitments persist

**External Factors**
- Changes in state, district, or school leadership
- Pandemic-related interruptions to implementation and policy
- Technology capacity
- USDOE IADA policy
- State accountability policy

**Long-term Goal**
- Improved reading comprehension among adolescent readers

**Design**

This is a mixed-methods study involving correlational analysis of quantitative data and exploratory analysis of qualitative data (Mertens, 2018). The study uses reports generated by the
LDOE to evaluate student participation and technology readiness fidelity in IADA implementation in two separate school years, 2019-2020 and 2021-2022. In determining outcomes, quantitative results are dominant. The study uses data gathered from a survey administered by the LDOE in 2019-2020 and 2021-2022, to evaluate Grade 7 ELA teacher principles and methods. In order to address limitations to validity discussed in chapter 5, the study two separate years of significant quantitative effects of a given construct in order to establish a correlation. The study then incorporates interviews of teachers whose students participate in IADA, in order to further validate quantitative findings and in order to explore mediating factors associated with quantitative outcomes.

While the study uses data from two school years, this is not a longitudinal study meant to show changes in the IADA effects over time. As discussed later in this chapter, the school district affiliations of non-treated survey respondents differ too greatly from 2019-2020 to 2021-2022 to allow for a stable control group over time. I thus limit tests of differences between the two years to differences between the treated groups in each year, indicating potential shifts in IADA conditions or implementation but not changes in IADA effects over time.

**Procedure**

This section of the chapter describes activities related to conducting the study. The section covers recruitment, sampling, and measures, along with processes for conducting the intervention, for collecting data, and for analyzing data.

**Participants**

The population for this study is Louisiana Grade 7 ELA teachers who use ELA Guidebooks 2.0 curricula in their classrooms. The size of this population annually is
CURRICULUM-SPECIFIC READING TEST

approximately 461.\textsuperscript{28} Prior research (Doan et al., 2022; Kaufman et al., 2016) indicates that surveyed Louisiana ELA teachers across grade levels were more likely than ELA teachers in other states to use instructional materials aligned to CCSS. Kaufman et al. (2016) also find that Louisiana teachers are more likely than ELA teachers in other states to affirm teaching principles embodied by CCSS, such as prioritizing access to grade level texts for all readers and teaching the meaning of texts rather than discrete reading skills. This study focuses on a narrow segment of that larger population: teachers in one grade level who teach using one highly-rated, CCSS-aligned curriculum. In focusing on this distinct population, the study seeks to discern unique correlations with the IADA test beyond principles and methods already adopted in the general Louisiana ELA teacher population.

December 2019, LDOE staff contacted ELA coordinators in the ten districts and one charter management organization participating in the IADA assessment, as well as coordinators in six school districts with at least one full year of Guidebooks 2.0 implementation experience (C. Johnson, personal communication, October 21, 2021).\textsuperscript{29} The LDOE requested that Grade 7 ELA teachers in these districts complete online forms of a 20-minute survey (Appendix B) as part of district- or school-based ELA trainings. In January 2021, LDOE staff distributed the same survey form by contacting district test coordinators in eight IADA-participating districts.

\textsuperscript{28} The LDOE does not publish a list of districts using ELA Curriculum Guidebooks. To estimate the size of the Grade 7 teacher population using Guidebooks, I assumed that the 40 districts with teachers who responded to the survey used in this study represented the set of districts using the curriculum. Based on those districts’ Grade 7 total enrollments, an estimate of Louisiana Grade 7 students using the curriculum annually is 33,000. Responding teachers reported that they teach an average of 2.6 sections of Grade 7 ELA. Applying an average class size of 25 Grade 7 students per class, there are 65 total Grade 7 students per teacher in the population and thus roughly 507 total teachers using Grade 7 ELA Guidebooks 2.0 in Louisiana.

\textsuperscript{29} The basis for selecting these six districts, beyond their experience in using the curriculum, was not discernible.
and one charter management organization. Rather than sending the survey to a small set of other districts using Guidebooks 2.0 as ELA Grade 7 curriculum (R. Cailland, personal communication, February 8, 2022), the LDOE sent the survey to all districts using ELA Guidebooks 2.0. The LDOE again requested that Grade 7 ELA teachers complete the form.

In order to include as many responses as possible within a narrow segment of the teacher population, the study uses a voluntary sampling framework, using the responses of teachers who completed the LDOE survey form. Teachers in the 2019-2020 and 2021-2022 treatment groups are all Grade 7 ELA teachers who responded to the survey in those respective years and whose district or charter organization is participating in each year’s Louisiana IADA tests. Non-treated teachers are respondents to each year’s survey who reported using ELA Guidebooks 2.0 but whose employing school district is not participating in the IADA test. Later in this chapter I discuss the group traits and establish where differences may create limitation to findings.

Survey response rates were high among treated teachers in 2019-2020 (77%) and lower among non-treated teachers (53%).\textsuperscript{30} Responses rates among treated teachers in 2021-2022 were lower than in the prior year (63%), as were response rates among non-treated teachers (39%). Lower responses rates in the 2021-2022 non-treated sample may be related to changes in the surveyed population, which the LDOE expanded from 6 districts to 31 districts, as discussed earlier in this chapter.

Before conducting analysis, I cleaned the survey data, excluding incomplete responses or responses from individuals who are not members of the population. In total I excluded 5

\textsuperscript{30} To calculate responses rates, I created an estimated number of Grade 7 ELA teachers in each responding district based on student population. See detailed assumptions involved in these estimates in the prior footnote.
responses to the 2019-2020 survey and 38 responses to the 2021-2022 survey. The resulting 2019-2020 sample includes 70 treated participants and 69 non-treated participants. The resulting 2021-2022 sample includes 49 treated participants and 132 non-treated participants. The difference in treatment group sizes between the two years may be due to fewer districts participating in Grade 7 IADA in 2021-2022. The difference in non-treatment group sizes between the two years may be due to the LDOE surveying a broader set of non-IADA districts in 2021-2022 than in 2019-2020.

To analyze the composition of each group, the survey requested that respondents report their years of teaching experience at the time of the survey, identifying as having fewer than three years of experience, between three and ten years of experience, or greater than ten years of experience. Table 4.3 shows that the 2019-2020 groups are generally similar in years of experience.

---

31 Because I required knowledge of a respondent’s affiliated school system to verify whether the respondent was participating in the IADA treatment, I first excluded any respondent in 2019-2020 (5) and 2021-2022 (7) who listed neither an affiliated school nor an affiliated district in completing the form. I next excluded one respondent to the 2021-2022 survey who indicated that she is not a teacher and one additional response from 2021-2022 whose responses were identical to another record, indicating that the response was likely a duplicate. I further excluded all 2021-2022 respondents (17) indicating on the survey that in a given week they never or almost never use ELA Guidebooks 2.0 in their classrooms. Of these 17 respondents, 15 taught in non-IADA districts. Survey responses provided no indication as to why these respondents did not use the curriculum routinely, even as the LDOE identified their districts as having adopted the curriculum and use it in other schools. That the excluded teachers are clustered in a small set of schools indicates that the decision not to use ELA Guidebooks 2.0 may have been made at the school level, rather than at the district level, and that a sample valuing consistency in curriculum usage could not reasonably include their responses. Last, I removed 12 respondents from the 2021-2022 sample either because their districts ceased participation in the Grade 7 IADA test or because their districts joined related 2021-2022 state pilots in other grade levels. These 2021-2022 teachers may have experienced indiscernible effects of IADA from participation in the past or from being a participant in new, related pilots, but their 2021-2022 experiences were not comparable to those of either treated participants in the fully operational 2021-2022 IADA test or non-treated participants with no IADA experience.
experience, while the 2021-2022 treatment group is more experienced than the non-treated group. To test the significance of these differences, I coded respondents’ categories on a scale of 0-2 for both years and conducted Mann-Whitney test of differences between means. There was no significant difference in 2019-2020 between the two groups ($p = 1.000$). The groups in 2021-2022 differ significantly in experience ($p = .006$). This may be partially a function of treated teachers responding to the same survey twice, two years apart; some respondents may be more experienced in 2021-2022 than when they first completed the survey in 2019-2020. Whatever the underlying factors, the difference in experience levels in 2021-2022 requires further examination of the role experience may play in any outcomes attributed to that years’ responses.

Table 4.3

*Years of Teaching Experience, 2019-2020 and 2021-2022*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 years of experience</td>
<td>13</td>
<td>13</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>%</td>
<td>19</td>
<td>19</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>&gt;3 years of experience, &lt;10 years of experience</td>
<td>24</td>
<td>20</td>
<td>11</td>
<td>48</td>
</tr>
<tr>
<td>%</td>
<td>34</td>
<td>29</td>
<td>22</td>
<td>36</td>
</tr>
<tr>
<td>&gt;10 years of experience</td>
<td>33</td>
<td>36</td>
<td>33</td>
<td>58</td>
</tr>
<tr>
<td>%</td>
<td>47</td>
<td>52</td>
<td>67</td>
<td>44</td>
</tr>
<tr>
<td>N</td>
<td>70</td>
<td>69</td>
<td>49</td>
<td>132</td>
</tr>
</tbody>
</table>

To understand the socioeconomic status of students in schools where participants teach, I imported the percentage of students in the schools of each respondent receiving free and reduced

115
lunch (FRL) subsidies from the LDOE Mutilstats reports for 2019-2020 (LDOE, 2020c) and 2021-2022 (LDOE, 2022). Mean school-level FRL for treated teachers in 2019-2020 was 71.32%; for non-treated teachers mean FRL was 78.26%. An independent samples t-test indicated a significant difference $t(137) = 2.425$, $p = .017$, between the two samples. Mean school-level FRL for treated teachers in 2021-2022 was 64.70%; for non-treated teachers mean FRL was 68.27%. Differences in 2021-2022 were not significant, $t(179) = 1.186$, $p = .237$.

Notably, the rate of FRL associated with both years’ samples exceeded the most recently reported U.S. average FRL (52.3%; NCES, n.d.). FRL rates for both samples also exceed the highest, most recent annual average FRL count in Louisiana (NCES, n.d.), taken in 2017-2018 (63.0%). In spite of differences between samples in 2019-2020, both samples in both years could be fairly characterized as disproportionately comprised of economically disadvantaged students.

As discussed previously, in 2019-2020 the LDOE distributed the survey to all districts participating in IADA and to six districts not participating in IADA that had at least a year of experience using ELA Guidebooks 2.0. In 2021-2022, LDOE again distributed the survey to all districts participating in IADA, but LDOE also distributed the survey to all districts with at least a year of experience using ELA Guidebooks 2.0. Treated participants in 2019-2020 were affiliated with ten school districts and one charter management organization; treated participants in 2021-2022 were affiliated with eight school districts and one charter management organization. Six school districts and one charter management organization had affiliated teachers in both years, and in both years teachers affiliated with the Ouachita Parish district, the largest of the group by students served, made up 51% of the group. While there was a shift in

---

32 LDOE did not disclose the basis for selecting non-IADA districts in either year.
the overall size of each group between 2019-2020 ($n=70$) and 2021-2022 ($n=49$), the set of districts represented in the group is relatively stable, and the large percentage of the group attributable to one large school system remains consistent.

Table 4.4

*District Affiliation of IADA-treated Groups, 2019-2020 and 2021-2022*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption</td>
<td>4</td>
<td>5.7</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Central</td>
<td>1</td>
<td>1.4</td>
<td>2</td>
<td>4.1</td>
</tr>
<tr>
<td>Claiborne</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4.1</td>
</tr>
<tr>
<td>Evangeline</td>
<td>10</td>
<td>14.3</td>
<td>4</td>
<td>8.2</td>
</tr>
<tr>
<td>Grant</td>
<td>2</td>
<td>2.9</td>
<td>4</td>
<td>8.2</td>
</tr>
<tr>
<td>Lincoln</td>
<td>5</td>
<td>7.1</td>
<td>2</td>
<td>4.1</td>
</tr>
<tr>
<td>Monroe</td>
<td>1</td>
<td>1.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ouachita</td>
<td>36</td>
<td>51.4</td>
<td>25</td>
<td>51</td>
</tr>
<tr>
<td>Rapides</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>18.4</td>
</tr>
<tr>
<td>Redesign</td>
<td>1</td>
<td>1.4</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>St. John</td>
<td>10</td>
<td>14.3</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td><strong>70</strong></td>
<td></td>
<td><strong>49</strong></td>
<td></td>
</tr>
</tbody>
</table>

As shown in table 4.6, non-treated group district affiliations show greater levels of variation between 2019-2020 and 2021-2022.

Table 4.5

*District Affiliation of Non-treated Groups, 2019-2020 and 2021-2022*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acadia</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>4.5</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Ascension</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>4.5</td>
</tr>
<tr>
<td>Beauregard</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>Bossier</td>
<td>4</td>
<td>5.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Caddo</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>14.4</td>
</tr>
<tr>
<td>Caldwell</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>Concordia</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>CSUSA</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td>Franklin</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>Iberia</td>
<td>2</td>
<td>2.9</td>
<td>8</td>
<td>6.1</td>
</tr>
<tr>
<td>Jefferson</td>
<td>46</td>
<td>66.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jeff. Davis</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5.3</td>
</tr>
<tr>
<td>Lafayette</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>9.8</td>
</tr>
<tr>
<td>Lafourche</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3.0</td>
</tr>
<tr>
<td>Morehouse</td>
<td>1</td>
<td>.8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Natchitoches</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3.0</td>
</tr>
<tr>
<td>Orleans</td>
<td>2</td>
<td>2.9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pointe Coupee</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>Richland</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3.0</td>
</tr>
<tr>
<td>SSD</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>St. Bernard</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>St. Charles</td>
<td>12</td>
<td>17.4</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>St. Helena</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.8</td>
</tr>
<tr>
<td>St. Landry</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>3.8</td>
</tr>
<tr>
<td>St. Martin</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>6.1</td>
</tr>
<tr>
<td>St. Mary</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>St. Tammany</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>11.4</td>
</tr>
<tr>
<td>Tensas</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>Terrebonne</td>
<td>1</td>
<td>1.4</td>
<td>13</td>
<td>9.8</td>
</tr>
<tr>
<td>Washington</td>
<td>1</td>
<td>1.4</td>
<td>1</td>
<td>.8</td>
</tr>
</tbody>
</table>

While teachers affiliated with eight districts were included in the 2019-2020 non-treated group, a majority (66.7%) were affiliated with one district, Jefferson Parish, the state’s largest parish by population. The 2021-2022 non-treated group included teachers from 25 districts and one charter management organization. Relatedly, the 2021-2022 group (n=132) is nearly twice the size of the 2019-2022 group (n=69). Only four districts were represented in both years’ non-treated groups.

While the 2019-2020 and 2021-2022 treated groups overlap a great deal in terms of affiliated districts and in terms of the predominant contribution of one district, Ouachita Parish,
non-treated groups demonstrated minimal overlap in district affiliations. The non-treated groups are separate groups of teachers, teaching under separate and distinct district leadership. While there are no data to verify the implications of these districts for the approaches teachers adopt, the difference in groups and the sheer variety in district affiliations within each year’s group, indicates that the study should not consider any question that requires longitudinal analysis of a non-treated group. For this reason, while the study will consider how and why treated teacher responses may change over time, the study will not consider the extent to which differences between treated and non-treated teachers change over time. Analyses examining differences between treated and non-treated groups will consider 2019-2020 and 2021-2022 separate and distinct years researching non-treated teachers who are very likely teaching under different and highly variable conditions.

Measures and Instrumentation

This section details measures related to process evaluation and outcomes evaluation. It will also outline instruments that the study uses to collect data, the indicators used for subsequent data analysis, and the tests conducted to yield findings.

Process Evaluation

Process evaluation indicators and data collection processes correspond with research questions 1 and 2 (see Table 4.7 for data collection matrix). The primary implementation activity in the logic model is administration and completion of the IADA test forms. A first indicator is the percentage of eligible students in treatment classrooms who complete three forms of the test. Because the test is designed to have a cumulative score based on completion of three forms, each eligible student should complete all three forms. Verifying this level of fidelity can be done via a review of an annually issued participation report submitted by the LDOE to the USDOE. The
USDOE requires that 95% of an eligible school or district population complete a state test (Dreillinger, 2016). This standard of participation allows an evaluation of the project’s implementation in given school districts, as districts with lower rates of test completion will not have fully implemented the intervention.\(^{33}\)

While the IADA test form is designed to be identical in content and in administrative conditions and procedures for all students, it also may be true that there are variations in the quality of test administration from site to site. Data related to irregularities in testing administration across the treatment group can be collected through the annual LDOE Test Security Irregularities Report presented to BESE on an annual basis (Dreillinger, 2016). The report notes data collected by on-site, external observers, as well as district staff, to identify instances in which administrators did not follow protocols or in which intentional malfeasance, such as plagiarism or post hoc answer changes, may have occurred (Dreillinger, 2016). While the LDOE issued no irregularities report for the 2019-2020 school year, a standard irregularities report will provide further evidence of implementation fidelity, at a classroom level rather than at a district level. Treated classrooms in which repeated administrative errors occur may reflect that administrative trainings or communications have been ineffective and that students in certain classrooms have had inconsistent testing experiences relative to other students of treatment.

\(^{33}\) As will be discussed, these are standard procedures entailed in state accountability systems. The suspension of these systems in 2020 due to the COVID-19 pandemic inevitably disrupted not only the implementation of state tests but also state data collection procedures, even where tests were partially completed. The pandemic presented challenges not only to actual implementation but also to the collection of implementation data, a limitation on evaluating the fidelity of 2019-2020 partial implementation. This limitation will be discussed in the discussion of implementation outcomes in chapter 5,
teachers. While these data will not provide evidence of district-wide implementation fidelity, they will be shared in chapter 5 to highlight potential inconsistencies.

Finally, because the IADA test form is provided digitally, participating schools must demonstrate adequate technology capacity in order to implement the intervention fully. The readiness of a school to implement online testing can be measured against both statewide average digital device access and a statewide goal that every school will possess at least one device for every student, established by the LDOE in 2018 (LDOE, 2020a). Prior to the pandemic, the LDOE issued annual Technology Footprint reports reporting device availability and Internet bandwidth that will be used to evaluate the technology context against these standards (LDOE, 2020b). In 2021-2022, the LDOE shifted its reporting format, making these data available through a web interface called the School Reopening Dashboard (LDOE, 2022a). Schools hosting an administration of the IADA form that are in parishes not meeting the state average device access or are not making progress toward statewide device access goals create a context for the intervention that may affect implementation or the ability of schools to sustain and grow IADA in future years.

Table 4.6

*Process Evaluation Data Collection Matrix*

<table>
<thead>
<tr>
<th>Process Evaluation Question (abbreviated)</th>
<th>Indicator</th>
<th>Data Source</th>
<th>Collection Tool</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are all eligible students completing all administrations?</td>
<td>% Eligible students completing test administration</td>
<td>LDOE</td>
<td>Pre-existing LDOE Participation Report</td>
<td>Once annually</td>
</tr>
</tbody>
</table>

34 Irregularity reports, as also discussed later, are components within accountability procedures that were significantly disrupted in 2020, further limiting implementation evaluation for that year.
### Outcomes Evaluation

Quantitative data, gathered via a survey form (see Appendix B) administered by the LDOE once in 2019 and once in 2021, are dominant in this study, generating data related to teacher principles and methods. The LDOE survey form was comprised of seven multi-select items addressing constructs reflected in the research questions. Four items on the survey were modified items originally developed for the American Teacher Panel (Opfer et al, 2016). Three items were original to the survey form.

Table 4.7

*Survey Items and Origins*

<table>
<thead>
<tr>
<th>Survey Question (abbreviated)</th>
<th>Response Options</th>
<th>Item Origin</th>
<th>Construct Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: How closely does your instruction follow Guidebooks units and lessons?</td>
<td>Multi-select: never, occasionally, often, almost always</td>
<td>Original to this survey</td>
<td>Text usage</td>
</tr>
<tr>
<td>Q2: For what percentage of a typical lesson do students use Guidebook texts?</td>
<td>Multi-select: never, occasionally, often, almost always</td>
<td>Original to this survey</td>
<td>Text usage</td>
</tr>
<tr>
<td>Q3: Which better describes your approach: teaching texts or teaching skills?</td>
<td>Binary: one of two statements</td>
<td>American Teacher Panel</td>
<td>Purpose of teaching text</td>
</tr>
<tr>
<td>Q4: Please select statements that accurately describe your view of what it means</td>
<td>Choose any number of five descriptive</td>
<td>Original to this survey</td>
<td>Purpose of teaching text</td>
</tr>
</tbody>
</table>
Qualitative data gathered via interviews of treatment teachers validated and expanded on quantitative findings related to the same constructs measured by the survey. I developed an interview protocol (Appendix C) to address constructs embodied by the study’s outcomes-focused research questions. The protocol contains six questions meant to mimic the substance and ordering of the seven substantive items on the LDOE survey. I combined the content of survey items 3 and 4 into one interview question; otherwise, each interview item is a recasting of the survey item, seeking descriptive responses while the survey sought multi-select responses.

In December and January of the 2021-2022 school year, administrators in three school systems that participated in IADA, Lincoln Parish, City of Monroe, and Redesign Schools, presented a written summary of the study to Grade 7 ELA teachers in affiliated schools; twelve volunteers then signed consent forms (Appendix D) and participated in 30-minute interviews.

<table>
<thead>
<tr>
<th>Survey Question (abbreviated)</th>
<th>Response Options</th>
<th>Item Origin</th>
<th>Construct Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>for a student to understand the meaning of a text.</td>
<td>statements, or choose none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q5: What percentage of time do students in your class read grade-level texts versus books on their reading levels?</td>
<td>Multi-select continuous ranges: 0-25%; 25-50%; 50-75%; 75-100%</td>
<td>American Teacher Panel</td>
<td>In-class reading activities</td>
</tr>
<tr>
<td>Q6: How often do students engage in the following activities in your class?</td>
<td>Multi-select: never, occasionally, often, daily</td>
<td>American Teacher Panel</td>
<td>In-class reading activities</td>
</tr>
<tr>
<td>Q7: To what extent do the following factors influence the texts and activities you choose for lessons?</td>
<td>Multi-select: not at all, to a slight extent, somewhat, a great deal</td>
<td>American Teacher Panel</td>
<td>Influence over teaching decisions</td>
</tr>
</tbody>
</table>

In December and January of the 2021-2022 school year, administrators in three school systems that participated in IADA, Lincoln Parish, City of Monroe, and Redesign Schools, presented a written summary of the study to Grade 7 ELA teachers in affiliated schools; twelve volunteers then signed consent forms (Appendix D) and participated in 30-minute interviews.
CURRICULUM-SPECIFIC READING TEST

via Zoom online meetings. Transcripts derived from recorded interviews via Trint software yielded qualitative data for triangulating quantitative findings and identifying potential mediating factors that may explain relationships among the intervention and quantitatively identified changes in teacher belief or behavior. Table 4.8 summarizes all outcomes data collection in the study.

Table 4.8

*Outcomes Data Collection Matrix*

<table>
<thead>
<tr>
<th>Outcomes Evaluation Question (abbreviated)</th>
<th>Construct</th>
<th>Data Source</th>
<th>Collection Tool</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do Grade 7 ELA students taking a state reading test focused on pre-determined texts use those texts in class with greater frequency than do peer students?</td>
<td>Text usage (method)</td>
<td>Teacher</td>
<td>Multi-select survey items 1 and 2; interview items 1 and 2.</td>
<td>Once annually</td>
</tr>
<tr>
<td>To what extent are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more likely to design lessons to teach the meaning of those texts rather than discrete reading skills?</td>
<td>Purpose in using texts (principle)</td>
<td>Teacher</td>
<td>Multi-select survey items 3 and 4; interview item 3</td>
<td>Once annually</td>
</tr>
<tr>
<td>Are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more or less likely than peer teachers to have students discuss or write about the significance of textual features; access texts of appropriate grade-level complexity; and discuss or write</td>
<td>Activities when using texts (method)</td>
<td>Teacher</td>
<td>Multi-select survey items 5 and 6; interview item 6</td>
<td>Once annually</td>
</tr>
</tbody>
</table>
The influence factors listed in survey item 7 include a list of programs, events, and brands, the definitions of which are not self-evident to readers not steeped in Louisiana education policy and customs. Table 4.9 lists and defines each factor included in survey item 7, as a reference for readers, particularly as they examine research question 6.

Table 4.9

*Survey Item 7 Influence Factors*

<table>
<thead>
<tr>
<th>Influence Factor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Louisiana Student Standards</td>
<td>Louisiana academic benchmarks discussed in chapter 1</td>
</tr>
<tr>
<td>Louisianabelieves.com</td>
<td>LDOE web site</td>
</tr>
<tr>
<td>ELA Guidebooks 2.0</td>
<td>Free curriculum used by all members of the sample</td>
</tr>
<tr>
<td>District or school requirements</td>
<td>Rules and procedures mandated by local school systems</td>
</tr>
<tr>
<td>End of year LEAP 2025 and LEAP Humanities</td>
<td>Brand names for end-of-year state test and IADA test</td>
</tr>
<tr>
<td>LEAP 360 assessments</td>
<td>Brand name for optional, interim test provided by state</td>
</tr>
<tr>
<td>District-wide or school-wide tests</td>
<td>Assessments required by school; system rather than state</td>
</tr>
</tbody>
</table>
CURRICULUM-SPECIFIC READING TEST

<table>
<thead>
<tr>
<th>Compass, TAP, CLASS or other teacher observation rubrics</th>
<th>Required performance management tool for teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-service preparation</td>
<td>Preparatory training prior to teaching career</td>
</tr>
<tr>
<td>Teacher Leader Summit</td>
<td>Annual statewide professional development event</td>
</tr>
<tr>
<td>Professional development experience in the last year</td>
<td>Training held throughout the course of the school year</td>
</tr>
<tr>
<td>Online teacher networks</td>
<td>Web sites offering resources for educators</td>
</tr>
</tbody>
</table>

**Threats to Validity**

The validating and exploratory function of the qualitative data gathering is particularly important given the predominance of survey data in this study. Self-reporting on surveys is a risk to construct validity (Shadish, Cook, & Campbell, 2002), but the mixed-methods design of the study ensures that there is not an exclusive reliance on one method of data collection (Creswell & Plano-Clark, 2018). Contamination within the sample is a further threat to construct validity (Shadish et al., 2002), given that treated teacher respondents may communicate with other teachers about the survey, including with control group members. The methodology addresses this by ensuring that non-treated participants do not teach in the same districts as treatment group teachers or teachers participating in a different grade level pilot of a curriculum-specific test. As discussed earlier in this chapter, inclusion in the treatment group required full and complete district participation in the Grade 7 operational test, across all schools, while participation in the non-treatment group required that there be no district participation in IADA or any related pilot. There is a final risk that treated teachers participating in the survey may have experienced more substantial or sophisticated district- or
school-based training related to curriculum implementation than have teachers in other districts using the same curriculum, and that a district’s choice to participate in IADA may be associated with the district’s academic sophistication. If true, this may be a source of selection bias in the treated group and thus a risk to the study’s internal validity. To minimize this risk, the non-treated group in both years is drawn entirely from districts with at least one year of full experience using ELA Guidebooks 2.0, minimizing the likelihood that greater familiarity with curricula creates advantageous conditions for treated participants over control participants.

**Data Analysis**

This section describes how I analyzed data in order to answer each research question. I gathered quantitative evidence from both 2019-2020 and 2021-2022 school years. I refer to findings associated with evidence from one school year as “promising”; I refer to findings associated with evidence from two school years as “compelling.” Qualitative evidence provides further validation and evidence of mediating factors that may explain results. In discussing findings later in the study, I emphasize compelling quantitative evidence validated by qualitative evidence.

**Process Evaluation**

To analyze student participation rates in the IADA assessment implementation (RQ1), I examined data indicating the percentage of Grade 7 students in each district or charter organization completing each seasonal administration of the test, made available by the LDOE for 2019-2020 (Z. Zhong, personal communication, April 13, 2022) and 2021-2022 (LDOE, 2022c). To determine deviations from generally held standards for fidelity and reach, I evaluated the extent to which districts or charter organizations exceeded or fell short of the 95% threshold for districtwide participation established by USDOE. Districts or charter organizations not
meeting this standard are worthy of further inquiry to determine the cause of failures to achieve adequate fidelity and reach, including examination of annual irregularity reports discussed earlier in this chapter and examination of technology access addressed by RQ2.

To determine adequacy of technology conditions in treated schools (RQ2), I drew on LDOE reports discussed earlier in this chapter (LDOE, 2020b; LDOE 2022a) in order to construct student-to-device ratios for IADA-participating schools. I then compared these ratios in each year to the LDOE statewide goal for student-to-device ratio (LDOE, 2020a) and the statewide average student-to-device ratio for each year.

To answer RQ2, I determined the extent to which IADA-participating schools in both years met or exceeded the statewide average and the state goal. I also calculated changes over time in participating schools’ device inventories and student-to-device ratios. I identified any school not meeting either statewide average or statewide goal standards for student-to-device ratios as potentially at risk for implementation fidelity.

**Outcomes Evaluation**

LDOE administered surveys and provided me with de-identified survey data. I then added school-level information from publicly available government sources to each respondent record in order to analyze the sample. This additional information included school-level FRL percentages, school federal intervention status, and district IADA participation status.

**RQ3: text usage.** To analyze the fidelity with which teachers use ELA Guidebook 2.0 texts to guide their lesson planning, as well as the frequency with which they use Guidebook texts in class, I coded survey data in the following way: 0 indicates “never”; 1 indicates “occasionally”; 2 indicates “often”; 3 indicates “always or almost always.” Using SPSS, I generated frequencies and descriptive statistics to determine potential differences among treated
and non-treated teachers. Because the survey yielded responses that are ranked and non-continuous, I used a robust non-parametric test, the Mann-Whitney test (MacFarland & Yates, 2016), to determine differences among mean responses. When means differences were present, I calculated effect size using the standardized test statistic $z$ and the sample size $n$ (MacFarland & Yates, 2016).\(^{35}\)

$$r = \frac{z}{\sqrt{n}}$$

I applied Cohen’s (1988) guidance on the magnitude of effect size calculations. I further used the Kolmogorov-Smirnov test to measure the shape of response distributions (Lehmann, 2006). Using both tests allowed me to evaluate differences in ways sensitive both to differences in median response and to response distributions (Lehmann, 2006). I used these tests to compare responses among the treated and non-treated groups in each of the two years; I also used both tests to compare differences between the treated group in 2019-2020 and the treated group in 2021-2022.

**RQ4: intent in teaching texts.** To determine differences in teachers’ intentions when teaching texts, I assigned values of 0 or 1 to each response to item 4 on the survey, which has a binary, nominal scale. I generated frequencies and descriptive statistics. I then used Fisher’s exact test, a non-parametric test of independence appropriate for binary variables (Lehmann, 2006), to determine relationships between IADA participation and intent responses. I used Fisher’s exact test also to measure the differences in relationship between intention and 2019-2020 IADA participation versus 2021-2022 IADA participation.

To determine differences in groups’ definitions of successful reading comprehension, I created a table to analyze the frequency and percentage of respondents supporting each

\(^{35}\) I calculate effect sizes for Mann-Whitney tests using this method throughout the study.
definition. To evaluate differences in definitions between groups, I created a binary scale of 0 for no response or 1 for affirmative responses, for each definition. I then calculated means for each definition and analyzed differences between treatment group and non-treatment group responses for each definition using Fisher’s exact test. I further used Fisher’s exact test to evaluate differences among treatment group response means between 2019-2020 and 2021-2022.

RQ5: classroom activities. To determine differences in the time treated and non-treated allot to teaching grade-level texts and reading-level texts in class, I generated frequency tables for responses to survey item 8. I coded responses to both prompts on an ordinal scale: 0 for “never or hardly ever/0-25%”; 1 for “occasionally/25-50%”; 2 for “often/50-75%”; and 3 for “always or almost always/75-100%”. Using those codes, I generated descriptive statistics. I then conducted Mann-Whitney tests to compare means, and, when means differences were present, I conducted Kolmogorov-Smirnov tests to compare distributions of responses between groups. I used these two tests again to measure differences in treated group responses in 2019-2020 and 2021-2022.

To determine differences in activities conducted while using texts, I assigned values on an ordinal scale for each response to each prompt in survey item 9: 0 for “never or hardly ever”; 1 for “occasionally”; 2 for “often”; and 3 for “almost always.” I developed frequency tables and descriptive statistics for all responses in both years. I then conducted Mann-Whitney tests to compare means, and, when means differences were present, I conducted Kolmogorov-Smirnov tests to compare distributions of responses between groups. I used these two tests again to measure differences in treated group responses in 2019-2020 and 2021-2022.

RQ 6. Influence over teacher plans. To evaluate the influence of factors over the activities chosen by teachers, I assigned a rating on an ordinal scale to each participant’s
response to each option in item 10 of the survey: 0 for “not at all”; 1 for “to a slight extent”; 2 for “somewhat”; and 3 for “a great deal.” I generated frequency tables and descriptive statistics for responses to each source of influence in both years. I then conducted Mann-Whitney tests to compare means, and, when means differences were present, I conducted Kolmogorov-Smirnov tests to compare distributions of responses between groups. I used these two tests again to measure differences in treated group responses in 2019-2020 and 2021-2022.

**Qualitative analysis.** I interviewed 12 volunteering educators via Zoom videoconference. The recordings were transcribed and edited via the transcribing software Trint. I generated a book (Appendix E) of *a priori* codes (Creswell & Plano-Clark, 2018) that mirrored the constructs measured in each survey item. I analyzed each interview transcript by attaching *a priori* codes to evidence in each interview transcript. I then synthesized each coded interview transcript by listing the most frequently supported codes in each interview, along with the specific evidence related to each code. I placed these frequently supported codes and specific evidence for each code into a synthesis matrix of all interviews that showed all evidence for all frequently supported codes. From this synthesis grid, I generated emergent codes (Creswell & Plano-Clark, 2018) that reflect new themes that emerged across interviews. These emergent codes were only identified where there was a predominant theme to respondents’ comments within a given a priori code. As one example, “pacing urgency” emerged as a new, repeated theme when teachers were prompted to discuss the *a priori* construct of curriculum fidelity.

**Conclusion**

Louisiana’s nascent IADA test offers the opportunity to research the effects of a curriculum-specific approach to state ELA tests on teachers’ principles and methods. Using two years of data generated by state-conducted surveys of teachers who use the ELA Guidebook 2.0
curriculum, this study seeks to discern correlations between IADA participation and the evidence-based, integrated model for teaching reading discussed in chapters 1 and 3. The study compares survey responses from Guidebook-using teachers whose students participate in IADA to the responses of Guidebook-using teachers whose students do not participate in IADA to discern differences. The study requires a repeated and significant difference among respondents in two separate years in order to establish a calculation.

The study further validates and explains quantitative results via emergent codes generated through analysis of interviews conducted with teachers and other educators in IADA-participating schools. This qualitative analysis does not yield findings on its own in this study; quantitative data are dominant in this study. But qualitative evidence provides further evidence of findings’ validity; it further helps to explain how mediating factors might contribute to or facilitate findings in the quantitative data.
Chapter 5

Results

This chapter describes findings in response to the study’s six research questions. As discussed in chapter 4, quantitative outcomes are dominant in the study. I find compelling evidence for a given factor only when significant quantitative effects occur in both years of the study. When effects of a given factor occur in one year but not two, I present the effects as promising evidence but not compelling correlation.

For each research question, I present quantitative evidence first. Where qualitative evidence offers validation of a quantitatively derived evidence or indicates a mediating factor associated with quantitative evidence, the qualitative evidence is included in the response to the research question. I present qualitative data related to both compelling quantitative evidence seen over two years and promising quantitative evidence seen only in one year. Qualitative evidence is only meant to validate or offer explanations; it is never presented as evidence of a finding on its own.

I summarize the array of evidence presented in a discussion section at the end of the chapter that includes a chart of all evidence presented. In this discussion section, I identify those findings with compelling evidence as opposed to instances of promising but nor compelling evidence. I further offer potential explanations for these differences, and I synthesize the evidence, drawing connections across the research questions.

The chapter starts with evidence responsive to two implementation research questions:

- RQ1: Are all eligible Grade 7 students in IADA-participating school districts consistently completing all available administrations of the assessment?
• RQ2: Do schools with participating teachers and students demonstrate adequate levels of technology readiness for implementing the IADA assessment?

I next present responses to outcomes questions related to teacher principles and methods:

• RQ3: Do Grade 7 ELA students taking a state reading test focused on pre-determined texts use those texts in class with greater frequency than do peer students?

• RQ4: To what extent are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more likely than peer teachers to design lessons to teach the meaning of those texts rather than discrete reading skills?

• RQ5: Are Grade 7 ELA teachers whose students take a state reading test focused on pre-determined texts more or less likely than peer teachers to have students discuss or write about the significance of textual features; access texts of appropriate grade-level complexity; or discuss or write about how texts relate to other texts or subject matters?

• RQ6: What factors do Grade 7 ELA teachers identify as influential in how they use texts in the classroom?

Implementation

This section answers two questions related to the reach and fidelity of the IADA intervention. The first question relates to the extent to which Grade 7 students enrolled in participating school districts completed the IADA test’s forms within the appointed testing windows. The second question relates to the availability of digital devices in participating
CURRICULUM-SPECIFIC READING TEST

schools, an essential condition for completion of the test. Data analyzed, as discussed below, are from publicly accessible reports issued by the LDOE.

RQ 1: Student Participation in IADA

ESSA, as with its NCLB predecessor, requires states to measure the learning of students in Grades 3 through 8 on standards-based tests (Dreillinger, 2016). In Louisiana, null scores for students who do not participate in the state test, remain included in the calculated performance score the state awards to schools and districts, creating an incentive for schools and districts to encourage participation (Dreillinger, 2016). Waivers issued by the federal government over the course of the COVID-19 pandemic allowed states to relax testing requirements and to not count 2021 testing results in school performance ratings (Ujifusa, 2021), but these waivers impacted neither 2019-2020 nor 2021-2022, the years in which this study collected data. Thus this section of the chapter will consider whether students in IADA-participating school systems completed the IADA test at a rate equal to or above the federal minimum standard of 95%.

Data from 2019-2020 were collected from LDOE reports showing participation rates in ten participating school districts and one charter management organization (Zhong, personal communication, April 13, 2022). Because the third administration in the IADA cycle was cancelled, Table 5.1 shows participation rates for each organization in each cycle. Mean participation in the fall administration was 99.4%; mean participation in the winter administration was 99.1%. In both testing windows, test participation exceeded minimum federal standards for participation.

Table 5.1

<table>
<thead>
<tr>
<th>District</th>
<th>Fall participants</th>
<th>% Grade 7</th>
<th>% Grade 7</th>
</tr>
</thead>
</table>

135
In 2021-2022, LDOE publicly reported the number and percentage of students participating in all state ELA tests, including Grade 7 IADA (LDOE, 2020b). LDOE did not report seasonal administration rates for IADA in 2021-2022. Table 5.2 presents participation totals in IADA-participating districts.

Table 5.2

<table>
<thead>
<tr>
<th>District</th>
<th>Participants</th>
<th>% Grade 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption</td>
<td>250</td>
<td>96.1</td>
</tr>
<tr>
<td>Central</td>
<td>362</td>
<td>100</td>
</tr>
<tr>
<td>Claiborne</td>
<td>122</td>
<td>100</td>
</tr>
<tr>
<td>Evangeline</td>
<td>449</td>
<td>100</td>
</tr>
<tr>
<td>Grant</td>
<td>240</td>
<td>100</td>
</tr>
<tr>
<td>Lincoln</td>
<td>386</td>
<td>100</td>
</tr>
<tr>
<td>Monroe</td>
<td>571</td>
<td>98.8</td>
</tr>
<tr>
<td>Ouachita</td>
<td>1537</td>
<td>99.9</td>
</tr>
<tr>
<td>Rapides</td>
<td>1663</td>
<td>98.6</td>
</tr>
<tr>
<td>Redesign</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>St. John</td>
<td>459</td>
<td>98.5</td>
</tr>
<tr>
<td>Total</td>
<td>6129</td>
<td>99.4</td>
</tr>
</tbody>
</table>

Winter participants

<table>
<thead>
<tr>
<th>Winter participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption</td>
</tr>
<tr>
<td>Central</td>
</tr>
<tr>
<td>Claiborne</td>
</tr>
<tr>
<td>Evangeline</td>
</tr>
<tr>
<td>Grant</td>
</tr>
<tr>
<td>Lincoln</td>
</tr>
<tr>
<td>Monroe</td>
</tr>
<tr>
<td>Ouachita</td>
</tr>
<tr>
<td>Rapides</td>
</tr>
<tr>
<td>Redesign</td>
</tr>
<tr>
<td>St. John</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

99.1
Test participation rates were incrementally higher in 2021-2022 than were reported in 2019-2020. Administrative conditions were different between the two years, however, given that 2021-2022 was not interrupted; districts had time and reason to engage in standard data review and make-up test procedures that were likely absent in 2019-2020. Nevertheless, in both years, participation rates exceeded federal minimum standards.

RQ2: Digital Device Access in IADA Schools

In its Statewide Education Technology Plan (2020a), the LDOE articulated a statewide goal of there being two public school students for every one qualified digital device. Accompanying Technology Footprint Snapshot reports (LDOE, 2020b) on device access in each school district demonstrated that statewide mean device access in schools was 1.1 students for every qualified device, a greater level of access than the statewide goal had established. Table 5.3 shows the extent to which schools in districts participating in 2019-2020 Grade 7 IADA exhibited levels of device access that exceeded statewide mean access and statewide goals for access at the time. The table first lists the number of schools involved in the IADA intervention. Of these schools, it then lists the numbers where there is greater access to devices than the statewide mean level of access and the number of schools in which access is better than the state’s goal; these two columns are not mutually exclusive, as the statewide mean is a more ambitious standard than the statewide goal in 2019-2020. The final column lists the number of schools that do not meet either standard, indicating low levels of device access.

Table 5.3

*Digital Device Access in 2019-2020 IADA-participating Schools*
While there was wide variation in the extent to which participating schools exceed statewide mean access (1.1:1), all but two participating schools (3.7%), Pierre Part Middle School in Assumption Parish and Oak Hill High School in Rapides Parish, provided students greater levels of access that the statewide goal for that school year.

The 2020 Statewide Education Technology Plan (LDOE, 2020a) articulated a goal that by the 2021-2022 school year, schools in Louisiana would provide mean device access of one student for every one qualified device, a greater level of device access than the state had previously set as its benchmark. Updated device access analysis from the LDOE School Readiness Dashboard (2022) indicated that by 2022, statewide device access was .72 students per each device, a greater level of device access than updated state goal required. Table 5.4

---

36 Device access data for City of Monroe and Redesign Schools were not publicly reported in 2019-2020. Results in this table do not include these school systems.

37 N/A indicates that data were not included in LDOE reports.
demonstrates the extent to which schools in districts participating in the 2021-2022 Grade 7 IADA test met or exceeded statewide goal access and statewide mean access.

Table 5.4

*Digital Device Access in 2021-2022 IADA-participating Schools*

<table>
<thead>
<tr>
<th>District</th>
<th>Total IADA schools</th>
<th>Device access in school at or better than state mean ($M = .72:1$)</th>
<th>Device access in school at or better than state goal (1:1)</th>
<th>Device access not at or better than state avg or goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Claiborne</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Evangeline</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Grant</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Lincoln</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Monroe</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Ouachita</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Rapides</td>
<td>13</td>
<td>3</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Redesign</td>
<td>2</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>45</strong></td>
<td><strong>17</strong></td>
<td><strong>35</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

Of 43 participating schools reporting device access, 8 (18.6%) failed to meet statewide goal access of one student per one device. Schools providing sub-standard access were concentrated in Evangeline Parish, where four of six (66.7%) schools failed to provide mean or goal levels of device access. This is a notable result in that half of schools failing to meet state expectations for device access were in one relatively small school system.

Districts participating in IADA consistently increased student access to devices between 2019-2020 and 2021-2022. Table 5.5 shows total devices and student-to-device ratios in all IADA-participating schools during the two school years studied.

Table 5.5
CURRICULUM-SPECIFIC READING TEST

**IADA School Device Access, 2019-2020 and 2021-2022**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption</td>
<td>756</td>
<td>1.65</td>
<td>Left IADA</td>
<td>Left IADA</td>
</tr>
<tr>
<td>Central</td>
<td>1510</td>
<td>.77</td>
<td>1510</td>
<td>.77</td>
</tr>
<tr>
<td>Claiborne</td>
<td>1102</td>
<td>.78</td>
<td>1601</td>
<td>.53</td>
</tr>
<tr>
<td>Evangeline</td>
<td>2648</td>
<td>1.25</td>
<td>3375</td>
<td>.99</td>
</tr>
<tr>
<td>Grant</td>
<td>790</td>
<td>.97</td>
<td>911</td>
<td>.89</td>
</tr>
<tr>
<td>Lincoln</td>
<td>1828</td>
<td>.81</td>
<td>2837</td>
<td>.52</td>
</tr>
<tr>
<td>Monroe</td>
<td>N/A</td>
<td>N/A</td>
<td>1565</td>
<td>1.01</td>
</tr>
<tr>
<td>Ouachita</td>
<td>4434</td>
<td>1.03</td>
<td>7889</td>
<td>.70</td>
</tr>
<tr>
<td>Rapides</td>
<td>7301</td>
<td>1.02</td>
<td>9364</td>
<td>.85</td>
</tr>
<tr>
<td>Redesign</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>St. John</td>
<td>2586</td>
<td>1.22</td>
<td>Left IADA</td>
<td>Left IADA</td>
</tr>
</tbody>
</table>

Of seven districts that participated in IADA both years, and for which the LDOE reported school-level device access in both years, six districts showed increases in device access in IADA schools. This is consistent with a statewide movement of mean access from 1.1 students per device in 2019-2022 to .72 students per device in 2021-2022. Furthermore, in the 2021-2022 school year, all eight IADA-participating school districts met or exceed the state’s 2021-2022 goal for device access levels. This compares with 75% of districts statewide that “met or exceeded” that standard in 2021 (Council for a Better Louisiana, 2021). The majority of IADA-participating school districts increased device access over the two years of this study, and all met statewide device access goals by year two of the study. This suggests that access to devices was not a barrier to IADA implementation fidelity. Adequate levels of digital device access further may have contributed to the high rates of IADA student completion reported earlier in the chapter.
Outcomes

In this section, I present data and analysis in response to four research questions on the principles and methods of Grade 7 ELA teachers. Quantitative data are dominant and were collected via an online survey LDOE administered once annually during 2019-2020 and 2021-2022 (Appendix B). As noted previously, I present results for the two years separately for each research question; I also present differences in treatment group responses between 2019-2020 and 2021-2022; I do not report comparisons among non-treated respondents over time.

I gathered qualitative data via interviews of teachers and school district staff implementing IADA. I present qualitative evidence only after presenting quantitative evidence. Qualitative evidence offers validation or potential explanation for quantitative results, but it does not on its own generate a finding.

RQ3: Guidebook Usage

Two survey questions provided data related to treated and non-treated teachers’ use of the ELA Guidebooks 2.0 curriculum on which the Grade 7 IADA test is based. The first question asked the extent to which, over the source of a typical week, a teacher’s lessons were consistent with the scheduled sequence of Guidebook lessons. The second asked the extent to which, within a typical lesson, a teacher’s students were actively engaging with a Guidebook text. Results from these questions are below; a presentation of mediating factors from qualitative data follows quantitative analysis of both survey results.

Weekly Guidebook Consistency

Survey responses from 2019-2020 IADA-treated teachers (n=70) and non-treated teachers (n=69) are in Table 5.6. The majority of teachers in both IADA-treated (73%) and non-
treated (64%) groups reported “almost always” teaching all required Guidebooks lessons in a typical week. Few teachers reported only occasional use, and no teachers reported not using the curriculum at all.

Table 5.6

In a Typical Week, Frequency of Guidebooks Instruction, 2019-2020

<table>
<thead>
<tr>
<th></th>
<th>Almost always: I teach all lessons in a unit</th>
<th>%</th>
<th>Often: I teach most lessons in a unit</th>
<th>%</th>
<th>Occasionally: I teach some lessons in a unit</th>
<th>%</th>
<th>Never or hardly ever: I do not use Guidebook lessons</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>IADA</td>
<td>51</td>
<td>73</td>
<td>16</td>
<td>23</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No IADA</td>
<td>44</td>
<td>64</td>
<td>23</td>
<td>33</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

As discussed in chapter 4, for purposes of analysis, I ranked responses on an ordinal scale of 0-3. Descriptive statistics indicated that IADA-treated teachers ($M=2.69$, $SD=.553$) and non-treated teachers ($M=2.64$, $SD=.484$) did not differ greatly in following Guidebook units. A Mann-Whitney test indicated no significant difference in means, $U(\text{IADA}=70, \text{NO-IADA}=69) = 2233$, $z = -.947$, $p = .344$.

Survey responses from 2021-2022 IADA-treated teachers ($n=49$) and non-treated teachers ($n=132$) are presented in Table 5.7. The majority of IADA-treated teachers (63%) reported highest fidelity to Guidebooks lessons in a typical week, as did a majority of non-treated teachers (51%). A greater share of the non-treated group (11%) reported only occasionally teaching some Guidebook lessons than was the case in the IADA-treated group (0%).

Table 5.7

In a Typical Week, Frequency of Guidebooks Instruction, 2021-2022
Descriptive statistics indicated that IADA-treated teachers ($M=2.65$, $SD=.481$) and non-treated teachers ($M=2.39$, $SD=.685$) showed potential differences in frequency of Guidebook lesson usage. A Mann-Whitney test indicated a significant difference in means, $U(N_{IADA}=49, N_{NO-IADA}=132) = 2636, z = -2.155, p = .031$, with IADA-treated teachers more likely to teach Guidebook lessons in a given week. Effect size for this finding was small ($r=.160$). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(181) = .870, p = .436$, indicating that distributions within the two groups did not differ.

To evaluate differences between IADA-treated Guidebook lesson usage in 2019-2020 ($M=2.69$, $SD=.533$) and 2021-2022 ($M=2.65$, $SD=.481$), I used a Mann-Whitney test to compare means. The test showed no significant difference between the years, $U(N_{2019}=70, N_{2021}=49) = 2636, z = -.702, p = .483$. Thus, while there is promising evidence of IADA-treated teachers using Guidebooks with greater frequency in one year, there is neither compelling evidence of this difference nor evidence of a change over time in the IADA-treated group.

**Guidebook Text Usage within Lessons**

Survey responses from 2019-2020 IADA-treated teachers ($n=70$) and non-treated teachers ($n=69$) are presented in Table 5.8. The majority of teachers in both IADA-treated (64%) and non-treated (69%) groups reported that in a typical lesson their students are almost
always reading, discussing, writing about, or otherwise using a Guidebook-assigned text. Few teachers reported only occasional use, and no teachers reported not using the texts at all.

Table 5.8

Percentage of Typical Lesson in Which Students Use Guidebook Text, 2019-2020

<table>
<thead>
<tr>
<th></th>
<th>Almost always: More than two thirds of lesson</th>
<th>Often: Between one third and two thirds</th>
<th>Occasionally: Approximately one third of lesson</th>
<th>Never or hardly ever: I do not use Guidebook lessons</th>
</tr>
</thead>
<tbody>
<tr>
<td>IADA</td>
<td>45</td>
<td>64</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>No IADA</td>
<td>48</td>
<td>69</td>
<td>19</td>
<td>28</td>
</tr>
</tbody>
</table>

Descriptive statistics indicated that when responses are ranked on an ordinal scale of 0-3 IADA-treated teachers ($M=2.57, SD=.484$) and non-treated teachers ($M=2.67, SD=.533$) did not differ greatly as groups in usage of Guidebook texts. A Mann-Whitney test indicated no significant difference in means, $U(N_{IADA}=70, N_{NO-IADA}=69) = 2260, z = -.793, p = .428$.

Survey responses from 2021-2022 IADA-treated teachers ($n=49$) and non-treated teachers ($n=132$) are presented in Table 5.9. The majority of IADA-treated teachers (76%) reported that students used Guidebook texts almost always during lessons, as did a majority of non-treated teachers (52%). A greater share of the non-treated group reported only occasionally using tests during lessons (13%) or hardly ever using texts (1%).

Table 5.9

Percentage of Typical Lesson in Which Students Use Guidebook Text, 2021-2022

<table>
<thead>
<tr>
<th></th>
<th>Almost always: More than one third</th>
<th>Often: Approximately one third of lesson</th>
<th>Never or hardly ever: I do not use Guidebook lessons</th>
</tr>
</thead>
<tbody>
<tr>
<td>IADA</td>
<td>76%</td>
<td>52%</td>
<td>13%</td>
</tr>
<tr>
<td>No IADA</td>
<td>76%</td>
<td>52%</td>
<td>13%</td>
</tr>
</tbody>
</table>
Descriptive statistics indicated that when responses are scored on an ordinal scale of 0-3
IADA-treated teachers ($M=2.76$, $SD=.434$) and non-treated teachers ($M=2.36$, $SD=.763$) showed
differences between mean usage of Guidebook texts within lessons. A Mann-Whitney test
indicated a significant difference in means, $U(N_{IADA} = 49, N_{NO-IADA} = 132) = 2344$, $z = -3.233$, $p = .001$, with IADA-treated teachers more likely have students use Guidebook texts for large
portions of a typical lesson than were non-treated teachers in the 2021-2022 school year. Effect
size for this finding is small ($r=.240$). A Kolmogorov-Smirnov test also indicated significant
difference in distribution of responses, $D(181) = .1.434$, $p = .030$. Within the non-treated group
($n=132$) were 17 participants responding “occasionally” (12.8%) and 2 participants responding
“never or almost never” (1.5%); by comparison, the IADA-treated group included no
respondents indicating text usage at either of the two lower levels on the scale.

To evaluate differences between IADA-treated guidebook lesson usage in 2019-2020
($M=2.57$, $SD=.627$) and 2021-2022 ($M=2.76$, $SD=.434$), I used a Mann-Whitney test to compare
means. The test showed no significant difference between the years, $U(N_{2019} = 70, N_{2021} = 49) = 1492$, $z = -1.486$, $p = .137$. As with weekly Guidebook usage, the sum finding is promising
evidence of a difference in Guidebook usage within a lesson, found in one year, but not
compelling evidence found in both years of the study. The qualitative evidence presented next
provides evidence that may help to interpret promising evidence; it does not validate a
compelling response to RQ3.
Qualitative Evidence of Guidebook Usage Factors

Interviewed teachers consistently reported high levels of ELA Guidebooks 2.0 usage, both in keeping with the curriculum over the course of a week and in using Guidebook texts within a given lesson. In a sentiment repeated by many interviewees, one Grade 7 teacher said of IADA, “It gives purpose to the Guidebooks that we're teaching. And I love that because there are some times where you're like, ‘I don't want to teach this textbook.’ Now you, you know, you have a reason to.” While interviewed teachers attributed high levels of Guidebooks usage to factors that extend beyond IADA—teachers frequently cited school district support for Guidebooks implementation as an important factor in their decision making—the IADA role in compelling Guidebooks usage emerged profoundly in two themes expressed often during interviews: the specific knowledge students need to perform well on this curriculum-specific test and pacing urgency, the mandate to complete specific learning by specific dates as a function of a state test that measures certain learning at certain dates. I will discuss pacing urgency here and will address the role of knowledge in preparing for the test when responding to research question 5.

One teacher and literacy coach said of pacing urgency, “Pacing has to be more intentional [with IADA] because if you don’t do all activities with that text, the students won’t be prepared for the test.” Not only was the substance of the curriculum represented on the test, but also students were required to demonstrate knowledge from the curriculum at specific points in time throughout the year, motivating teachers to keep with the curriculum’s intended schedule and ostensibly to maximize learning productivity. Another teacher articulated a similar relationship between pacing urgency and what students need to succeed on a unit-based test:
[IADA] pretty much levels the playing field in that every kid comes with the same knowledge or they've had the same exposure to the knowledge. I will say that, as far as just operational day-to-day, those [three testing] windows, they can be challenging because in reality, you can't always complete a Guidebook lesson in a day. That's not realistic always. And so with those windows, if you're not doing a lesson a day, you start to feel the pressure and you start to get stressed. And then you're faced with, okay, what can I speed up? Can I cut something? You have to make decisions that you don't want to have to make because they need it all.

This was a duality interviewed teachers expressed consistently on pacing urgency: it enhanced the focus on Guidebooks, but it also necessitated trade-offs or did not allow for nuances in how students develop over the course of the year. Several teachers in particular expressed concerns with the timing of the October test administration window, noting both that the turbulence of the early school year does not allow for a full two months of instructional time and that IADA requires writing and fluency skills they would like more time to develop as the year starts. “At one of the professional developments,” said one Grade 7 teacher, “I said, I'm not sure my kids aren't ready to write in October. I'm not ready for you to judge me on how well my kids are writing in October. And the person running the PD goes, ‘Oh, we understand. They're not ready. And we take that into account.’ And my thought was, well, if we know they're not ready to test in October then why do we test them in October?” A director of curriculum and instruction said that pacing urgency can be in tension with the confidence with which a curriculum-specific test could otherwise imbue historically challenged student groups: “If the teacher shows up daily, I think, you know, there's a sense of comfort [on the test], like I said, that it means [students] are familiar with the characters, familiar with the story…. But there could be anxiety if we didn't
cover the whole book. Like my teacher wasn't here or I had to miss like three weeks. It’s hard to make up time in that way on this test. And, you know, we talked about that before, but those are some of the like the positives, and the double-edged sword to it as well.”

RQ4: Intentions in Teaching Texts

Two survey questions provided data related to treated and non-treated teachers’ intentions in teaching texts. As noted in chapter 4, these questions related to the principles teachers hold about what students should learn by engaging with texts, rather than the methods teachers employ in teaching texts. The first question discussed below asked respondents to make a binary choice between their intention to teach reading skills such as summarizing or to teach the meaning of texts. The second question asked respondents to select any number of statements from a list of possible aspirations they hold for student learning.

Comprehension Skills Versus Textual Meaning

As discussed in chapter 4, the survey asked respondents to choose one statement of two options in order to describe the respondents’ approach to teaching texts. One statement reads, “I prefer to teach particular skills and strategies (locating the main idea, describing author’s purpose), and then organize instruction around those skills and strategies.” The other statement reads, “I prefer to teach particular texts (novels, poems, articles), and then organize instruction around helping students to comprehend and analyze those texts.” Table 5.10 provides 2019-2020 responses from IADA treated and non-treated respondents.

Table 5.10

Preference to Teach Skills and Strategies versus Textual Meaning, 2019-2020

<table>
<thead>
<tr>
<th>I prefer to teach skills</th>
<th>%</th>
<th>I prefer to teach</th>
<th>%</th>
</tr>
</thead>
</table>

148
CURRICULUM-SPECIFIC READING TEST

<table>
<thead>
<tr>
<th></th>
<th>I prefer to teach skills and strategies</th>
<th>I prefer to teach particular texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>IADA</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>No IADA</td>
<td>59</td>
<td>73</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>105</td>
</tr>
</tbody>
</table>

The majority of U.S. reading teachers continue to use texts to teach students discrete, isolated skills, such as locating a text’s main idea, rather than assisting students with comprehending a whole text’s meaning, as also prescribed by CCSS (Opfer et al., 2016). It is thus notable that the majority of respondents (64%), all of whom use ELA Guidebooks 2.0, indicated a preference for organizing lessons about the meaning of particular texts. However, a Fisher’s exact test indicated no difference in the proportion of responses between IADA-treated and non-treated respondents ($p = .483$). Table 5.11 shows that similar frequencies also occurred between IADA-treated and non-treated groups in 2021-2022.

Table 5.11

Preference to Teach Skills and Strategies versus Textual Meaning, 2021-2022

The share of 2021-2022 IADA-treated respondents favoring organizing lessons around particular texts (65%) was consistent with 2019-2020 responses from treated participants (67%). A Fisher’s exact test indicated no difference in the proportion of responses between IADA-treated and non-treated respondents ($p = .241$). A Fisher’s exact test further showed no difference in
proportion of responses among IADA treated participants from 2019-2020 to 2021-2022. There thus was neither a difference between treated and non-treated teachers’ intent in teaching texts in either given year nor a difference in treated teachers’ intentions in teaching texts over time.

**Definitions of Student Comprehension**

To further evaluate the principles that guide teachers’ approach to teaching texts, a survey item asked respondents to select all listed statements that accurately describe the respondent’s perspective on what it means to understand a text. Repeated differences between groups could indicate a correlation between IADA participation and how teachers approach reading comprehension. Respondents selected from up to four statements or could list “none of the above.” Table 5.12 shows frequencies for 2019-2020 responses.

Table 5.12

*Teachers Endorsing Definitions of Understanding a Text, 2019-2020*

<table>
<thead>
<tr>
<th></th>
<th>IADA</th>
<th>%</th>
<th>IADA</th>
<th>%</th>
<th>IADA</th>
<th>%</th>
<th>IADA</th>
<th>%</th>
<th>IADA</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student can accurately paraphrase what the text is saying</td>
<td>59</td>
<td>84</td>
<td>58</td>
<td>83</td>
<td>45</td>
<td>64</td>
<td>58</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student can articulate opinions backed by textual evidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student analyzes text by examining figurative language or textual structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student describes a text’s relationship to other texts, art, or histories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While both groups agreed with a definition of comprehension focused on articulating opinions backed by textual evidence at high rates, IADA-treated participants supported the definition at a lower rate (83%) than did non-treated teachers (93%). A Fisher’s exact test indicated the difference was significant \(p = .075\). IADA-treated participants also supported a definition of comprehension focused on examining figurative language and textual structure at a lower rate.
CURRICULUM-SPECIFIC READING TEST

(64%) than do non-treated teachers (75%). A Fisher’s exact test indicated the difference is not significant ($p = .202$). All other differences were also insignificant. Table 5.13 presents 2021-2022 responses.

Table 5.13

*Teachers Endorsing Definitions of Understanding a Text, 2021-2022*

<table>
<thead>
<tr>
<th></th>
<th>Student can accurately paraphrase what the text is saying</th>
<th>Student can articulate opinions backed by textual evidence</th>
<th>Student analyzes text by examining figurative language or textual structure</th>
<th>Student describes a text’s relationship to other texts, art, or histories</th>
</tr>
</thead>
<tbody>
<tr>
<td>IADA</td>
<td>42</td>
<td>86</td>
<td>92</td>
<td>33</td>
</tr>
<tr>
<td>No IADA</td>
<td>113</td>
<td>86</td>
<td>125</td>
<td>95</td>
</tr>
</tbody>
</table>

The rate at which treated teachers in 2021-2022 supported a definition related to other texts, art, and histories differed between IADA-treated teachers (71%) and non-treated teachers (59%). A Fisher’s exact test indicated the difference is not significant ($p = .169$). Relatedly, the percentage of IADA-treated teachers responding favorably to the definition involving other texts, art, and histories in 2021-2022 (71%) was lower than the percentage of treated teachers supporting that definition in 2019-2020 (83%). A Fisher’s exact test indicated that this difference was not significant ($p = .177$). When tested, no other differences were significant. The results presented about indicate no compelling evidence of differences between the treated and control groups. There further was also no evidence of differences within the control group over time.

Qualitative Evidence of Intentions in Teaching Texts

Interviews yielded no consistent themes related to teachers’ intentions in teaching texts. Asked how they define successful comprehension in their classrooms and what purpose teaching
texts serves for students, interviewees provided wide-ranging responses that did not indicate the presence of a predominant principle or indicate a factor related to the absence of principle. This lack of a principle distinguishing principles of IADA-treated teachers aligns with quantitative results presented above. I will discuss the absence of compelling differences in response to this question in the discussion section of this chapter.

**RQ5: Classroom Activities When Using Texts**

Two survey questions provided data related to activities in which treated and non-treated teachers have students engage when using texts. The first question asked respondents to quantify the extent to which they have students engage with texts of appropriate complexity for the student’s grade level and the extent to which they have students engage in texts whose complexity corresponds with a given student’s reading performance level. The second question asked respondents to quantify the frequency with which they have students engage in activities relevant to an integrated reading model’s emphasis on knowledge derived from both within the text and knowledge beyond the text.

**Text Complexity Levels**

To evaluate the approach teachers use when determining the complexity of texts they assign in class, one survey item asked respondents to quantify the proportion of in-class time students are alternately engaged with texts of a complexity equivalent to the students’ grade level versus texts of a complexity equivalent to students’ reading levels. Respondents selected a proportion of time across a continuum of four potential responses, each one of which included a range of time spent: 0-25%, 25-50%, 50-75%, and 75-100%. Repeated differences between
groups could indicate a correlation between IADA participation and how teachers approach text selection in the classroom. Table 5.14 shows frequencies for 2019-2020 responses.

Table 5.14

*Percent of Time Students Engage Grade-level versus Reading-level Texts in-class, 2019-2020*

<table>
<thead>
<tr>
<th>Grade-level</th>
<th>IADA</th>
<th>No IADA</th>
<th>IADA</th>
<th>No IADA</th>
<th>IADA</th>
<th>No IADA</th>
<th>IADA</th>
<th>No IADA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Always or almost always (75-100%)</td>
<td>%</td>
<td>Often (50-75%)</td>
<td>%</td>
<td>Occasionally (25-50%)</td>
<td>%</td>
<td>Never or hardly ever (0-25%)</td>
<td>%</td>
</tr>
<tr>
<td>Grade-level</td>
<td>27</td>
<td>39</td>
<td>29</td>
<td>41</td>
<td>12</td>
<td>18</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>No IADA</td>
<td>26</td>
<td>38</td>
<td>29</td>
<td>42</td>
<td>13</td>
<td>19</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Reading-level</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>16</td>
<td>32</td>
<td>46</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>No IADA</td>
<td>3</td>
<td>4</td>
<td>14</td>
<td>20</td>
<td>26</td>
<td>38</td>
<td>26</td>
<td>38</td>
</tr>
</tbody>
</table>

In 2019-2020, IADA-treated and non-treated teachers reported nearly identical rates of grade-level text assignment. It is notable that the majority of respondents in the entire sample (70%) report always or often assigning grade-level texts, as this runs counter to national trends among teachers, who tend to assign reading-level texts (Opfer et al., 2016). But a Mann-Whitney test indicated no difference among groups in grade level text assignments, $U(N_{IADA} = 70, N_{NO-IADA} = 69) = 2400.5$, $z = -.066$, $p = .948$.

Means differed modestly in assigning reading level tests between IADA-treated teachers ($M=1.11, SD=.941$) and non-treated teachers ($M=.94, SD=.889$). A Mann-Whitney test indicated no significant difference, $U(N_{IADA} = 70, N_{NO-IADA} = 69) = 2189$, $z = -1.008$, $p = .314$. Table 5.15 shows frequencies for 2021-2022 responses.
CURRICULUM-SPECIFIC READING TEST

Table 5.15

| Percent of Time Students Engage Grade-level versus Reading-level Texts in-class, 2021-2022 |
|---------------------------------|---|---|---|---|---|---|
|                                | Always or almost always (75-100%) | %  | Often (50-75%) | %  | Occasionally (25-50%) | %  | Never or hardly ever (0-25%) | %  |
| Grade-level                    | IADA | 26 | 53 | 18 | 37 | 5 | 10 | 0 | 0 |
|                                | No IADA | 47 | 36 | 60 | 45 | 19 | 14 | 6 | 5 |
| Reading-level                  | IADA | 3 | 6 | 6 | 12 | 25 | 50 | 16 | 32 |
|                                | No IADA | 10 | 8 | 27 | 20 | 53 | 40 | 42 | 32 |

In 2021-2022, mean scaled responses on grade-level text assignment from IADA-treated teachers ($M=2.43, SD=.677$) and non-treated teachers ($M=2.12, SD=.820$) indicated differences in the frequency with which groups assign grade-level texts. A Mann-Whitney test indicated a significant difference, $U(N_{IADA}=49, N_{NO-IADA}=132) = 2579.5, \ z=-2.264, p = .025, with IADA-treated teachers reporting a greater share of in-class time using grade-level texts than non-treated teachers. Effect size of this difference was small ($r=.168$). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(181) = 1.043, p = .226$.

Mean ranked responses on reading-level text assignment from IADA-treated teachers ($M=.94, SD=.827$) and non-treated teachers ($M=1.04, SD=.911$) indicated only minor differences in approach to assigning reading-level texts. A Mann-Whitney test indicated no significant difference, $U(N_{IADA}=49, N_{NO-IADA}=132) = 3059.5, \ z=-.593, p = .533$. Mean responses on grade-level text assignment among IADA-treated teachers indicated potential differences
CURRICULUM-SPECIFIC READING TEST

between 2019-2020 respondents (M=2.16, SD=.810) and 2021-2022 respondents (M=2.43, SD=.677). A Mann-Whitey test indicated a significant difference, \( U(N_{2019} = 70, N_{2021} = 49) = 1408, z = -1.801, p = .072 \), with 2021-2022 IADA-treated respondents assigning grade-level texts for greater portions of lessons than 2019-2020 IADA-treated respondents. Effect size of this difference was small (\( r = .165 \)). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, \( D(119) = .778, p = .580 \). Mean responses on reading-level text assignment among IADA-treated teachers indicated modest differences between 2019-2020 respondents (M=1.11, SD=.941) and 2021-2022 respondents (M=.94, SD=.827). A Mann-Whitney test indicated no significant difference, \( U(N_{2019} = 70, N_{2021} = 49) = 1554.5, z = -.933, p = .351 \). In discussing results, I will address the promising though not compelling evidence of differences on grade-level text assignment between the two groups.

Classroom Activities

Respondents indicated the frequency with which they had students engage in specific classroom activities when engaged with texts. Activities included using evidence from the text to analyze a text, analyzing the structure of a text, analyzing two texts with similar themes, and connecting literacy content to other subject matters, such as social studies. Table 5.16 presents the frequency of responses in 2019-2020.

Table 5.16

<table>
<thead>
<tr>
<th>Activity</th>
<th>Daily</th>
<th>%</th>
<th>Often (1-3 times per week)</th>
<th>%</th>
<th>Occasionally (1-3 times per month)</th>
<th>%</th>
<th>Never or hardly ever</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use evidence from text</td>
<td>IADA</td>
<td>50</td>
<td>71</td>
<td>20</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>No IADA</td>
<td>54</td>
<td>78</td>
<td>13</td>
<td>19</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>
IADA-treated teachers engaged students more frequently in analyzing textual structure ($M=2.29$, $SD=.663$) than did non-treated teachers ($M=2.10$, $SD=.667$). IADA-treated teachers also engaged students more frequently in analyzing two texts with related themes ($M=2.19$, $SD=.666$) than did non-treated teachers ($M=1.69$, $SD=.670$). IADA-treated teachers more frequently integrated other subjects into literacy instruction ($M=1.96$, $SD=.824$) than did non-treated teachers ($M=1.67$, $SD=.741$).

Mann-Whitney tests showed two of these means differences to be significant. IADA-treated teachers engaged students significantly more frequently in analyzing two texts with similar themes, $U(N_{IADA} = 70, N_{NO-IADA} = 69) = 1803, z = -2.841, p = .004$. Effect size of this difference was small ($r=.241$). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(139) = .997, p = .273$. IADA-treated teachers also significantly more frequently connected literacy instruction to other content, $U(N_{IADA} = 70, N_{NO-IADA} = 69) = 1939.5, z = -2.138, p = .033$. Effect size was small ($r = .181$). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(139) = .914 p = .374$. 

<table>
<thead>
<tr>
<th>Daily</th>
<th>%</th>
<th>Often (1-3 times per week)</th>
<th>%</th>
<th>Occasionally (1-3 times per month)</th>
<th>%</th>
<th>Never or hardly ever</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyze text structure</td>
<td>IADA</td>
<td>28</td>
<td>40</td>
<td>34</td>
<td>49</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>No IADA</td>
<td>18</td>
<td>26</td>
<td>38</td>
<td>55</td>
<td>12</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Two texts w/ same theme</td>
<td>IADA</td>
<td>21</td>
<td>30</td>
<td>37</td>
<td>53</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>No IADA</td>
<td>11</td>
<td>16</td>
<td>37</td>
<td>54</td>
<td>21</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>Connect to other subject matter</td>
<td>IADA</td>
<td>20</td>
<td>29</td>
<td>27</td>
<td>39</td>
<td>22</td>
<td>31</td>
</tr>
<tr>
<td>No IADA</td>
<td>10</td>
<td>14</td>
<td>27</td>
<td>39</td>
<td>31</td>
<td>44</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 5.17 shows the frequencies of select classroom activities when using texts in 2021-2022.

Table 5.17

*Frequency of Classroom Activities When Using Texts, 2021-2022*

<table>
<thead>
<tr>
<th>Activity</th>
<th>Daily</th>
<th>% Often (1-3 times per week)</th>
<th>% Occasionally (1-3 times per month)</th>
<th>% Never or hardly ever</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use evidence from text</td>
<td>IADA</td>
<td>35</td>
<td>71</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>No IADA</td>
<td>107</td>
<td>81</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>Analyze text structure</td>
<td>IADA</td>
<td>9</td>
<td>18</td>
<td>32</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>No IADA</td>
<td>48</td>
<td>36</td>
<td>58</td>
<td>44</td>
</tr>
<tr>
<td>Two texts w/ same theme</td>
<td>IADA</td>
<td>10</td>
<td>20</td>
<td>29</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>No IADA</td>
<td>24</td>
<td>18</td>
<td>71</td>
<td>54</td>
</tr>
<tr>
<td>Connect to other subject matter</td>
<td>IADA</td>
<td>4</td>
<td>8</td>
<td>23</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>No IADA</td>
<td>27</td>
<td>20</td>
<td>52</td>
<td>39</td>
</tr>
</tbody>
</table>

In 2021-2022, IADA-treated teachers did not engage students more frequently in analyzing textual structure ($M=2.02, SD=.595$) than did non-treated teachers ($M=2.18, SD=.729$). IADA-treated teachers also did not more frequently integrate other subjects into literacy instruction ($M=1.59, SD=.705$) than did non-treated teachers ($M=1.70, SD=.906$). IADA-treated teachers again engaged students more frequently in analyzing two texts with related themes ($M=2.00, SD=.645$) than did non-treated teachers ($M=1.88, SD=.721$). A Mann-Whitney test indicated that this difference was not significant, $U(N_{IADA}=49, N_{NO-IADA}=132) = 2965, z = -.962, p = .341$. 
The evidence that there are differences between the groups in the frequency with which they have students assign two texts was consistent but remains only promising and not compelling.

Mean responses from IADA-treated participants in 2021-2022 indicate pronounced differences from IADA-treated respondents in 2019-2020. Table 5.18 presents descriptive statistics for IADA-treated teachers in the two years.

Table 5.18

<table>
<thead>
<tr>
<th></th>
<th>Text evidence</th>
<th>Text structure</th>
<th>Two texts</th>
<th>Other subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2019</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.70</td>
<td>2.29</td>
<td>2.19</td>
<td>1.96</td>
</tr>
<tr>
<td>N</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>SD</td>
<td>.455</td>
<td>.663</td>
<td>.666</td>
<td>.824</td>
</tr>
<tr>
<td><strong>2021</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.71</td>
<td>2.02</td>
<td>2.00</td>
<td>1.59</td>
</tr>
<tr>
<td>N</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>SD</td>
<td>.456</td>
<td>.595</td>
<td>.645</td>
<td>.705</td>
</tr>
</tbody>
</table>

Means differences between years in analyzing text structure (.27), analyzing two texts of similar themes (.19), and including other subjects in literacy instruction (.37) all indicate that IADA-treated teachers report less usage of reading activities in 2021-2022 than in 2019-2020. A Mann-Whitney test indicated that differences in analyzing text structure are significant, $U(N_{2019} = 70, N_{2021} = 49) = 1336, z = -2.292, p = .022$, with 2021-2022 teachers less likely to have students analyze text structure. Effect size for this difference was small ($r = .210$), and a Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(119) = 1.161, p =$
A Mann-Whitney test also indicated that differences in using content from other subjects in the ELA classroom were significant, $U(N_{2019} = 70, N_{2021} = 49) = 1307, z = -2.352, p = .019$, with 2021-2022 teachers less likely use other subject area content. Effect size for this difference was small ($r = .216$), and a Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(119) = 1.172, p = .128$. Differences in having students compare other, related texts were not significant, $U(N_{2019} = 70, N_{2021} = 49) = 1461.5, z = -1.527, p = .127$.

**Qualitative Classroom Evidence**

Evidence gathered from interviews yielded two emergent themes that may provide insight into evidence presented in this section. The first emergent theme, like texts, includes teachers’ descriptions of their efforts and their districts’ efforts to have students draw thematic connections based on knowledge gleaned from thematically related texts. The theme of like texts may explain emerging evidence of differences between groups in their use of two similar texts, which was significant in 2019-2020 and present but not significant in 2021-2022. The second emergent theme, social studies changes, relates to the LDOE shift toward requiring IADA students to take a separate social studies test, in addition to IADA. This shift away from considering IADA a test of social studies knowledge unto itself may have contributed to significant declines in the treated group’s inclusion of other subject matter between 2019-2020 and 2021-2022.

**Like Texts.** This chapter will later discuss the broad and specific influences of standardized tests and their mechanics over teachers, in responding to research question 6. Gathered here is qualitative evidence that IADA-treated teachers are compelled to engage students in drawing thematic connections between texts that have been sequenced specifically because of those thematic connections. This is relevant to promising quantitative evidence
presented above indicating differences between the frequency with which the two groups have students engage with similar texts. In interviews, teachers repeatedly stated that the IADA test, which asks students not only to analyze passages presented on the test form but also to draw on knowledge of texts not presented on the test form, creates an incentive for routinely and explicitly making these connections in classroom discussions and in written responses. “Being able to track and connect those stories,” said one head of curriculum, “because at the end of the day, you're going to have to compare and contrast multiple stories that you're covered. So [IADA] gives us links and gives meaning throughout the entire year, when you're doing the Innovative assessment versus just LEAP 2025.” Another teacher contrasted this focus on cross-text comparisons with prior teaching practices influenced by the conventional state test: “We also spend a lot of time synthesizing text because that's where we struggle so much. We spend more time now when we're pulling from our anchor text and then some of the other reading that we do. We work really hard and spend much more time than I had anticipated just synthesizing two texts.”

The impulse in IADA-treated teachers to create connections across texts is so strong that teachers and districts use time outside of class or outside of time spent using Guidebooks texts to foster text connections and further build students’ knowledge within a given unit’s knowledge domain. “Okay so for example with my seventh grade, we were headed into the spring and my unit gets done,” said one teacher. “So I'm not going to pull in anything like move on to another unit or anything because I know what we're fixing to be tested on. And so I pulled more things that were tied to the curriculum, that unit, and I made sure that's what we were focusing on.”

Given additional time in the school year, with Grade 7 students participating in IADA, the teacher assigns additional texts related to the Guidebook unit’s knowledge domain. In the
interview, she contrasts this with how she approaches teaching students participating in the conventional state ELA test. “However, with my eighth graders, I went ahead and started doing some cold reads and dissecting prompts that were tied to cold reads for research and simulation task and literary analysis. Because I wanted them to have some practice… So [the test] changes things a little bit in terms of what direction I take.”

One district official described the district’s effort to replicate this cross-text practice outside the curriculum across all schools, by creating unit-aligned “book bins” from which students select grade-level texts for independent reading in class and out-of-class. Interviewed teachers from that district consistently noted and reinforced the centrality of that strategy to their effort to help students build knowledge that can later be represented on the IADA test. Teachers consistently stated that this and other efforts to build cross-text connections were a function of specific and new tasks on the IADA test form. They rarely commented on their own philosophy or their own opinion of new the test’s direction. Building knowledge across texts was necessary for students to prepare for specific requirements of IADA, as one teacher demonstrated in describing the case of Grade 7 students who enroll in or return to the school mid-year and are thus required to take both relevant IADA unit test(s) and the end-of-year LEAP:

I feel like the writing in the curriculum is very different for LEAP and Innovative [test]… I have to give the LEAP ELA test to a handful of students this year for people who are not here for both windows of the Innovative [test], so I feel like those kids are not going to be prepared for their writing because [IADA] writing is definitely about two texts. I just feel [on LEAP] you have the text in front of you, you're pulling textual evidence. So I feel like those kids, I don't even know what's going to happen with them. But then I feel like since they do not have the text or the Innovative [test] anchor text [in front of them],
it's like we're changing the way writing has been done since these kids got in like third grade. Since third grade, we've been teaching them to pull textual evidence.

This challenge of writing based on built, stored knowledge, rather than, as one teacher called it, “henpecking for evidence,” was frequently present in teachers’ stories about teaching in the IADA context and preparing for specific requirements of the text. One teacher who expressed support for the knowledge-building, cross-text approach IADA compelled, gave voice to the challenges students have recalling and applying knowledge in writing:

Oh, the application portion is a big difference. You know, in the past, our students would have two or three different texts. They were all dissimilar. Now they have their texts, but they read a new text and they have to apply the same knowledge that they learned throughout the unit. And not only that, but in their writing as well. And our students are struggling with the with the application piece because it's something that we haven't done before, but it's necessary for what they need when they graduate and when they leave us. So that's one of the main differences and that's why we're struggling. And that's when we incorporate more knowledge texts, because that's something that traditionally our students have not done. They haven't been forced to read a text then make connections on their own.

This chapter will later present evidence of the role of the IADA test as a source of influence over teachers approaches to teaching texts. This section has presented initial quantitative and qualitative evidence that the specific format of the test may have compelled specific shifts toward cross-text connections to build knowledge.

**Social Studies Changes.** This section has included quantitative evidence of significant differences in IADA-treated teachers’ classroom activities between 2019-2020 and 2021-2022.
These differences in particular indicated that IADA-treated teachers were less likely in 2021-2020 to have students analyze a text’s structure and to introduce other subject matter, such as social science content, into the literacy context. Presented here is repeated qualitative evidence of a change in statewide social studies policy and implementation, which relates in particular to differences in IADA-treated teachers’ use of other subject matters over time.

Early IADA-related communications from LDOE, and subsequent reporting on IADA, emphasized the integration of social studies content into the IADA test as a prominent feature of the initiative (Sentell, 2018). Not only did inclusion of social studies content affect the knowledge required for students to perform on the test, but also it obviated the need for the state to require IADA-participating districts to administer both a state ELA test and a state social studies test (Sentell, 2018). In 2021, BESE adopted new social studies standards for grades K-12, including a new grade-level sequence for teaching historical periods, geographies, and cultures (Sentell, 2022, March 8). In 2022, LDOE administered distinct Grade 7 social studies tests to students in IADA-participating districts, a shift from past policy that had exempted these districts from the traditional social studies Grade 7 LEAP (Sentell, 2022, August 3). While interviewed teachers did not express concrete knowledge of the implications of this shift for IADA tests specifically, their technical understanding of how standards influence curriculum, as well as informal communications they had received, indicated that integration of social studies content for purposes of IADA testing would no longer be feasible.

Interviewed educators frequently reported that building coherent bodies of knowledge to prepare for the IADA test was more easily facilitated in units with non-fiction anchor texts.38

38 While teachers reported in interviews that they valued non-fiction texts for ease of knowledge-building, this orientation was not reflected in districts’ choices of units. In 2021-2022, only two of ten participating districts chose a non-fiction anchor text for the fall unit; this meant that only
One teacher articulated the promise of integrating non-fiction Guidebook units with social studies instruction by describing the knowledge of U.S. slavery and the Civil War era her students would possess when entering Grade 8 because of their engagement with the Grade 7 Guidebooks unit anchored by Elizabeth Keckley’s (2012) enslavement and Civil War narrative, *Behind the Scenes*. This knowledge would not only be represented and reinforced in a Grade 8 Guidebooks unit and text on the sugar trade, but also by Grade 8 social studies standards focused on antebellum agriculture, commerce, and enslavement in Louisiana. This synergy, the teacher said, would now be disrupted by state-mandated shifts in the period of history covered in Grade 8.

Another school had dispensed with separating Grade 7 ELA and social studies classrooms altogether. Prompted by IADA, the school instead chose to integrate social studies and literature into a class called “humanities”:

[IADA] set up for a partnership between the two subjects that might have been sometimes related, but most of the time isolated. Our humanities teachers worked in tandem together and planned and were participating in the English section as well as the social studies…. So it us to shift in that way. But then it was even more meaningful right on top of it, because you had to plan together. It wasn't an isolation thing. And I think it needs to go back to where the social studies is embedded with English. That was a powerful piece. That was my attraction to the Innovative [test] to begin with.

Teachers mentioned this sense of regret frequently when the issue of social studies arose in interviews. One Grade 7 teacher discussed assigning additional non-fiction texts in parallel to

-------------------

two districts had students read more than one non-fiction anchor text throughout the course of 2021-2022.
the Keckley (2012) text, such as *Narrative of the Life of Frederick Douglass* (1845), again to bolster students’ base of knowledge and ability to compare like texts. “I’ve always taught a novel and then a text very similar to what was with the Guidebooks, and I partnered with [the social studies teacher] a lot to do some cross curricular stuff. I think that a big opportunity. That's what made me very excited about the Innovative [test], was that tie between social studies and ELA. And of course, now that's gone to the wayside.” What was lost, these teachers said, is not their ability to add knowledge-building texts as much as the chance to co-create an integrated approach with a social studies teacher. With social studies again dedicated to its own standards implementation and its own test, no longer overlapping with the sequence of ELA texts and knowledge domains, that incentive to co-create has been lost.

**RQ6: Factors Influencing Teaching Approach**

A final survey question asked respondents to rate the extent to which 12 external factors influenced their approaches to planning lessons. The question asked respondents to evaluate the level of influence each factor presents, on a four-point continuum from “a great deal” to “not at all.” Data provided evidence both for identifying factors that most influence treated and non-treated teacher groups and for discerning where treated and non-treated groups differ. Results indicating factors that influence teachers also provided important context useful to synthesizing findings from across the study, per the discussion session at the end of this chapter.

**2019-2020 Influence Factors**

In 2019-2020, factors identified by the highest percentage of IADA-treated teachers as influencing their lesson plans “a great deal” were district requirements (91%), state ELA
CURRICULUM-SPECIFIC READING TEST

standards (86%), ELA Guidebooks 2.0 (83%), LEAP 2025 or LEAP Humanities (79%), and louisianabelieves.com (70%). The four highest influence factors by this measure for IADA-treated teachers matched the four highest factors for non-treated teachers; only non-treated teachers’ responses to louisianabelieves.com (46%) did not rank as highly influential for non-treated teachers. Table 5.19 presents mean scaled figures for each factor’s influence over both groups.

Table 5.19

Mean Influence of Factors over Lesson Plans, 2019-2020

<table>
<thead>
<tr>
<th>Factor</th>
<th>Group</th>
<th>Mean</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>State ELA standards</td>
<td>IADA-treated</td>
<td>2.81</td>
<td>.490</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.74</td>
<td>.585</td>
</tr>
<tr>
<td>Louisianabelieves.com</td>
<td>IADA-treated</td>
<td>2.61</td>
<td>.644</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.04</td>
<td>1.104</td>
</tr>
<tr>
<td>School or district requirements</td>
<td>IADA-treated</td>
<td>2.89</td>
<td>.401</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.86</td>
<td>.463</td>
</tr>
<tr>
<td>LEAP 2025 or LEAP Humanities</td>
<td>IADA-treated</td>
<td>2.69</td>
<td>.692</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.51</td>
<td>.760</td>
</tr>
<tr>
<td>LEAP 360&lt;sup&gt;40&lt;/sup&gt;</td>
<td>IADA-treated</td>
<td>1.21</td>
<td>1.190</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.13</td>
<td>.999</td>
</tr>
</tbody>
</table>

<sup>39</sup> These are brands with which teachers would be familiar, signifying the conventional state test and the IADA test respectively.

<sup>40</sup> As noted in chapter 4, LEAP 2025 is the name of the conventional, state-administered end-of-year testing series. LEAP Humanities is jargon for IADA. LEAP 360 is a free and optional test districts may choose to administer throughout the course of the school year.
I tested all constructs for differences in mean influence. Three factors indicated means differences in which non-treated teachers report greater influence than do IADA-treated teachers: online networks, teacher evaluations, and LEAP 360. A Mann-Whitney test indicated significant differences in the influence of online teacher networks, \( U(N_{IADA} = 70, N_{NO-IADA} = 69) = 1489, z = -4.112, p < .001 \), with non-treated teachers experiencing greater influence. Effect size for this difference was moderate \( (r = .349) \), and a Kolmogorov-Smirnov test indicated significant difference in distribution of responses, \( D(139) = 2.087, p < .001 \). A Mann-Whitney test also

<table>
<thead>
<tr>
<th>Factor</th>
<th>Group</th>
<th>Mean</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>District assessments</td>
<td>IADA-treated</td>
<td>2.41</td>
<td>.925</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.42</td>
<td>.864</td>
</tr>
<tr>
<td>ELA Guidebooks 2.0</td>
<td>IADA-treated</td>
<td>2.76</td>
<td>.600</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.88</td>
<td>.365</td>
</tr>
<tr>
<td>Pre-service preparation</td>
<td>IADA-treated</td>
<td>1.04</td>
<td>.984</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>1.09</td>
<td>1.197</td>
</tr>
<tr>
<td>Teacher evaluation</td>
<td>IADA-treated</td>
<td>1.77</td>
<td>1.182</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.16</td>
<td>1.009</td>
</tr>
<tr>
<td>Teacher Leader Summit</td>
<td>IADA-treated</td>
<td>.97</td>
<td>1.129</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>1.06</td>
<td>1.162</td>
</tr>
<tr>
<td>Prof. development</td>
<td>IADA-treated</td>
<td>1.94</td>
<td>1.058</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>2.22</td>
<td>.855</td>
</tr>
<tr>
<td>Online teacher network</td>
<td>IADA-treated</td>
<td>.67</td>
<td>.896</td>
</tr>
<tr>
<td></td>
<td>Non-treated</td>
<td>1.36</td>
<td>1.043</td>
</tr>
</tbody>
</table>
CURRICULUM-SPECIFIC READING TEST

indicated significant differences in the influence of teacher evaluation systems, $U(N_{IADA} = 70, N_{NO-IADA}= 69) = 1988.5, z = -1.901, p = .057$, with non-treated teachers experiencing greater influence. Effect size for this difference was small ($r = .161$), and a Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(139) = .991, p = .280$.

Differences in the influence of LEAP 360 were significant, $U(N_{IADA} = 70, N_{NO-IADA}= 69) = 1402, z = -4.445, p < .001$, with non-treated teachers experiencing greater influence. Effect size for this difference was moderate ($r = .377$), and a Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, $D(139) = 2.172, p < .001$.

Two factors indicated means differences in which IADA-treated teachers report greater influence: louisianabelieves.com and LEAP. A Mann-Whitney test indicated that IADA-treated teachers were more influenced by louisianabelieves.com, in 2019-2020 than were non-treated teachers, $U(N_{IADA} = 70, N_{NO-IADA}= 69) = 1733.5, z = -3.238, p = .001$. Effect size for this difference was small ($r = .274$). A Kolmogorov-Smirnov test indicated a significant difference in distribution of responses, $D(139) = 1.393, p = .041$. There were no other significant differences in which 2019-2020 IADA-treated teachers reported greater influence than non-treated teachers.

2021-2022 Influence Factors

In 2019-2020, factors identified by the highest percentage of IADA-treated teachers as influencing their lesson plans “a great deal” were district requirements (96%), ELA Guidebooks 2.0 (94%), LEAP 2025 or LEAP Humanities (84%), state ELA standards (82%), and louisianabelieves.com (65%). Four of these high-influence factors for IADA-treated teachers were also among the five highest-influence factors for non-treated teachers, by this measure; only non-treated teachers’ responses to the influence of LEAP (47%) did not rank as highly
influential for non-treated teachers. District assessment influence ranked as a top-five influence factor for non-treated teachers instead (58%). Table 5.20 presents mean scaled figures for each factor’s influence over both groups.

Table 5.20

<table>
<thead>
<tr>
<th>Mean Influence of Factors over Lesson Plans, 2021-2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>--------------------------------</td>
</tr>
<tr>
<td>State ELA standards</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Louisianabelieves.com</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>School or district requirements</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LEAP 2025 or LEAP Humanities</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LEAP 360</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>District assessments</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ELA Guidebooks 2.0</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Pre-service preparation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Teacher evaluation</td>
</tr>
</tbody>
</table>
### CURRICULUM-SPECIFIC READING TEST

<table>
<thead>
<tr>
<th>Factor</th>
<th>Group</th>
<th>Mean</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Leader Summit</td>
<td>Non-treated</td>
<td>1.29</td>
<td>1.176</td>
</tr>
<tr>
<td>IADA-treated</td>
<td>.49</td>
<td>.681</td>
<td></td>
</tr>
<tr>
<td>Non-treated</td>
<td>.58</td>
<td>.891</td>
<td></td>
</tr>
<tr>
<td>Prof. development</td>
<td>IADA-treated</td>
<td>2.22</td>
<td>.855</td>
</tr>
<tr>
<td>Non-treated</td>
<td>1.94</td>
<td>1.058</td>
<td></td>
</tr>
<tr>
<td>Online teacher network</td>
<td>IADA-treated</td>
<td>1.36</td>
<td>1.043</td>
</tr>
<tr>
<td>Non-treated</td>
<td>.67</td>
<td>.896</td>
<td></td>
</tr>
</tbody>
</table>

As in 2019-2020, LEAP 360, the optional, statewide interim assessment series, significantly influenced non-treated teachers more than IADA-treated teachers, \( U(N_{IADA} = 49, N_{NO-IADA} = 132) = 2737, z = -1.650, p = .099 \). Effect size for this difference was small (\( r = .122 \)). A Kolmogorov-Smirnov test indicated no significant difference in distribution of responses, \( D(181) = .945, p = .333 \). This is a repeated, compelling finding of a statewide interim assessment’s influence having a negative correlation with IADA participation.

Three factors influenced IADA-treated teachers significantly more in 2021-2022 than non-treated peers. LEAP 2025 and LEAP Humanities, the statewide, mandated end-of-year test, influenced IADA-treated teachers significantly more than non-treated peers, \( U(N_{IADA} = 49, N_{NO-IADA} = 132) = 2042, z = -4.276, p < .001 \). Effect size for this difference was moderate (\( r = .318 \)), and a Kolmogorov-Smirnov test indicated a significant difference in distribution of responses, \( D(181) = 2.149, p < .001 \). ELA Guidebooks 2.0 influenced IADA-treated teachers significantly more than peers, \( U(N_{IADA} = 49, N_{NO-IADA} = 132) = 2705, z = -2.545, p = .011 \). Effect size for this difference was small (\( r = .189 \)), and a Kolmogorov-Smirnov test indicated no significant
difference in distribution of responses, $D(181) = .947, \ p = .331$. Last, a Mann-Whitney test indicated significant differences in the influence of online teacher networks, \( U(N_{IADA} = 49, N_{NO-IADA} = 132) = 2303.5, z = -3.216, p = .001 \). Effect size for this difference was small ($r = .239$), and a Kolmogorov-Smirnov test indicated a significant difference in distribution of responses, $D(181) = 2.189, \ p < .001$.

**Differences in Influence between 2019-2020 and 2021-2022**

Results presented in this section have indicated that the factors most influencing IADA-treated teacher classroom lessons were generally consistent between 2019-2020 and 2021-2022. Table 5.21 presents results for IADA-treated teachers only over both years.

Table 5.21

*Influence Factor Descriptive Statistics for Treated Teachers, 2019-2020 and 2021-2022*

<table>
<thead>
<tr>
<th>Factor</th>
<th>Year</th>
<th>N</th>
<th>M</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>State standards</td>
<td>2019</td>
<td>70</td>
<td>2.81</td>
<td>.490</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.80</td>
<td>.456</td>
</tr>
<tr>
<td>Louisianabelieves.com</td>
<td>2019</td>
<td>70</td>
<td>2.61</td>
<td>.644</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.49</td>
<td>.820</td>
</tr>
<tr>
<td>District requirements</td>
<td>2019</td>
<td>70</td>
<td>2.89</td>
<td>.401</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.96</td>
<td>.200</td>
</tr>
<tr>
<td>LEAP 2025/Humanities</td>
<td>2019</td>
<td>70</td>
<td>2.69</td>
<td>.692</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.76</td>
<td>.662</td>
</tr>
<tr>
<td>LEAP 360</td>
<td>2019</td>
<td>70</td>
<td>1.21</td>
<td>1.190</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>1.27</td>
<td>1.169</td>
</tr>
<tr>
<td>District tests</td>
<td>2019</td>
<td>70</td>
<td>2.41</td>
<td>.925</td>
</tr>
</tbody>
</table>
### CURRICULUM-SPECIFIC READING TEST

<table>
<thead>
<tr>
<th>Factor</th>
<th>Year</th>
<th>N</th>
<th>M</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.18</td>
<td>1.149</td>
</tr>
<tr>
<td>ELA Guidebooks 2.0</td>
<td>2019</td>
<td>70</td>
<td>2.76</td>
<td>.600</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>2.94</td>
<td>.242</td>
</tr>
<tr>
<td>Teacher evaluation</td>
<td>2019</td>
<td>70</td>
<td>1.04</td>
<td>.984</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>.57</td>
<td>.866</td>
</tr>
<tr>
<td>Pre-service</td>
<td>2019</td>
<td>70</td>
<td>1.77</td>
<td>1.182</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>1.35</td>
<td>1.052</td>
</tr>
<tr>
<td>TL Summit</td>
<td>2019</td>
<td>70</td>
<td>.97</td>
<td>1.129</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>.49</td>
<td>.681</td>
</tr>
<tr>
<td>Prof. dev.</td>
<td>2019</td>
<td>70</td>
<td>1.94</td>
<td>1.048</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>1.43</td>
<td>1.041</td>
</tr>
<tr>
<td>Online network</td>
<td>2019</td>
<td>70</td>
<td>.67</td>
<td>.896</td>
</tr>
<tr>
<td></td>
<td>2021</td>
<td>49</td>
<td>1.31</td>
<td>1.245</td>
</tr>
</tbody>
</table>

Among highly influential factors, only the influence level of ELA Guidebooks 2.0 showed significant differences between the two years, $U(N_{2019}=70, N_{2021}=49) = 1520$, $z = -1.828$, $p = .068$. Effect size for this difference was small ($r = .168$). While the most influential factors remained highly influential over both years, Mann-Whitney tests indicated that four comparatively low-influence factors related to support services for teachers decreased significantly in influence between 2019-2020 and 2021-2022: teacher evaluation, $U(N_{2019}=70, N_{2021}=49) = 1347$, $z = -2.052$, $p = .040$; pre-service preparation, $U(N_{2019}=70, N_{2021}=49) = 1235$, $z = -2.787$, $p = .005$; Teacher Leader Summit, $U(N_{2019}=70, N_{2021}=49) = 1365$, $z = -2.075$.
p = .038; and professional development, $U(N_{2019} = 70, N_{2021} = 49) = 1245, z = -2.630, p = .009$.

The influence of online teacher networks on IADA-treated teachers increased significantly from 2019-2020 to 2021-2022, $U(N_{2019} = 70, N_{2021} = 49) = 1296, z = -2.621, p = .009$.

**Qualitative Influence Evidence**

Interviews yielded three emerging themes that provided both validation of quantitative findings and evidence of mediating factors related to quantitative findings. Promising evidence of LEAP test influence on IADA-treated teachers was validated by interviewed teachers and was related to the emergence of test form preparation as a theme. Compelling evidence of low LEAP 360 influence was validated by interviewed teachers and was related to the emergent theme of curriculum-embedded assessments. Last, evidence of diminished influence of several low-influence factors, as well as the increase in online teacher network influence, was validated and was related to the emergent theme of pandemic shutdowns. Evidence of these themes is presented below.

**Test Form Preparation.** Interviewed teachers and district officials frequently acknowledged the outsized influence the state ELA test has on teachers’ lessons and on the tasks they have students complete. “I won’t sugarcoat it,” said one teacher, “because singularly the most important thing is whatever test we're giving, because it is how I'm judged as a teacher. It's how your students are compared to other students. It's how your school is compared to other schools. Everything centers around whatever test we're giving now.” According to this teacher, the influence of the test was attributable to its power as an instrument of the state and of the district. “You know if your students aren't doing well on that test,” this teacher later said, “your administrators and school board officers are going to be knocking on your door going, ‘What's going on? Why are they not doing well? We need them to do better.’”
CURRICULUM-SPECIFIC READING TEST

The nature of the test, that teacher later said, was to transparently reveal student performance, in particular via their writing. The concreteness of seeing what skills and knowledge students displayed shaped his teaching more than did abstract standards describing skills and knowledge:

I mean, I don't I look at, ‘We didn't cover this standard in this unit, and we need to work on this standard.’ I look at, ‘Okay, they're not doing very well at, you know, writing their introduction or they're not pulling the right evidence or commentaries’…. And all that kind of stuff is what I look for and not really of, ‘Oh, we're not meeting this particular standard.’ Because at this point I know what the test is going to require them to do, and I know what day-to-day-life in the class is going to require them to be able to do. And that's what we focused on for our folks.

This teacher was clear that teaching Grade 7 ELA was largely centered on preparing students for a test. IADA tested students’ understanding of the substance of entire texts and their capacity to analyze and synthesize whole tests, the teacher said. If it did not test this knowledge and these skills, “I would definitely have to change [my teaching] because, you know, fortunately for me, my style and my opinion of what they should learn matches with what the test asked the lot of them.”

Teachers repeated this notion that the substance of daily lessons was affected by the specific knowledge and tasks required of IADA, in particular when compared with the conventional state ELA test, and that it would create differences in teaching methods, even among teachers using the same. “If you know you're given Innovative [tests],” said one teacher, “then you're probably following the curriculum more to a tee…. when I was doing LEAP, we would do a little bit more skills-based practice. Like maybe we'd spend a little more time, maybe
CURRICULUM-SPECIFIC READING TEST

at the end of the lesson or beginning a lesson working on just like some maybe some different
texts and working on theme or tone or mood, but you could kind of supplement and work on
more skills-based [activities], where this Innovative [test], I feel like we're more just pounding
the background knowledge, the background knowledge aspect.” This teacher’s example
provided a notable and clear example of how preparation for different tests engenders divergent
uses of lesson time, even when using the same curriculum. In the past, this teacher said, teachers
would use time outside of the curriculum to prepare students for questions they were likely to be
asked on the state test, related to specific reading skills, demonstrated via texts unrelated to the
curriculum. In the IADA-context, teachers spend that time “pounding the background
knowledge,” which in this district, as presented earlier in this chapter, includes using
thematically-grouped book-bins or bringing in additional content and texts that are specifically
and intentionally connected to the curricular unit. Another teacher captured this new,
curriculum-focused approach to test preparation under IADA: “I would say there's a very definite
shift in our approach to instruction since the Innovative [test]. I mean, we used to devote time to
test prep, and we would just do practice LEAP tests. We don't do that anymore. We devote our
time to diving into the unit and making sure that students have a strong understanding, as much
background knowledge as we can possibly give them.”

In particular, interviewed teachers emphasized moving away from two specific activities
related to test preparation. The first activity is repeated exposure to “cold reads,” textual
passages unrelated to the curriculum, followed by having students complete skills-based multi-
select test items. Teachers uniformly said that IADA participation diminishes the need or
incentive to expose students to cold reads. “Like at that point [in the school year close to the
testing window], some cold read passages and things like that we normally would do,” said one
teacher. “But when I have the Innovative [test] in play, it takes all of that out of the equation for me. Because my units are what I’m going to be tested on.” It is worth noting that the break from cold reads and related test preparation activities was not absolute. As noted in chapter 4, the IADA test form includes brief cold read passages and test items in response. Several interviewed teachers also noted that tests students will take in future years, including Grade 8 state ELA tests and later college admissions tests, are largely made up of cold reads and related test items. Teachers acknowledged these factors and expressed occasional ambivalence about not specifically preparing students for these realities. “I wish we did more multiple choice,” said one. “But at the same time I’m supposed to teach them the curriculum…. Innovative [test] has multiple choice as well. Guidebooks is lacking multiple choice.”

The second test preparation activity teachers eschewed was a teaching a style of writing that emphasized articulating opinions and identifying specific passages to support opinions. IADA essay tasks, teachers noted, require students to draw on knowledge they can recall, most often from outside the passage in front of them on the test form. Teachers repeated the notion that prior to IADA citing passages was part of a “formula” they routinely reinforced in teaching writing, including in students’ responses to Guidebook tasks. “In junior high,” said one teacher, “you have to teach procedures, like it's like formula for writing. And your better writers can go from there, but your weaker writers, they need like, you know, some formula to formalize that. [With IADA] we're just really throwing this formula for a loop.” The legacy writing formula, as it were, was tied to the requirements of specific tasks on the legacy state test and the specific rubrics via which teachers could assess the quality of a response. One interviewed teacher described an evolution, from embracing this formula and rubric to embracing IADA writing tasks:
We started it with the PARCC\textsuperscript{41} test. You know, that was kind of the original two-part question. ‘Answer this question and then whatever proves this question.’ And for a long time, I didn't like that. And I'm still kind of on the fence with it because I said, ‘You know, if a kid is not very good at pulling the right piece of evidence,’ he misses half of the question. So that is one of the things that I don't like. I did like the Innovative [test] versions because some of them had the part Bs and some of them didn't. Those few kids that are really not good at finding the right evidence, then they don't have to miss half of every question. But I think the difference is these [IADA writing] questions really end up judging the kids on ‘did you really understand?’ as opposed to the old school, you know, ‘who was the main character? What is the theme? And, you know. What happened after, you know, Susie met Joe?’ or whatever the old basic comprehension stuff. It's now having to really understand what you've read.

As with the move away from cold reads, teachers noted potential risks to students of shifting the writing style away from passages from a text on the form and toward recalled evidence from the curriculum. “So now we've shifted like seventh grade randomly,” said one interviewed teacher, “and so I feel like the teachers are probably shifting more, at least in seventh grade since it's [not a pilot]. They're probably shifting some of their writing in the Guidebooks by saying, don't look at your text and try to answer this exit ticket… But like these sixth graders, I mean, they didn't even do the [IADA] field test. They never even saw this [IADA] test. And then this year they hear, it's live for you and we're going to completely change your way of writing and test you. And it’s in October.” This teacher validated that the operational Grade 7 IADA had precisely the

\textsuperscript{41} The Partnership for Assessment of Readiness for College and Career (PARCC) test items have been used in Louisiana ELA tests since 2014.
effect policymakers intended; teachers taught students to build and recall knowledge rather than to hunt for passages. But in the absence of a multi-grade-level approach, across the entire system, the teacher said, students could experience incoherent teaching approaches as a result, especially when learning to write in response to texts.

Curriculum-embedded Assessments. Related to the coherence of students’ annual testing experience is the coherence of the tests administered by schools, districts, and states over the course of a given school year. As presented earlier in this chapter, on surveys, IADA-treated teachers were repeatedly and significantly less likely than non-treated teachers to report being influenced by LEAP 360, the state of Louisiana’s digital, no-stakes system for interim skills tests administered over the course of a school year, at the discretion of local school systems. This represents the most compelling finding in this study, backed by repeated, significant means differences. Interviewed teachers validated the finding and repeatedly explained that IADA has caused them to rely more on assessments embedded within the ELA Guidebook 2.0 curriculum than on externally developed skills tests. As with quantitative evidence, the qualitative evidence was clear and unambiguous. One interviewed teacher said, “I no longer use the LEAP practice test for my seventh graders, and I used to rely heavily on that for us to prepare for LEAP in the spring. I'm still pulling that for my eighth graders, but I no longer do that for my seventh.” Another teacher explained the phenomena not by disavowing test preparation but by explaining that the curriculum became her vehicle for test preparation:

“T'm a strong believer in students being familiar with the format of a test and how to attack a test. How to home in on what the question is really asking and how to narrow down their answer choices. I've seen the benefit of practicing that skill with them. But now what we do is we use our section quizzes from the Guidebook units, we use our
practice cold reads to practice those skills instead of pulling in all these practice LEAP test and just testing them to death. We build it into the unit with the unit-related stuff.”

Because the test rewards knowledge gleaned from the curriculum, there was less incentive for teachers to spend time on test preparation activities outside of the curriculum. “The best part is we don't have to think about the separate tests happening and preparing kids in a way for something that's foreign to them,” said one teacher. “I think there's a sense of comfort to teachers knowing that, you know, the curriculum I am actually working with right here, right now in my classroom is actually what my kids are going to be exposed to [on the test]. And we're going to get meaningful information from [IADA] throughout the year, I think makes you be able to invest so much more into that curriculum.” For IADA-treated teachers, curriculum and test reinforce one another. That alignment minimized the need or incentive to seek outside sources of test preparation information or activities. This is not to say that test-preparation activities did not occur in IADA-treated classrooms; it is instead to say that, to the extent they did happen, they were more likely to be a function of texts, tasks, and activities embedded within the curriculum.

**Pandemic Shutdowns.** As discussed in the executive summary, effects of the global COVID-19 pandemic are not a subject of this study. Neither survey respondents nor interviewees were specifically asked about the pandemic or the school shutdowns and online schooling that resulted. However, the issue arose in several interviews, in the context of factors that this study has shown to have had significant declines in their influence over IADA-treated teachers between 2019-2020 and 2021-2022. These declining influence factors are professional development, pre-service preparation, teacher evaluation systems, and the annual Teacher Leader Summit. Online teacher networks significantly increased its influence over IADA-treated
teachers. “I feel like the pandemic has had an impact on the frequency of PD,” said one teacher. “I know last year meeting virtually was not the same as meeting in person. Our district is innovative; we've scheduled PD throughout the year. But it's a different experience being in-person and actually being able to sit beside a peer and say, ‘Hey, we should annotate our guidance. Hey, I mean, just look at this assessment piece.’ And so I, I'm excited to get back to face-to-face.” Another teacher painted a more dire picture: “We have not had face-to-face PD. Only two days at the beginning of the school year for ELA. All PD now is Zoom-based. Teachers will turn off the screen. A lot of teachers are disinterested. They’ll just look at their phones.” These teachers’ experiences validate survey results showing changes in the influence of face-to-face services for IADA-treated teachers. While an important finding on its own, this change in influence may provide context for other changes seen in survey results. This chapter turns now to a discussion that synthesizes these and other findings.

Discussion

This study provides an initial analysis of implementation and outcomes of a statewide policy implementation. The study has addressed the most stubborn of U.S. education challenges via the largest-scale sort of education intervention, and it has presented evidence that a conceptually simple shift in the nature of state reading tests—from curriculum-agnostic tests to curriculum-specific tests—is associated with promising shifts in teacher practice at large scale and on a rapid timeline. The study indicates the importance of further IADA research, and it should further foster interest among state policymakers in the design of the ineffectual reading tests that persist in their states, as well as in a new alternative that is grounded in longstanding evidence of how children actually learn to comprehend texts.
CURRICULUM-SPECIFIC READING TEST

This section first summarizes outcomes. It then synthesizes them, in order to draw connections among results and to highlight particular insights they may provide. Figure 5.1 summarizes all significant quantitative results and all emergent qualitative codes. To distinguish between compelling evidence and promising evidence, I underline evidence related to the one compelling finding in the study.

*Figure 5.1 Quantitative and qualitative results summary*

<table>
<thead>
<tr>
<th>Research question</th>
<th>2019-2020 differences</th>
<th>2021-2022 differences</th>
<th>2-year change (IADA-treated)</th>
<th>Emergent codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R3: Text usage</td>
<td></td>
<td>Guidebook weekly usage (IADA-treated more likely)</td>
<td></td>
<td>Pacing urgency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Guidebook text use in lesson (IADA-treated more likely)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ4: Intent</td>
<td>Intent to have students articulate opinions supported by evidence (non-treated more likely)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ5: Activities</td>
<td>Other texts w/ similar themes (IADA-treated more likely)⁴²</td>
<td>Other subject area (IADA-treated more likely)</td>
<td>Other subject area (less likely in 21-22)</td>
<td>Like texts</td>
</tr>
<tr>
<td></td>
<td>Other subject area (IADA-treated more likely)</td>
<td>Grade-level texts (IADA-treated more likely)</td>
<td>Grade-level texts (more likely in 21-22)</td>
<td>Social studies changes</td>
</tr>
<tr>
<td></td>
<td>Grade-level texts (more likely in 21-22)</td>
<td>Analyze text structure (less likely in 21-22)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⁴² Other texts showed non-significant means differences in 2021-2022 ($MD=0.12, p=0.217$) and validating qualitative evidence.
### RQ6: Influence

<table>
<thead>
<tr>
<th>LEAP 360 (IADA-treated less influenced)</th>
<th>LEAP 360 (IADA-treated less influenced)</th>
<th>Curriculum-embedded assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online networks (IADA-treated less influenced)</td>
<td>Online networks (IADA-treated more influenced)</td>
<td>Online networks (more influence in 21-22)</td>
</tr>
<tr>
<td>Teacher eval (IADA-treated less influenced)</td>
<td>Teacher eval (less influence in 21-22)</td>
<td>Test form preparation</td>
</tr>
<tr>
<td>State web site (IADA-more influenced)</td>
<td>LEAP tests (IADA-treated more influenced)</td>
<td>Preparation (less influence in 21-22)</td>
</tr>
<tr>
<td></td>
<td>Guidebooks (IADA-treated more influenced)</td>
<td>Professional; development (less influence in 21-22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TL Summit (less influence in 21-22)</td>
</tr>
</tbody>
</table>

Results included one compelling correlation, demonstrated via repeated significant differences, and validated by qualitative results, between IADA participation and lower levels of LEAP 360 test influence on teachers’ decisions. Other constructs, including fidelity of Guidebook usage,

---

43 Repeated significant quantitative findings with validating qualitative evidence

44 LEAP test influence showed notable means differences in 2019-2020 ($p=.150$) and validating qualitative evidence.
usage of thematically similar texts, and influence of LEAP tests and Guidebooks on teachers’ decisions, indicated promising but not repeated results that were supported by qualitative evidence. Several significant differences in IADA-treated group responses from 2019-2020 to 2021-2022 were also supported by qualitative results, including changes in the use of content from other subject matters and changes in the influence of face-to-face support services related to the COVID-19 pandemic.

Results Synthesis

To describe the problem U.S. adolescent reading in chapter 1, I used the theoretical framework of a multi-layered ecological system. Using this framework allowed me to explore the problem at multiple levels: the issue’s macrosystem, which is its national and global context; its exosystem, the network of events and institutions influencing its actors; and its microsystem, which includes relationships and interactions between parent or teacher and child. The study that followed examined the IADA intervention at all three levels: the intervention’s implementation by a state government, the influence of large-scale policies and instruments on teaching methods, and the principles and methods teachers actually adopted in the classroom. In discussing findings, I return to the ecological systems framework to synthesize the evidence presented, distinguishing among implications at different levels of the education system.

Microsystem: Teacher Principles and Methods

This study has presented promising, and in one instance compelling, evidence of correlations between IADA participation and adoption of classroom methods that embody and reinforce a knowledge-building reading curriculum. This alignment of curriculum design and classroom method was most evident in IADA-treated teachers’ reliance on curriculum-embedded tests rather than interim tests provided outside of the curriculum. This was a repeated
quantitative finding, reinforced by teachers repeatedly detailing their reliance on tests from within the curriculum. While it was less evident in quantitative data, there were promising indications that IADA-participating teachers reinforced units’ knowledge domains, using additional texts and additional time. There were also promising indications that IADA-participating teachers used the Guidebooks curriculum for greater portions of their lessons, proceeding with urgency through the curriculum because the test requires of students a strong base of knowledge than can best be achieved. While only the finding on the influence of interim tests outside the curriculum yields compelling evidence of a correlation, the aforementioned factors combine to present a picture of focus and efficiency in the teacher that researchers should consider in studying Louisiana’s IADA model further.

A decade into the state’s push to advance high-quality curricular materials in public schools, while Louisiana ELA teachers are highly likely to use such materials (Doan et al., 2022; Opfer et al., 2016), it remains the case that most Louisiana ELA teachers do not believe that the curricular materials they use are sufficient to prepare students for the state ELA test (Doan et al, 2022). The implication is that most Louisiana ELA teachers remain under the influence of multiple, often conflicting, sources of guidance, each with its own demands on time in the classroom. There is some evidence in this study that IADA participation reduces that misalignment, as teachers indicate significantly less concern with tests that don’t embody the curriculum and some evidence of more frequent use of the curriculum itself. Greater alignment between curriculum and test could mean more time to focus on a singular task, guided by the curriculum, building coherent bodies of knowledge to an extent not possible when pulled in the direction of multiple, powerful influences. In such a classroom, the teacher’s job is to do the work of the curriculum, and the curriculum and teacher share the burden of producing learning
given limited time with students, a complex interplay that should be accounted for in the design of future quantitative evaluations of teacher efficacy or curriculum efficacy.\textsuperscript{45}

It should be noted that IADA-treated teachers indicated no particular principle or belief that distinguished them from other users of ELA Guidebooks 2.0. The IADA test’s potential or compelling influence on teachers was evident, but it was technical rather than philosophical. Teachers viewed the test as a source of direction when they eschewed external interim tests or when they added related texts, for example, and they followed those directions responsively. This responsiveness may satisfy the ambitions of policymakers and advocates seeking particular changes to classroom practice, but it may also imply a risk of repeating the kind of mechanical response to tests found in past, NCLB-era testing regimes. Teachers’ focus on building knowledge and fostering textual meaning may become fetishized in the manner of skills-based test preparation for cold reads, if teachers lack the perspective and their own background understanding of the role of knowledge and its relationship to issues of equity and fairness. Some interviewed teachers, it is worth noting, emphasized that their ability to achieve pacing urgency or to use outside texts and subject matter was related to the extent to which they were teaching student sections designated as gifted or honors. While not an outcome of this study, this observation reflects the potential peril of forcefully emphasizing a particular dimension of reading instruction in state policy without grounding teachers in a belief system that fosters responsiveness to the needs of unique students. This is all the more important given the weak influence treated teachers attributed to pre-service and in-service preparation activities in this study. If teachers are to forge principles for why they teach texts, the role of knowledge, and

\textsuperscript{45} Blazar et al. (2019) do not account for testing conditions in concluding that there is no distinction among curricular products, for example.
how these priorities impact struggling readers, they may require systems of support beyond those they are currently experiencing, lest the IADA test become a distant, technical mandate more than a reinforcing prompt to professional and intellectual growth.

**Exosystem: Influences on Teacher Principles and Methods**

The Louisiana IADA intervention was, essentially, an effort to conjoin two frequently disconnected sources of influence over teacher practice and student learning: the curriculum and the test. IADA showed promising if not fully compelling signs of being distinctively influential. In 2021-2022 IADA-participating teachers reported significantly greater influence of the substance of state tests, for example, than did non-treated peers. In the first year of the study, they reported significantly higher rates of using other, related texts in the classroom, which interviewed teachers attributed to the influence of the test design. In the second year of the study, they reported greater rates than their peers of using the Guidebooks curriculum in class, which interviewed teachers also attributed to the influence of the test’s requirements and the urgent pacing that they pursued as a result.

However, much of this chapter has presented results that show little distinction in principles and methods between IADA-treated teachers and other teachers using Guidebooks. There was, for example, no distinction in teachers’ beliefs about the goal of teaching texts. While there were occasional differences in methods and curriculum fidelity, there was only one instance of compelling evidence demonstrating a correlation. Even that instance of compelling evidence, of a relationship between IADA participation and limited influence of interim tests provided outside the curriculum, spoke to the potentially powerful influence of the curriculum as much as to the test on its own. Given these results, it is plausible to ask whether the Guidebooks curriculum and how it was implemented in Louisiana has been a definitive influence over how
teachers approach their work, to the exclusion of other sources of influence. One way to think of the IADA influence could be, therefore, less as a direct source of influence and more as an amplifier to the already present influence of the curriculum.

This tight alignment among elements of systems that have such influence places a particular burden on policymaking bodies to design and communicate a coherent vision, and to understand the implications of even modest shifts in design. As discussed earlier in this chapter, shifts in Louisiana social studies standards and testing policies disrupted structural reforms at the school level to conjoin social studies and ELA instruction. This risked both confusion among practitioners and a threat to knowledge-building strategies participating schools and districts had developed. Perhaps relatedly, results indicated that teachers in the second year of the study were significantly less likely to bring other subject matter into the ELA classroom.

Other state-level policies were present in the context of IADA implementation as well. Teachers were aware that IADA-like, curriculum-specific tests in other grade levels were only pilot efforts, and that accountability results in those grade levels still were a function of traditional LEAP tests largely made up of cold reads and skills-based prompts. The LDOE ceased in 2021-2022 to provide statewide professional development specific to the IADA test, for all participating schools and teachers. Perhaps relatedly, treated teachers reported greatly diminished influence of professional development generally in 2021-2022, just as they reported significantly lower levels of employing evidence-based methods in that year, such as analyzing textual structure. These large-scale shifts in social studies policy, accountability policy, and professional development implementation, all highlight the intricacy of state efforts to exert influence over teacher principles and methods. The coin of state influence has two sides: one an unmatched opportunity for coherence and focus, the other a long tail of confusion in light of only
CURRICULUM-SPECIFIC READING TEST

partially built systems. Further documentation and study of IADA should examine this
understudied issue of state policy systems design, its complexity, and the time and structures
necessary to achieve coherence amidst tightly connected policies.

Macrosystem: U.S. and Louisiana Context

In the backdrop of this study is the context of a global health pandemic, resulting school
shutdowns across the U.S., and a climate of political and social unrest. This study was
suspended mid-year in 2019-2020 and then delayed for a second year in 2020-2021. In that time,
state tests were largely suspended in the U.S. Teachers communicated with students online for
months at a time. The nation’s history and literature became a regularly debated political issue.

There are outcomes in this study that relate specifically to these disruptions. Most
notably, the influence of face-to-face support systems such as professional development, teacher
evaluation, and educator preparation, plummeted in the second year of the study. The influence
of online teacher networks concurrently increased. It is not hard to see in these results the shift
in how teachers interacted and were supported during this tumultuous time.

Other social factors elude this study but are worth noting for their importance in this time.
How the nationwide suspension of testing and accountability policies impacted schools and
teachers, for example, is beyond the capacities of this study but must be examined. Similarly,
teachers’ and students’ experiences with digitally delivered curricula and other instructional
materials, such as ELA Guidebooks 2.0, will be an important area of study in future years, as
will the pandemic’s undoubtedly sweeping effects on children and families. Much changed in
American schools over the course of this study, much of it having to do with curriculum, testing,
and cultural knowledge; a robust research response is essential.
Implications

While standardized tests are frequently the source of outcomes data for quantitative research, literature examining the effects of particular large-scale standardized test systems is notably sparse. Perhaps because so little has changed in state testing regimes since the implementation of NCLB, research has tended to focus on the consequences of status quo tests (Au, 2007; Hirsch, 2016; Moon, Brighton, Jarvis, & Hall, 2007; Musoleno & White, 2010), rather than on the effects of shifts in testing policy or testing forms. This dissertation has prominently featured studies of NCLB testing and accountability policy (Dee & Jacob, 2011; Lee & Reeves, 2010), an important study of IB curriculum and test effects on postsecondary outcomes (Conley et al., 2014), and a paper synthesizing literature on international assessment models (Steiner, 2018). The creation of IADA models, in Louisiana and elsewhere, creates a valuable opportunity for researchers to study the inputs, outputs, and outcomes associated with a notable shift in state testing approach. The results of this study point researchers toward several compelling or promising avenues for future study that should be pursued as a rare opportunity to evaluate innovation in state testing.

As discussed in chapter 1, state testing systems are largely a creation of federal laws. The greatest opportunity the U.S. system has for improvement in the design of its tests lies in the periodic reauthorizing of ESEA, which is presently up for reauthorization. IADA was a product of the current version of that law, ESSA, and nascent results such as those in this study should inform a dialogue among policymakers about the substance and design of state assessments, as stipulated by future versions of the federal law. Louisiana’s IADA creates the hope for a model that has substantive and helpful impacts for teachers and students alike. But changes to state testing series are not without costs and not without the need for thoughtful federal policy.
Leaders at the USDOE and on Capitol Hill should consider further investment in stimulating and studying new assessment models and should consider how federal policy could steer states toward thoughtful, evidence-based approaches to assessment system design.

Perhaps most critically, state education agency and state board of education leaders, who are ultimately responsible for assessment system design, should consider this study both for what it says about the merits of Louisiana’s approach and for the case study it creates in attempts at coherent assessment system design. In the wake of the pandemic, states have been called on to lead academic recovery efforts at great scale. Their options for academic influence at statewide scale are limited. Coherent approaches necessitate substantive and communications alignment across these elements, a challenge to agency leaders known for short tenures in politicized contexts.

**Limitations to Findings**

As discussed in chapter 4, there are limits to the design of this study that should give readers caution in generalizing its findings. There is a prominent risk of selection bias in the sample (Shadish, Cook, & Campbell, 2002), this being a study of select districts participating in a voluntary reform in a single state. Furthermore, data collected are entirely self-reported, which risks validity (Shadish, Cook, & Campbell, 2002). I addressed these limitations by not exclusively relying on one method of data collection (Creswell & Plano-Clark, 2018) and by requiring that results repeat over two years in order to establish a compelling correlation. But the findings should be evaluated with the self-reporting limitation in mind.

The role of school district influence, notably the highest source of influence of teacher decisions according to both treated and non-treated survey respondents in both years, presents another challenge to internal validity. The study does not account for the specific influences of
district curriculum implementation. That every district represented in the survey has used ELA Guidebooks 2.0 for at least one prior school year provides some level of comparability among respondents, but future researchers should consider identifying and controlling for district implementation factors in order to limit the contamination from unseen effects of varying district approaches.

As noted previously, both the initial IADA administration and this study were disrupted by the 2020 suspension of in-person schooling and state testing. While some likely effects of the pandemic were discernible in the differences between teacher influence factors from year to year, many effects of the pandemic disruption likely influenced the principles and methods of teachers in in ways the study cannot fully discern. Relatedly, as discussed in chapter 4, there are differences in the 2019-2020 and 2021-2022 Grade 7 IADA implementations: modest changes to the list of participating districts, changes in the LDOE approach to professional development, ambiguity on school accountability policy, and expansions of the concept to other grade levels and curricula. These are additional sources of caution that require that the study treat the two years as distinct implementations, a limit to the generalizability of any one-year result. I have addressed this limitation by not referring to outcomes in any single year as compelling evidence of a correlation throughout the study.

Last, as noted earlier, this study is part of a voluminous literature on the effects of policies and tests on teachers. It does not ask or answer the paramount question of how the IADA reform affected students, or whether IADA achieved its goal of better educating young readers. As IADA-participating students progress in their educations, and as the population of students participating in curriculum-specific ELA tests in Louisiana grows, it is essential that
CURRICULUM-SPECIFIC READING TEST

effects on reading skills and student experiences be researched. This study asks important questions as a starting point, but it speaks only to the initiative’s means and not its truest ends.

Conclusion

This study analyzed the problem of stagnant adolescent reading levels in the U.S. at multiple levels of society and the education system, establishing that the nation’s policy response, standards-based reform, had failed to achieve a coherent, evidence-based approach to teaching reading to middle grade students. Critically, the skills-based tasks required of students on state reading tests have become not just a feature of measure instruments but instruction to teachers, many of whom have made teaching these discrete, isolated skills a daily feature of the American classroom. This practice contrasts with evidence that teaching students to comprehend texts should be rooted in an effort to make sense of a whole text; that by making sense of a text, the reader builds the base of knowledge he possesses; and that this newly possessed background knowledge opens doors for comprehending additional texts of even greater complexity.

Louisiana embraced this approach by supporting the curricular products that embrace this centrality of knowledge and that feature activities via which to develop knowledge. The IADA test Louisiana developed with its partners sought to enhance the pedagogical effects of one such curricula and to end deleterious practices encouraged by conventional state ELA tests. This study sought to discern correlations between IADA participation and teachers approaches to teaching texts by collecting survey data over two years and by validating quantitative outcomes with qualitative data collected in interviews.

The study found that IADA participation is compellingly correlated with diminished influence of interim assessments provided outside the curriculum; participating teachers were more likely than even peers using the same curriculum to rely on curriculum-embedded
assessments instead of externally provided tests. Evidence indicated promising but not yet compelling evidence in other factors, such as participating teachers’ fidelity to the curriculum, adoption of methods important to an integrated approach to reading comprehension, and adherence to the test design and its requirement that students possess knowledge of what they have read.

The study took place during a tumultuous period in which schools were closed for extended periods of time due to the COVID-19 pandemic, and during which U.S. politics were taken up with debates on race and teaching history. Results indicated likely effects related to these issues, including the diminished influence of face-to-face support systems for participating teachers over the course of the pandemic and the diminished role of social studies content for participating teachers given nationally covered changes to Louisiana’s social studies standards.
References


Chambers, R. (2013). The common core text exemplars – a worthy canon or not? *Voices from
the Middle, 21(1), 48-52. Retrieved from https://library.ncte.org/journals/vm/issues/v21-1


CURRICULUM-SPECIFIC READING TEST


CURRICULUM-SPECIFIC READING TEST


Del Rio, G.M.N. (2021, May 15). The University of California will no longer consider ACT and


CURRICULUM-SPECIFIC READING TEST

98(2), 382-393. doi:10.1037/0022-0663.98.2.382


CURRICULUM-SPECIFIC READING TEST

404. doi:10.1086/669852


doi:10.1162/DAED_a_00078


doi:10.1037/0012-1649.43.3.620


204
CURRICULUM-SPECIFIC READING TEST

Research, Inc.

Joftus, S. (2002). Every child a graduate: A framework for an excellent education for all middle
and high school students. Washington, DC: Alliance for Excellent Education. Retrieved
from https://all4ed.org/reports-factsheets/every-child-a-graduate-a-framework-for-an-
excellent-education-for-all-middle-and-high-school-students/

Kamil, M.L. (2012). Reading preparedness for college and technical professions. Washington,
DC: National Assessment Governing Board. Retrieved from
https://www.nagb.gov/content/nagb/assets/documents/what-we-do/preparedness-
research/judgmental-standard-setting-studies/paper-the-standard-for-minimal-academic-
preparedness-in-reading-to-enter-a-job-training-program.pdf

Improving adolescent literacy: Effective classroom and intervention practices: A practice
and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

https://www.brookings.edu/research/never-judge-a-book-by-its-cover-use-student-
achievement-instead/

Kane, T., Owens, A., Marinell, W., Thal, D., & Staiger, D. (2016). Teaching higher: Educators’
perspectives on common core implementation. Cambridge, MA: Center for Education
report.pdf


Louisiana Department of Education. (n.d.). Common core state standards to Louisiana state standards crosswalk. Retrieved September 1, 2022 from
CURRICULUM-SPECIFIC READING TEST

https://www.louisianabelieves.com/docs/default-source/academic-standards/ela-ccss-gle-crosswalk


Louisiana Department of Education. (2020). Louisiana receives federal waiver for required
standarized tests and accountability systems for 2019-2020 school year. Retrieved from
https://www.louisianabelieves.com/newsroom/news-releases/2020/03/20/louisiana-
secures-federal-waiver-for-required-standardized-tests-and-accountability-systems-for-
2019-2020-school-
year#:~:text=This%20means%20Louisiana%20will%20not,for%20the%20current%20school%20year

statewide-educational-technology-plan.pdf?sfvrsn=3e999e1f_6


operational.pdf?sfvrsn=d7fd6418_2


Statistics for the Biological Sciences Using R. Springer, Cham. Retrieved from https://doi.org/10.1007/978-3-319-30634-6_4


National Center for Education Statistics (2017). *Public school expenditures*. Retrieved April 15,
CURRICULUM-SPECIFIC READING TEST


National Center for Education Statistics (n.d.) Table 204.10 Number and percentage of students eligible for free and reduced-priced lunch, by state. Retrieved from https://nces.ed.gov/programs/digest/d20/tables/dt20_204.10.asp


CURRICULUM-SPECIFIC READING TEST


Radunzel, J. & Noble, J. (2012). Predicting long-term college success through degree completion
CURRICULUM-SPECIFIC READING TEST

using ACT composite score, ACT benchmarks, and high school grade point average.


Slavin, R.E., Lake, C., Cheung, A., & Davis, S. (2009). Beyond the basics: Effective reading programs for the upper elementary grades. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education. Retrieved from


Curriculum-Specific Reading Test

21. doi:10.1111/1467-9817.12042


https://nepc.colorado.edu/thinktank/review-evaluation-tfa-texas


doi:10.1177/1086296X16632329
Appendix A

Needs Assessment Survey Form

As you likely know, your school district is participating in an Innovative Assessment Pilot initiative, trying out a new form of standardized test to determine whether the test assists teachers in improving their support of students’ reading skills. The questions in this survey are designed to help researchers and the state Department of Education understand what materials and activities teachers of English language arts (ELA) in districts piloting the tests are using in their classrooms. The survey also is designed to help researchers understand how teachers get information to determine which materials and activities they should use in the classroom.

The survey is designed to take 20 minutes of your time.

As a participant, you remain anonymous. Your name and your individual responses will not be shared or published in any manner. This survey is also not an evaluation of your knowledge or performance.

Thank you again for completing this survey and for your work to support Louisiana’s children.

Question 1: Identify the frequency with which you use the following types of instructional materials, such as curriculum or classroom handouts, for your ELA classroom lessons this year [Options: Never use and/or never heard of; Rarely (1x per month or less); Occasionally (2–3x per month); Often (1–2x per week); Daily or almost daily (3–5x per week)]

- Materials developed and/or selected by my district (including textbooks, novels, or any other materials the district circulates to teachers across the district for use in classroom lessons)
- Materials I developed and/or selected myself (including materials developed in collaboration with other teachers but not formally circulated to teachers across the district for use in classroom lessons)
- Leveled readers/texts (different reading materials for students at different reading levels)
- Trade Books
- EngageNY materials
- Accelerated Reader (Renaissance Learning)
- Book It! Program
- Core Knowledge Language Arts/CKLA (Amplify)
- Developing Core Proficiencies (Odell Education)
- ELA Guidebooks (LearnZillion)
- Expeditionary Learning
- Great Source (Houghton Mifflin Harcourt)
- Houghton Mifflin Reading (Houghton Mifflin Harcourt)
- Harcourt Reading (Houghton Mifflin Harcourt)
- Journeys (Houghton Mifflin Harcourt)
- Literacy by Design (Houghton Mifflin Harcourt)
CURRICULUM-SPECIFIC READING TEST

- Literature or Elements of Literature (Holt McDougal)
- MacMillan (McGraw-Hill)
- Making Meaning (Developmental Studies Center)
- McDougal Littell Language of Literature (Houghton Mifflin Harcourt)
- RAZ-Kids (Learning A–Z)
- Read 180 (Scholastic)
- Reading A–Z (Learning A–Z)
- Reading Street (Pearson Scott Foresman)
- Reading Wonders (McGraw-Hill)
- SpringBoard English Language Arts (College Board)
- SRA Reading or Open Court Reading (McGraw-Hill)
- Storytown (Harcourt)
- Prentice Hall Literature (Pearson Prentice Hall)
- Text Talk/Direct Instruction (Scholastic)
- Treasures (Macmillan McGraw-Hill)
- Trophies (Harcourt)

Question 2: Roughly how many hours do you spend on your own developing and/or selecting instructional materials, such as handouts, books, or other texts, in a typical week?

- Less than 1
- 1–2
- 3–4
- 5–6
- 7–8
- 9–10
- More than 10

Question 3: Which one of the two statements below better describes your approach to teaching students to comprehend what they read?

- Teach particular texts (novels, poems, articles) and then organize instruction around helping students to comprehend and analyze those texts
- Teach particular skills and strategies (locating the main idea, describing author’s purpose), then organize instruction around those skills and strategies

Question 4: How often do lessons in your classroom that involve reading a text include the following activities? [Options: Never/Rarely/Occasionally/Often/Always]

- Discussing texts in small-group or whole-class settings
- Locating the main idea of a text or summarizing the text
- Learning about historical context or other knowledge relevant to the text
- Developing persuasive arguments about the text
- Responding in writing to questions that are similar to questions on the LEAP test
• Learning or using vocabulary used in the text

Question 5: Please indicate the extent to which each of the following factors influence which ELA materials, such as texts or classroom handouts, you typically use in your classroom. If a factor is not applicable, please choose “not at all.” [Options: Not at all; Minimally; Somewhat; Significantly; Very significantly]

• Louisiana Student Standards for ELA
• Louisianabelieves.com
• School or district requirements
• End-of-year LEAP tests
• District-wide or school-wide assessments
• LEAP 360 assessments
• Compass, TAP, CLASS, or other observation rubric
• Teacher Leader Summit
• National Council of Teachers of English Language Arts and Reading Education Standards
• Online teacher networks, blogs, or forums
• Pre-service preparation (college of education or alternative certification programs)
• Professional development experience in the last year
• Other (please describe)
CURRICULUM-SPECIFIC READING TEST

Appendix B

Survey Form

1. I have served as a K-12 teacher in public or private schools for
   • 0-3 years
   • 3-10 years
   • More than 10 years

   The school at which I work is named: _________________________

2. During the current semester, I am teaching the following number of sections of grade 7 English language arts (ELA):
   • 0
   • 1
   • 2
   • 3
   • 4
   • 5
   • More than 5

3. A length of a typical ELA class period in my school is
   • 30-40 minutes
   • 41-60 minutes
   • 61-90 minutes
   • More than 90 minutes

4. In a typical week of teaching, how closely does your instruction of unit follow Guidebooks lessons and activities? (Source: NWEA pilot teacher survey)
   • Never or hardly ever: I do not use the Guidebook lessons and activities.
   • Occasionally: I teach some Guidebook lessons and activities in a given unit.
   • Often: I teach most Guidebook lessons and activities in a given unit.
   • Almost always: I teach all Guidebook lessons and activities in a given unit.

5. In a typical lesson in your classroom, for roughly what percentage of the lesson do students actively engage with a Guidebook text (reading it, writing about it, or discussing it) rather than doing another activity (learning background knowledge before reading a text, responding to cold reads, practicing particular reading skills, e.g.)? (Source: self-made)
   • Never or hardly ever: I never or almost never use Guidebook texts in class.
   • Occasionally: students spend approximately a third of their time engaging with a Guidebook text.
CURRICULUM-SPECIFIC READING TEST

- Often: students spend between one third and two thirds of the time in a typical lesson on Guidebook texts (up to 30 minutes in a 45-minute lesson).
- Almost always: more than two thirds of the time in a typical lesson.

6. Which one of the two statements below better describes your approach to teaching classes of 7th grade students to comprehend what they read? (Source: ATP)

- I prefer to teach particular texts (novels, poems, articles), and then organize instruction around helping students to comprehend and analyze those texts.
- I prefer to teach particular skills and strategies (locating the main idea, describing author’s purpose), and then organize instruction around those skills and strategies.

7. Please select all of the statements below that accurately describe your view of what it means for a student to understand the meaning of a text in your classroom. You may select one statement or more, in order to reflect your view of what it means to understand a text. If none of these statements accurately describe your view, please select only the last option. (Self-made)

- None of these statements describe what it means for a student to understand a text.
- The student can accurately paraphrase what the text is saying in the student’s own words.
- The student can articulate opinions about what the text is saying and support those opinions with evidence from the text.
- The student forms a hypothesis about what a text means by analyzing figurative language or the text’s structure.
- The student explains the text’s importance by describing its relationship to other texts, to works of art, or to history.

8. Taking into account all the reading students do in a typical class period, what percent of the time does your class of students read: (Source: ATP)

- The same text written at the grade level you teach?
  - Never or hardly ever (0–25%)
  - Occasionally (25-50%)
  - Often (50-75%)
  - Always or almost always (75-100%)
- Different texts, depending on students’ individual reading levels?
  - Never or hardly ever (0–25%)
  - Occasionally (25-50%)
  - Often (50-75%)
  - Always or almost always (75-100%)
9. How often do you ask your students to engage in the following practices during class? [Options: Never or hardly ever; Occasionally (1–3 times per month); Often (1–3 times per week); Daily or almost daily] (Source: ATP)

- Use evidence from a text to support opinions drawn based on passages within the text
- Analyze the structure of texts, including how specific sentences, paragraphs and larger portions of text relate to each other and the whole
- Connect literacy instruction to other content (e.g., science, social studies)
- Analyze how two or more texts address similar themes

10. Please indicate the extent to which each of the following factors influences your use of texts and the activities you choose for lessons. If a factor is not applicable, please choose “not at all.” [Options: Not at all; To a slight extent; Somewhat; A great deal] (Source: ATP)

- Louisiana Student Standards for ELA
- Louisianabelieves.com
- School or district requirements
- LEAP 2025 or LEAP Humanities
- LEAP 360
- District or school ELA assessments
- ELA Curriculum Guidebooks 2,0
- Preservice preparation in a college of education or an alternative certification program
- Compass, TAP, CLASS, or another system of observation and feedback
- Teacher Leader Summit
- Professional development experiences in the last year
- Online teacher networks, blogs, or forums
- Other (Please specify).
Appendix C

Interview Protocol

Thank you for taking the time out of your busy schedule to do this interview. As you may know, the state of Louisiana has implemented an ELA curriculum statewide called ELA Guidebooks 2.0. In some parishes, the state has also worked with local leadership to implement a new Innovative assessment that is based on the Guidebooks curriculum. In order to understand the effects of curriculum and tests on how teachers think about their work, the Louisiana Department of Education recently asked teachers in this school and others to complete short surveys related to their practice. If you completed that survey, thank you for doing that.

The purpose of today’s interview is to understand issues covered on that survey in greater depth than a survey allows. Hearing your explanations will provide additional information that may provide insights for researchers they would not have had simply looking at the survey results. Your answers might also raise questions that the researchers wish to explore in different studies.

The interview is being recorded in order to assure accuracy. Quotes from the interview may be used in a study, but neither your name nor anything about your individual identity will be attached to any quote or opinion in the study. Your identity will remain completely confidential and will not be shared or published.

I expect the interview will take about 20 minutes. Before we start, do you have any questions? Great, let’s begin.

1. Your school uses ELA Guidebooks 2.0. Generally speaking, how closely do you follow Guidebook lessons and activities in your classroom?

2. Speaking of the Guidebook texts themselves, for how much of your lessons are students typically working with the text itself – writing about it, reading it, or talking about it?

3. When you have students working with the text, what does success mean for you? Is it that students can tell you what’s happening in the text? Is it that they are demonstrating skills like being able to summarize or identify the main idea? Or is it something other than this?

4. You obviously make a lot of decisions related to what gets taught and how it gets taught. When you think about the influences on the decisions you make, what factors rank high in your mind? And what makes you say that? Please focus particularly on academic factors such as standards, assessments, curriculum, professional development, pre-service preparation, and so on.

5. Regarding the influence of the Innovative assessment test on your decisions, does it play a big role? Why or why not?
6. Take me into your thought process about how the Innovative assessment test influences your teaching decisions. What elements or characteristics of the Innovative assessment tests show up in your teaching or the assignments your students complete, if any?

7. Is there anything else you would like to add about the factors that influence your teaching approach? Anything else about the Innovative assessment?
Appendix D

Interview Consent Form

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Informed Consent Form

Title: Effects of Text-specific Reading Assessment on Teacher Belief and Practice

Principal Investigator: Dr. David Steiner, Johns Hopkins University

Date: <<INSERT>>

PURPOSE OF RESEARCH STUDY:

- The purpose of the study is to identify the effects of using a state reading test that measures students’ knowledge of specific texts, like the new assessment pilot in Louisiana, on the beliefs and actions of teachers. The study also seeks to identify factors that may explain why and how tests affect the beliefs and actions of teachers.

PROCEDURES:

- We are asking if you would participate in a 20-minute interview related to your practices and beliefs in the classroom.

RISKS/DISCOMFORTS:

- The risks associated with participation in this study are no greater than those encountered in daily life.

BENEFITS:

- There are no direct benefits to you from participating in this interview (financial benefits, for example).
- However, this study may benefit teachers and students if the results lead to a better understanding of how a test that measures students’ knowledge of actual books students have read impacts the way teachers teach.
CURRICULUM-SPECIFIC READING TEST

VOLUNTARY PARTICIPATION AND RIGHT TO WITHDRAW:

- Your participation in this study is entirely voluntary: You choose whether to participate. If you decide not to participate, there are no penalties, and you will not lose any benefits to which you would otherwise be entitled.
- If you choose to participate in the study, you can stop your participation at any time, without any penalty or loss of benefits. If you want to withdraw from the study, please contact the state’s institutional review board at laura.boudreaux@la.gov.
- If we learn any new information during the study that could affect whether you want to continue participating, we will discuss this information with you.

CONFIDENTIALITY:

- Any study records will be kept confidential.
- To protect your confidentiality, the researchers will use only your first and last initials to identify your records, rather than your full name.
- Any records that identify you, such as this form or interview notes, will only be available only to people working on the study.

COMPENSATION:

- You will not receive any payment or other compensation for participating in this study.

IF YOU HAVE QUESTIONS OR CONCERNS:

- You can ask questions about this research study now or at any time during the study, by talking to the researcher(s) working with you or by emailing Dr. David Steiner (d.steiner@jhu.edu).
- If you have questions about your rights as a research participant or feel that you have not been treated fairly, please call the Homewood Institutional Review Board at Johns Hopkins University.

SIGNATURES

WHAT YOUR SIGNATURE MEANS:

- Your signature below means that you understand the information in this consent form. Your signature also means that you agree to participate in the study.
- By signing this consent form, you have not waived any legal rights.
<table>
<thead>
<tr>
<th>Participant's Signature</th>
<th>Date</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Signature of Person Obtaining Consent (Investigator or HIRB Approved Designee)</th>
<th>Date</th>
</tr>
</thead>
</table>
## Appendix E

### Qualitative Codebook

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXT USAGE</td>
<td>GUIDEBOOKS</td>
</tr>
<tr>
<td></td>
<td>GRADE LEVEL</td>
</tr>
<tr>
<td></td>
<td>LEVELED</td>
</tr>
<tr>
<td>INTENT</td>
<td>SKILLS</td>
</tr>
<tr>
<td></td>
<td>MEANING</td>
</tr>
<tr>
<td>ACTIVITIES</td>
<td>EVIDENCE</td>
</tr>
<tr>
<td></td>
<td>TEXT STRUCTURE</td>
</tr>
<tr>
<td></td>
<td>OTHER SUBJECTS</td>
</tr>
<tr>
<td></td>
<td>TWO TEXTS</td>
</tr>
<tr>
<td>INFLUENCES</td>
<td>STANDARDS</td>
</tr>
<tr>
<td></td>
<td>LEAP</td>
</tr>
<tr>
<td></td>
<td>LEAP 360</td>
</tr>
<tr>
<td></td>
<td>DISTRICT</td>
</tr>
<tr>
<td></td>
<td>DISTRICT TEST</td>
</tr>
<tr>
<td></td>
<td>CURRICULUM</td>
</tr>
<tr>
<td></td>
<td>PD</td>
</tr>
<tr>
<td></td>
<td>PREP</td>
</tr>
<tr>
<td></td>
<td>STATE WEB SITE</td>
</tr>
<tr>
<td></td>
<td>ONLINE NETWORK</td>
</tr>
<tr>
<td></td>
<td>SUMMIT</td>
</tr>
<tr>
<td></td>
<td>EVAL</td>
</tr>
<tr>
<td>IMPLEMENTATION</td>
<td>IMPLEMENT</td>
</tr>
</tbody>
</table>