EVALUATING THE PRAGMATIC ANALYTICAL READING LEVEL INSTRUCTION FRAMEWORK: A MIXED METHODS RESEARCH AND DEVELOPMENT CASE STUDY

Katherine Elizabeth Bradarich
University of Missouri-St. Louis, kebradarich@gmail.com

Follow this and additional works at: https://irl.umsl.edu/dissertation

Part of the Education Commons

Recommended Citation
EVALUATING THE PRAGMATIC ANALYTICAL READING LEVEL INSTRUCTION FRAMEWORK: A MIXED METHODS RESEARCH AND DEVELOPMENT CASE STUDY

By

KATHERINE BRADARICH

B.A., University of California-Los Angeles, 1983
M.A.T., National-Louis University, 2001

A DISSERTATION

Submitted to the Graduate School of the UNIVERSITY OF MISSOURI-ST LOUIS
In partial fulfillment of the requirements
For the Degree

DOCTOR OF PHILOSOPHY

In

EDUCATION

May, 2012

Advisory Committee

Gwendolyn Turner, Ed.D.
Chairperson

Wolfgang Althof, Ph.D.

Robert Burnett, Ph.D.

Kathleen Haywood, Ph.D.

Copyright, Katherine Bradarich, 2012
Abstract

In light of the escalating literacy demands of the 21st Century workplace, and the reality the adolescent literacy rates remain stagnant despite considerable focus on them, there appears to be a need for translation of the role that reading comprehension plays in the growth of literacy of adolescent learners. A framework that recognizes the inherent complexities of reading at the secondary school level and provides a systematic and targeted means for flexible instruction to remediate the reading comprehension deficits of a diverse population of struggling adolescent readers was developed. The Pragmatic Analytical Reading Level Instruction (PARLI) framework integrates the essential constructs from the fields of education, cognitive science, and neuroscience as they relate to reading comprehension among adolescent learners.

A mixed methods evaluation multi-case study was conducted to provide a formative evaluation of the Pragmatic Analytical Reading Level Instruction (PARLI) framework. In particular, the ten cases of the quintain were middle school students identified as struggling readers. The PARLI framework was effective with nine of the 10 cases of the quintain when all data are considered. When exclusively considering quantitative data from reading assessments, seven of the ten reached proficiency on two out of three measures. Further development and evaluation of the PARLI framework is recommended.

Keywords: Adolescent literacy, transdisciplinary approach, instructional framework, struggling readers, mixed methods.
Dedication

I have the good fortune to be blessed in knowing a number of remarkable people to whom I wish to dedicate this work, and I will do so in narrative form beginning with the person who struck the first spark, Dr. Annette Dahlman. Annette brought me to the University of Southern California as part of her doctoral work when I was six years old, from which point on I knew that I would pursue my own doctorate someday. I am thankful that I was able to share my gratitude with Annette before her recent passing.

My brothers, George, Frank, Paul, and Joe, made me strong and courageous enough to take chances because I knew they were there to challenge me and build me up. I cannot imagine celebrating any significant accomplishment in my life without these four wonderful men.

My parents, George and Florine Perkovich, populated my formative years with the most incredible array of fascinating and accomplished people including a cadre of amazing friends who showed me the full spectrum of what it means to be a productive member of society. My parents and these friends taught me that purpose is found less in the circumstances of our birth and more in what we do within our circumstances, and this principle has informed every good decision I have made in my life.

I married the greatest guy I have ever met when I was too young to realize how fortunate I was; I know now. Joe Bradarich has given me loving support, tangible and intangible, for nearly three decades. Our journey has taken us through the hell of cancer, the physical separation of hundreds of miles, and the incredibly long and demanding process of realizing the dream that fills the pages before you.

Our amazing daughters, Laura and Rebecca, relentlessly encouraged me to chase my dream until I captured it; I am proud and humbled to have produced two such beautiful and gracious young women, and I hope I can continue to support their dreams as they continue to embody and support mine.

I am not sure if I would have found my way out of the corporate world and into teaching without Bobbi Snow, who saw the teacher in me and drew me into the field. She has been an incredible mentor, sister, and friend throughout this journey.

And finally, to my wonderful friends and champions Barbara McKay, Beverly Lofton, and Shelley Hoffman, who saw me through some of the roughest days of my life, encouraged me to be the best teacher I am capable of being, and listened to my endless monologues about this dissertation.

For all these and the unnamed people whose small acts of kindness gave me the strength to persist, day to day, I thank you. Truly, I am the most fortunate of women.

- Dr. Katherine Bradarich
Acknowledgments

Enduring thanks to my dissertation committee—Gwendolyn Turner, Wolfgang Althof, Kathleen Haywood and Richard Burnett—for their insightful advice throughout the dissertation process. This study could not have been completed without their wise council.

When Dr. Gwendolyn Turner agreed to be my Doctoral Chair I truly hit the lottery. I cannot imagine a better shepherd to guide me through the unfamiliar terrain of the dissertation process. Dr. Turner knew when to push and when to affirm, and I thank her for making the dissertation process so rewarding and enjoyable.

Dr. Wolfgang Althof met me when I didn’t know enough about research to understand how little I knew; a few minutes into our first class I realized that he would significantly expand my knowledge and skill. At times when I wanted to take an easy path, I thought of Dr. Althof and redoubled my efforts.

Dr. Kathleen Haywood was kind enough to provide constructive critical feedback on the structure and formatting as well as the content of the study.

Dr. Richard Burnett’s knowledge in the field of reading and his thoughtful insights kept me from veering off the road down too many side streets.

Finally, Dr Neal Sarahan hired me for my dream job when I was ABD and gave me ample time and resources to complete my writing. As if that was not enough, he also provided me with constructive feedback on the dissertation.

Thank you all for your guidance and the vital support you gave; together we have created a piece of work of which I am proud. It has been my honor to work with you on this labor of love.
# TABLE OF CONTENTS

Chapter 1 – Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of the Study</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>7</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>8</td>
</tr>
<tr>
<td>Hypothesis and Research Questions</td>
<td>8</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>9</td>
</tr>
<tr>
<td>Delimitations of the Study</td>
<td>12</td>
</tr>
<tr>
<td>Limitations</td>
<td>12</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>14</td>
</tr>
<tr>
<td>Significance of Study</td>
<td>16</td>
</tr>
<tr>
<td>Organization of Dissertation</td>
<td>17</td>
</tr>
</tbody>
</table>

Chapter 2- Literature Review

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adolescent Literacy in the United States</td>
<td>19</td>
</tr>
<tr>
<td>Reading Comprehension and the Struggling Reader</td>
<td>23</td>
</tr>
<tr>
<td>Theoretical Basis of PARLI</td>
<td>26</td>
</tr>
</tbody>
</table>
### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Study: Holly</td>
<td>247</td>
</tr>
<tr>
<td>Case Study: Bella</td>
<td>265</td>
</tr>
<tr>
<td>Case Study: Alice</td>
<td>283</td>
</tr>
<tr>
<td>Summary of Chapter</td>
<td>299</td>
</tr>
</tbody>
</table>

**Chapter 6 - Discussion**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of the Study: Problem and Purpose</td>
<td>301</td>
</tr>
<tr>
<td>Hypothesis and Research Questions</td>
<td>301</td>
</tr>
<tr>
<td>Review of the Methodology</td>
<td>302</td>
</tr>
<tr>
<td>Limitations</td>
<td>302</td>
</tr>
<tr>
<td>Major Findings by Case</td>
<td>305</td>
</tr>
<tr>
<td>Cross-case Analysis</td>
<td>320</td>
</tr>
<tr>
<td>Implications for Action</td>
<td>327</td>
</tr>
<tr>
<td>Recommendations for Further Research</td>
<td>333</td>
</tr>
<tr>
<td>References</td>
<td>334</td>
</tr>
<tr>
<td>Appendix A - Sample PARLA Assessment</td>
<td>355</td>
</tr>
<tr>
<td>Appendix B - MARS Survey Measure</td>
<td>366</td>
</tr>
<tr>
<td>Appendix C - MAASCM Survey Measure</td>
<td>368</td>
</tr>
</tbody>
</table>
Appendix D- Permissions............................................................... 371
Appendix E- Initial Codes.......................................................... 375
Appendix F- Definition of Codes and Coding Rules...................... 382
Appendix G- Event Flow Diagrams.............................................. 393
List of Figures................................................................. xi
List of Tables................................................................. xii
# List Of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Theoretical Basis for the PARLI Framework</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>The Nature of Skill Development</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>Growth Curves</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>Cyclical Spurts of Development</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>Skill Scale Illustration of Tiers of Representations &amp; Abstractions</td>
<td>38</td>
</tr>
<tr>
<td>6</td>
<td>Cortical Network Cycle of Representations &amp; Abstractions</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>Study Diagram</td>
<td>75</td>
</tr>
<tr>
<td>8</td>
<td>Distribution of PABs in Case 1</td>
<td>139</td>
</tr>
<tr>
<td>9</td>
<td>Case 1- Tracking PABS Across the Study</td>
<td>140</td>
</tr>
<tr>
<td>10</td>
<td>Distribution of PABs in Case 2</td>
<td>155</td>
</tr>
<tr>
<td>11</td>
<td>Case 2- Tracking PABS Across the Study</td>
<td>156</td>
</tr>
<tr>
<td>12</td>
<td>Distribution of PABs in Case 3</td>
<td>173</td>
</tr>
<tr>
<td>13</td>
<td>Case 3- Tracking PABS Across the Study</td>
<td>174</td>
</tr>
<tr>
<td>14</td>
<td>Distribution of PABs in Case 4</td>
<td>190</td>
</tr>
<tr>
<td>15</td>
<td>Case 4- Tracking PABS Across the Study</td>
<td>191</td>
</tr>
<tr>
<td>16</td>
<td>Distribution of PABs in Case 5</td>
<td>207</td>
</tr>
</tbody>
</table>
Figure 17. Case 5 - Tracking PABS Across the Study ................................. 208

Figure 18. Distribution of PABs in Case 6 ............................................. 224

Figure 19. Case 6 - Tracking PABS Across the Study ............................ 225

Figure 20. Distribution of PABs in Case 7 ............................................. 242

Figure 21. Case 7 - Tracking PABS Across the Study ............................ 244

Figure 22. Distribution of PABs in Case 8 ............................................. 259

Figure 23. Case 8 - Tracking PABS Across the Study ............................ 260

Figure 24. Distribution of PABs in Case 9 ............................................. 277

Figure 25. Case 9 - Tracking PABS Across the Study ............................ 279

Figure 26. Distribution of PABs in Case 10 .......................................... 294

Figure 27. Case 10 - Tracking PABS Across the Study ......................... 295
List Of Tables

Table 1. The MST Taxonomy of Personal Agency Beliefs………………….. 32

Table 2: Considering the Skill Scale and the PARLI Framework Together….. 41

Table 3. Think-aloud Protocol Categories…………………………………… 101

Table 4. Study Timeline from Implementation to Analysis………………….. 105

Table 5. Matrix of Measures Used for Mixed Methods Analysis……………. 122

Table 6. Reflections on MARSI survey by reading strategy for Case 1……… 142

Table 7. Frequencies of Levels of think-aloud protocols for Case 1………….. 142

Table 8. Case 1: Frequencies of Levels present in work by unit……………… 143

Table 9. Case 1: Frequencies of Quality present in work by unit……………… 143

Table 10. Reflections on MARSI survey by reading strategy for Case 2……… 158

Table 11. Frequencies of Levels of think-aloud protocols for Case 2………… 159

Table 12. Case 2: Frequencies of Levels present in work by unit……………… 160

Table 13. Case 2: Frequencies of Quality present in work by unit……………… 160

Table 14 Reflections on MARSI survey by reading strategy for Case 3……… 176

Table 15. Frequencies of Levels of think-aloud protocols for Case 3………… 176

Table 16. Case 3: Frequencies of Levels present in work by unit……………… 177
Table 34. Case 8: Frequencies of Levels present in work by unit……………

Table 35. Case 8: Frequencies of Quality present in work by unit…………

Table 36. Reflections on MARSI survey by reading strategy for Case 9……

Table 37. Frequencies of Levels of think-aloud protocols for Case 9………..

Table 38. Case 9: Frequencies of Levels present in work by unit………………

Table 39. Case 9: Frequencies of Quality present in work by unit……………

Table 40. Reflections on MARSI survey by reading strategy for Case 10……

Table 41. Frequencies of Levels of think-aloud protocols for Case 10………..

Table 42. Case 10: Frequencies of Levels present in work by unit………………

Table 43. Case 10: Frequencies of Quality present in work by unit……………. 
Chapter 1: Introduction

Overview of the Study

Data collected in recent years indicate that too many young adults in America are not proficient readers (Planty et al., 2009), prompting an increased interest in adolescent literacy. The shift to the standards-based, high accountability environment of public schooling, and the stated national good (National Center for Education Statistics [NCES], 2010) that all students learn at high levels have provided the impetus for a focus on struggling readers, although more attention has been placed on younger readers than adolescent readers (Biancarosa & Snow, 2004). When students reach middle school, they must navigate increasingly complex texts (Greenleaf & Hinchman, 2009). Schools face both a tremendous opportunity to build academic literacy and the risk of the consequences of failing to do so with these struggling readers.

Useful discourse about a concept as broad and deep as literacy must first begin with a definition of the terms. Comparing literacy rates requires an understanding of historical meanings as well. In the United States, and in many places throughout the world, literacy was defined in the past as the ability to read or write at all, and was measured through self-reporting as well as reading tests.

Background

In 1940, the U.S. Census began counting as literate any adults with a certain number of years of school attendance (Kutner et al., 2007). However, the definition of literacy used by the National Assessment of Educational Progress (NCES, 2009) defines literacy more practically as “using printed and written information to function in society, to achieve one’s goals, and to develop one’s knowledge and potential” (Walton, 1989, p.
2). Similarly, the National Assessment of Adult Literacy (NAAL) expands on this pragmatic definition, breaking literacy into three types, each of which is measured on a scale of 0 to 500:

Prose literacy: The knowledge and skills needed to search, comprehend, and use information from continuous texts. Prose examples include editorials, news stories, brochures, and instructional materials.

Document literacy: The knowledge and skills needed to search, comprehend and use information from noncontinuous texts. Document examples include job applications, payroll forms, transportation schedules, maps, tables, and drug and food labels.

Quantitative literacy: The knowledge and skills needed to identify and perform computations using numbers that are embedded in printed materials. Examples include balancing a checkbook, figuring out a tip, completing an order form, and determining the amount of interest on a loan from an advertisement. (Kutner et al., 2007, p. iii)

For the current study, concern was restricted to prose and document literacies, with specific emphasis on school-based, or academic, literacy and not the various forms of literacy explored in students’ lives outside of school. This focus was appropriate when one considers that literacy is a requirement of school success. School success determines achievement of diplomas, and this validation of competence is, in turn, associated with opportunity and advancement in many fields, especially as we continue to progress in the knowledge age (U. S. Dept. of Labor, 2000).

In 1983, Chall put forth a sequential developmental stage model of literacy that continues to inform the discussion of literacy development today. These stages include:

- Stage 0: Prereading, pseudo-reading;
- Stage 1: Initial reading and decoding;
- Stage 2: Confirmation and fluency;
- Stage 3: Reading for learning the new;
• Stage 4: Multiple viewpoints;
• Stage 5: Construction and reconstruction (Chall, 1983). When reading is considered as a developmental process, the opportunity for more productive conversations about struggling adolescent readers is expanded. Accordingly, Jacobs (2008) advocates a re-focusing on the stages of the reading model that Chall (1983) put forth to serve four important goals: (a) clarifying purposes and timing of particular skills K through 12, (b) recognizing that explicit skill instruction needs to continue beyond elementary, (c) shifting the focus to appropriate scaffolding in the first place to reduce the need for later remediation, and (d) understanding that the reading stage framework eliminates arguments about reading skills versus processes (Jacobs, 2008). Reading comprehension, both prose and document, needs to be explored through a developmental lens so that the reading demands of the secondary level can be adequately staged and scaffolded as a deliberate part of the daily schedule.

Recently there has been an increase in the focus on adolescent literacy in the United States, fueled in no small part by the frustrations of middle and high school teachers faced with increasing numbers of students arriving “without the requisite knowledge, skills, or disposition to read and comprehend the materials placed before them” (Snow, 2002, p. iii). Adolescent literacy poses a challenge to contemporary American education that engenders a fair amount of consensus on how to take positive action toward improvement. However, while a number of meta-analyses on adolescent literacy suggest that there is a known solution in what to do to address the current crisis
(Biancarosa & Snow, 2004; Kamil, 2003; Kamil et al., 2008; Phelps, 2005; Phillips, 2005; Scammacca et al., 2007; Slavin, Cheung, Groff, & Lake, 2008; Torgesen, Houston, Rissman, Decker et al., 2007), recent national results highlight the gap on a broad scale between knowing what to do and actually getting it done. In the most recent National Assessment of Educational Progress (NCES, 2010), a third (28%) of eighth grade students were proficient or above in reading comprehension, 2% were advanced, 43% were basic, and 26% were below basic (U.S. Dept. of Education, 2010). While these performance summaries indicate a slight improvement from the previous year’s report (Planty et al., 2009), the percentage of students not demonstrating competence in literacy is alarming. As Kamil (2003) explains, the levels of literacy (prose and document combined) for the eighth grade readers are categorized as follows:

- **NAEP basic** at eighth grade is a score of 243 and describes a student who can:
  - demonstrate literal understanding of what they read,
  - make some interpretations,
  - identify specific aspects of the text that reflect overall meaning,
  - extend the ideas in the text by making simple inferences,
  - recognize and relate interpretations and connections.

- **NAEP proficient** at eighth grade is a score of 281 and describes a student who can:
  - show an overall understanding of text, including inferential as well as literal information,
  - if grade level text, should be able to extend ideas in text by making clear inferences from it,
should be able to draw conclusions,
should connect to own experiences (including other reading),
should be able to identify some of devices authors use in composing text.

- **NAEP advanced** at eighth grade is a score of 323 and describes a student who can:
  - describe the more abstract themes and ideas of overall text,
  - analyze both meaning and form and support analyses explicitly with examples from text,
  - extend text information by relating it to their experiences and world events.

Improving adolescent literacy in the United States has been a persistent problem, with little change in literacy rates over several decades (Torgesen et al., 2007). National efforts have included the *No Child Left Behind Act* (NCLB, 2008) and the more recent *Elementary and Secondary Education Act of 2010* (U.S. Department of Education, 2010). Researchers have evaluated what is being done and how well it is working. The *Striving Readers* initiative, which aims to improve adolescent literacy rates in Title I schools while promoting significant contributions to the adolescent literacy research base, uses competitive grants funded by the U.S. Department of Education and has reported limited success to date (U.S. Department of Education, 2009).

A variety of meta-analyses summarize a broad range of adolescent literacy research that incorporates both research focused specifically on reading instruction and research that considers a more comprehensive literacy domain (Biancarosa & Snow, 2004; Kamil, 2003; Kamil et al., 2008; Phelps, 2005; Phillips, 2005; Scammacca et al., 2007; Slavin et al., 2008; Torgesen et al., 2007). One particular study (Bates, Breslow,
Hupert, 2009) highlights efforts being put forth to incorporate research-based best-practices in five states, focusing on the various ways states are implementing wide-scale changes in literacy instruction, but does not address the academic results. As such, findings from this research do not offer conclusive evidence to guide adolescent literacy instruction.

The literacy demands of life beyond school are escalating, as demonstrated through labor market analyses and the daily realities of the population’s constantly linked-in lives. The literacy rates are not keeping pace (Casner-Lotto & Barrington, 2006; U.S. Dept. of Labor, 2000; U.S. Dept. of Labor, 2002). In secondary school students engage in little sustained reading during the course of an average day (Greenleaf & Hinchman, 2009). Subject area teachers are expected to cover vast amounts of content and do not generally have the training to teach literacy tasks (Lewis & Moorman, 2007; Phelps, 2005; Sturtevant, 2003). Not surprisingly, these factors combine with the belief that the majority of content area teachers, when confronted with students who either struggle with reading or who are simply reluctant to do so, default to teaching the content in ways that excuse students from the challenging work of academic literacy resulting in missed opportunities to build these complex and powerful skills (Greenleaf & Hinchman, 2009).

In 2002 the RAND Reading Study Group (as cited in Snow, 2002) identified areas for research focus, including (a) in the area of comprehension instruction, the need to understand how instruction about strategies to improve reading comprehension leads to students’ wide application of these strategies to successfully tackle comprehension tasks; (b) exploration of how we construct informal reading comprehension assessments that
support teachers in identifying and addressing the specific needs of students with low comprehension; and (c) how measures of motivation and engagement in reading can be linked to the development of comprehension skill to inform classroom instruction (Snow, 2002).

Literacy demands of life beyond school are great and current levels of adolescent literacy are inadequate to these needs. A broad research base is available for literacy professionals to use to inform effective instruction among adolescent learners.

**Statement of the problem**

There appears to be a need for translation and understanding of the role that reading comprehension plays in the growth of literacy of adolescent learners. A framework is needed that integrates research findings from human development and the development of reading comprehension, especially in adolescent literacy. This framework needs to recognize the complexities of reading at the secondary school level and provide a systematic and targeted means for flexible instruction to remediate the reading comprehension deficits of a diverse population of struggling adolescent readers.

Contributions from theories of reading (Côté & Goldman, 2004; Hillocks, 1980; Hillocks & Ludlow, 1984; Pressley & Afflerbach, 1995; Spiro, 2004), motivation (Ford, 1992, 1995; Gordon Rouse, 2001; Gordon Rouse & Cashin, 2000), engagement (Guthrie, 2004; Guthrie, McRae, & Klauda, 2007; Guthrie, Wigfield, Barbosa et al., 2004; Guthrie, Wigfield, & VonSecker, 2000; ), and agency (Bandura, 2001, 2006, 2008) have identified essential framework constructs. Using these constructs, this author designed the Pragmatic Analytical Reading Level Instruction (PARLI) framework. The PARLI framework uses research from the fields of education, cognitive science, and
neuroscience to improve academic literacy among adolescents currently 6 months to 2 years behind grade level in reading comprehension, as measured by standardized tests of reading.

**Purpose of the study**

The intent of this study was to conduct a formative evaluation of the Pragmatic Analytical Reading Level Instruction (PARLI) framework using a case study methodology with a mixed methods design. In the study, two standardized tests of reading comprehension, the Gates-MacGinitie Reading Test (GMRT-4), and the Scholastic Reading Inventory (SRI) were used. The Metacognitive Awareness of Reading Strategies Inventory (MARSI), and the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM), a survey of metacognitive awareness of reading, and a survey measure of agency and motivation regarding academic settings were used to measure the relationship between implementation of the PARLI framework and reading comprehension, metacognition, agency, and motivation regarding academic literacy. At the same time, the cognitive processes of reading comprehension (through metacognition) and the process of developing agency regarding academic literacy was examined using observations, think-aloud protocols, and learning reflection logs with eighth grade struggling readers at a Midwestern middle school.

**Hypothesis and research questions**

Struggling adolescent readers are students who demonstrate skill deficits in reading comprehension when compared to grade level expectations. The PARLI intervention protocol was designed to remediate those skill deficits.
**Research questions.** Is the Pragmatic Analytical Reading Level Instruction (PARLI) framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

*Sub-questions for the PARLI evaluation case study.*

1) *Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*

2) *Do students participating in PARLI report a shift in agency and motivation?*

3) *Do students participating in PARLI demonstrate improved metacognition?*

4) *Does student performance on the assessments form a pattern of development?*

5) *How can measurement tools, including observations, used with struggling readers result in better understanding of these students and how to optimize their learning opportunities?*

**Hypothesis.** Participation in the PARLI curriculum framework for one academic semester would result in growth of reading comprehension among struggling readers.

**Theoretical framework**

motivation and engagement research (Bandura, 2006; Ford, 1992; Guthrie, 2004; Guthrie et al., 2007; Immordino-Yang & Damasio, 2007); and (e) neuroscience research as it pertains to reading, learning, and motivation and engagement (Bunge, Klingberg, Jacobsen, & Gabrieli, 1999; Cooke et al., 2001; Cutting, Eason, Young, & Alberstadt, 2009; Dehaene, 2009; Della Chiesa, Davis, Miyamoto, & Momii, 2007; Giedd et al., 1999; Schmalhofer & Perfectti, 2007; Shaywitz et al., 2004). The PARLI framework uses the research of the five fields listed; however, the core theoretical work that inspires the framework comes from an integration of Hillocks’ (Hillocks, 1980; Hillocks & Ludlow, 1984) taxonomy of skills in the interpretation of fiction, and Fischer’s (1980) Dynamic Skill Theory. Both of these theories develop complex and thorough hierarchical models that explicate both the process and outcome of the complex and dynamic skills that make up reading comprehension. It is perhaps explanatory of the lack of progress when considering the current levels of documented success with struggling adolescent readers, that neither Fischer (1980) nor Hillocks (1980) and Hillocks and Ludlow (1984) appear in the adolescent reading literature with any frequency. Using these theoretical underpinnings of hierarchical and taxonomic levels of reading comprehension (Hillocks, 1980; Hillocks & Ludlow, 1984), and levels and tiers of cognitive development of complex skills (Fischer, 1980), predicates a commitment to the development of an integrated, dynamic frame. It is a complex and dense task.

PARLI is a curriculum interaction frame that is engaging and built for the practical transfer of comprehension skill from the remedial reading setting to multiple
academic contexts in the middle school. It starts with how humans actually learn, in
concert with what comprehension is and incorporates what is engaging to adolescents.

Most graduate reading certification programs at universities across the country
and packaged intervention systems sold by educational publishers, focus on instruction to
build competence with the use of discrete, specific comprehension strategies. Some of
these strategies have been found to have functional merit, and a number of experts in the
field support using them in a coordinated fashion (Alvermann, 2002; Beers, 2003;
Daniels & Steineke, 2004; Harvey & Daniels, 2009; Harvey & Goudvis, 2007; Robb,
2008). However, it is precisely these strategies that have largely failed the struggling
reader. Therefore, PARLI starts with the premise that struggling readers at the middle
school level lack a framework that enables them to fit individual strategies within an
overall schema of reading comprehension. As Torgesen et al., (2007) conclude, the
literacy instruction received by students who arrive at middle school struggling to
comprehend, obviously did not work. Competent readers have managed to fit these
strategies into an overall schema for comprehension. They are able to evaluate, almost
instantaneously, which strategy is called for in a particular reading comprehension
context, draw on that strategy, and understand what they read (Alexander, 2008; Beers,
2003; Harvey & Daniels, 2009). The struggling readers are not likely to generalize from
reading class to content area classes unless teachers explicitly teach, support, and
elaborate the strategies’ use with content area texts (Torgesen et al., 2007). Add to this a
growing body of research regarding the level of disengagement that comes with specific
skill-and-drill remedial instruction that is documented in the educational and cognitive
research literature (Alvermann, 2002; Fischer & Immordino-Yang, 2002; Langer, 2001), and it is possible to understand why some adolescents remain unsuccessful.

**Delimitations of the study**

The study was conducted in one middle school in a Midwestern suburban school district. At its broadest scope, the study encompassed eighth grade students. Two of the standardized reading measures (the Missouri Assessment Program or MAP, and the Gates-MacGinitie Reading Test or GMRT-4) used in this study included all eighth grade students attending the middle school that was the focus of the PARLI framework intervention. The third standardized reading measure, the Scholastic Reading Inventory (SRI), is used for the study cohort of cases being studied as a multicase (the quintain) and their grade level peers who are also receiving reading intervention services in the study district. The PARLI implementation and qualitative aspects of the study were focused on one group of struggling eighth-grade readers (10 students), who would otherwise have received the same reading intervention services to remediate their current below-grade level comprehension skills as the identified eighth grade struggling readers in the other four middle schools across the district.

**Limitations**

Because purposeful sampling was used in the quantitative portions of the study, the researcher cannot say with confidence that the sample was representative of the population (Creswell, 2003). Purposeful-criterion sampling was selected for the qualitative elements of this study, limiting the ability to generalize. Furthermore, this small sample participating in the PARLI framework diminished the statistical power of the analysis, in addition to the effects of the purposeful nature of this sample on
generalizability. As such, the focus was on analytical generalization as explained in subsequent sections of this dissertation. Because of the nature of qualitative research, the data obtained using qualitative methods may be subject to different interpretations by different readers. Finally, the study features the researcher as the teacher implementing the framework, which may have resulted in bias. Utilization of a panel of literacy experts to analyze audio recorded think-aloud protocols at both the start and conclusion of this study provided control for this potential bias.

Bracketing was also used to address potential bias. According to Crotty (as cited in Ahern, 1999) bracketing is “the means by which researchers endeavor not to allow their assumptions to shape the data collection process and the persistent effort not to impose their own understanding and constructions on the data” (p. 407). A chief mechanism that was incorporated in this study involved bracketing as a part of the audit trail through regular, reflective memos. Some of the considerations were to assess whether students were being candid during think-aloud protocols, or telling what they think the researcher wanted to hear. The reality was that the researcher, as the teacher, had power in this situation that may influence student behavior.

The researcher’s belief in the essential nature of motivation in learning also required sensitivity and vigilant reflection to ensure that the data was central to the analysis. The teacher/researcher role has potential as a role conflict, as does the researcher/framework developer dichotomy. That said, because the framework was designed to be flexible and responsive to student needs, the researcher’s commitment to maintaining awareness of these potential role conflicts may have mitigated any potential bias in this area. As a teacher who works with struggling adolescents, this researcher was
rarely neutral toward students. There was no attempt made to become neutral, as this would impede teaching effectiveness, but rather to clearly articulate this position and use reflections as a way to give it voice in an open and frank manner. Throughout the process, neither the researcher nor the literacy panel of experts recognized any bias.

Definition of terms

**Academic literacy.** Academic literacy is reading and writing effectively in academic settings at the appropriate grade level, and encompasses both prose and document literacy.

**Adolescent literacy.** For this study, adolescent literacy was defined as the complex system of reading skills among young adults that encompasses academic literacy. Writing is included only as it serves development and articulation of reading comprehension.

**Agency.** Agency, for the purposes of this study, is the active expression of self-efficacy. Agency comes from Bandura’s social cognitive theory. More specifically, “to be an agent is to influence intentionally one’s functioning and life circumstances” (Bandura, 2008, p. 16). There are four properties of agency: intentionality, temporal extension of agency through forethought, self-reactiveness, and self-reflectiveness. Finally, “agency embodies the endowments, belief systems, self-regulatory capabilities and distributed structures and functions through which personal influence is exercised,” (Bandura, 2001, p. 2)

**Literacy.** The effective combination of both the receptive language skills involved in reading and the expressive language skills involved in writing about what is read.
Metacognition. This study used the definition of metacognition from the study of executive function as that which refers to “The ability to stand back and take a bird’s-eye view of oneself in a situation. It is the ability to observe how you problem solve. It also includes self-monitoring and self-evaluative skills” (Dawson, 2010, p. 1).

Motivation. The definition of motivation from motivation systems theory (MST) was used for this study: “In MST, motivation is defined as the organized patterning of three psychological functions that serve to direct, energize, and regulate goal-directed activity: personal goals, emotional arousal processes, and personal agency beliefs” (Ford, 1992, p. 3).

Pragmatic. Both the linguistic meaning and the standard meaning were included in defining pragmatic. The linguistic meaning of the term is as a noun meaning the study of language as it is used in a social context, including its effect on the interlocutors. In daily language use, pragmatic is an adjective that means pertaining to a practical point of view or dealing or concern with facts or actual occurrences. In this framework, pragmatic was used as an adjective to describe the focus on the practical comprehension of text language in everyday content-area settings.

Quintain. Quintain is the group of cases considered together. This term was coined by Stake (2006) for considering the collective in a multi-case study. This umbrella term is used to refer to the cases being studied when they are being considered as one overall case.

Reading comprehension. Reading comprehension is a complex concept that for the purposes of this study was defined as “the process of simultaneously extracting and constructing meaning through interaction and involvement with written language” (Snow,
Comprehension entails three elements: “the reader who is doing the comprehending, the text that is to be comprehended, and the activity in which comprehension is a part” (p. 11).

**Struggling readers.** Struggling readers are those adolescents who arrive at middle school reading 6 months or more below grade level. The subset of struggling readers that was the focus of this study were those identified as being between 6 months and 2 years below grade level, based on standardized tests of reading comprehension.

**Significance of the study**

The current study addressed several key issues that are considered important in the literacy research agenda (Snow, 2002). By using the breadth and depth of multiple case studies to understand these struggling readers individually and as a group, in all their complexity, this study can contribute to the theoretical knowledge base on how motivation, agency, and metacognition can inform literacy instruction for struggling adolescent readers.

This knowledge can play a role in the building of educator capacity to address adolescent literacy issues by offering a detailed window through which to view the struggling reader. To the extent that students participating in a PARLI curriculum format demonstrated improved metacognition through both classroom observation and self-report, this study provided a contribution to theory and practice by illustrating how experimental research in the cognitive field can inform both research and practice in the complex, real-life environment of middle school.

The PARLI results as the quintain moved through the units displayed a cyclical spurt growth pattern consistent with Fischer’s Dynamic Skill Theory. Along with the
growth in reading comprehension exhibited in the quintain as a whole, results suggest that the PARLI framework can provide a means for attending to how the complex, dynamic skill of reading comprehension development becomes visible through every day work products. This focus on teaching and learning in dynamic and engaging ways can enable middle schools to more effectively design and implement curricula to support the growth of all adolescent readers.

The greater progress of the quintain in the unit of longest duration in the PARLI framework implementation suggests that the PARLI framework may also have broader application as a tool to build reading competency at the secondary school level through professional development of teachers and implementation within the content area courses. This potential application is consistent with the Standards interdisciplinary approach (Council of Chief State School Officials [CCSSO] & National Governor’s Association [NGA], 2010), as well as recommendations included in the Rand Reading Study Group report (Snow, 2002).

**Organization of the dissertation**

This dissertation is organized into six chapters, a reference list, and appendices. Chapter 1 of this dissertation provides an introduction and overview of the topic of research. Chapter 2 provides an overview of the selected literature and the theoretical framework that informs the present study. The researcher integrated work from four fields of study to approach the instructional framework being evaluated. Chapter 3 explains the PARLI framework. Chapter 4 describes the research design and methodology including sample selection, description of setting, data collection, and analysis. Chapter 5 provides the results of the study, while Chapter 6 is devoted to the
discussion of those results. All resources used in this study are the original work of the author, or permission to use copyrighted materials has been obtained by the author.
Chapter 2: Literature Review

In this chapter, a review of selected theoretical and research literature on the topic of this study, *Evaluating the PARLI framework*, is presented. Major variables and concepts across sources are analyzed to discern what is currently known, as well as what is still to be discovered. The examination of the literature begins with and exploration of general adolescent literacy research and specific analysis of the research base with regard to reading comprehension and the struggling reader. The analysis then utilizes five basic research lenses: The Nature of Reading, Research-Based Instructional Strategies, Cognitive Development, Motivation and Engagement, and Neuroscience, as they pertain to learning and reading comprehension development in particular.

Adolescent literacy in the United States

While the current challenge is unquestionably great, it is important to note that educators have been facing the challenge of developing adolescent literacy to match the demands of the workplace for a very long time (Torgesen et al., 2007), with data for several decades revealing no major shifts in literacy rates. The issue now, at the beginning of the 21st Century, is how much the nature of work has changed from earlier decades. As the macroeconomic structure continues its progression from a once predominantly agricultural basis, to a predominantly industrial one, and now to a global economy based on information exchange, there is a need for a highly literate workforce.

In the most recent data (U.S. Dept. of Labor, 2009), the trend established a decade ago (U.S. Dept. of Labor, 2000, 2002) continues with the projections that the fastest growing jobs between 2008 and 2018 are those requiring college degrees. Even more noteworthy is that the majority of new jobs being created require considerable knowledge
gained from on-the-job training. This reality means that workers in these jobs will need good, basic reading, communication, and mathematics skills to get a job and grow a career. In the recent survey of American businesses, Casner-Lotto and Barrington (2006) found that more than a third of the businesses report deficiencies in reading comprehension among high school graduates. Among 2-year and 4-year college graduates, the deficiencies that draw the most attention are writing skills, with between a quarter and a half of the responding employers expressing these concerns for both 2-year college and 4-year college graduates respectively.

Importantly, the nature of education in the United States has changed from its history as a mechanism of sorting people into categories, with some designated for post-secondary education and many not, to a stated purpose of educating all citizens, as exemplified in the overarching No Child Left Behind Act of 2001 (NCLB). As the United States evaluates this legislation and moves toward reauthorization, the Obama administration is promoting stringent standards with a goal of every student being college and career ready by 2020 (U.S. Department of Education, 2010). The blueprint for reauthorization of the Elementary and Secondary Education Act (U.S. Department of Education, 2010) states goals in four areas:

(1) Improving teacher and principal effectiveness to ensure that every classroom has a great teacher and every school has a great leader; (2) Providing information to families to help them evaluate and improve their children’s schools, and to educators to help them improve their students’ learning; (3) Implementing college- and career-ready standards and developing improved assessments aligned with those standards; and (4) Improving student learning and achievement in America’s lowest-performing schools by providing intensive support and effective interventions (U.S. Department of Education, 2010, p. 3).

This developing movement on the national education reform front toward college and career readiness for all is illustrated by the Common Core State Standards for
English Language Arts and Literacy in History/Social Studies & Science (Standards) being proposed by the Council of Chief State School Officers (CCSSO) and the National Governors Association (NGA) that “builds on the foundation laid by states in their decades-long work on crafting high-quality education standards” (CCSSO & NGA, 2010, p. 1). With “college and career readiness” being the endpoint of the Standards, they describe “a vision of what it means to be a literate person in the twenty-first century” (CCSSO & NGA, 2010, p. 1). In light of the demands of our times, a flexible and multifaceted method of adolescent literacy instruction is needed that actually teaches the transfer that allows young adults to systematically conquer texts across disciplines and contexts. In addition, to develop full-fledged literacy, the reader needs to be able to articulately voice his or her understandings, synthesize the information, and apply it to novel circumstances. Toward these ends, educators throughout the United States are implementing secondary literacy practices incorporating the suggestions from research in the educational field.

Bates, Breslow, and Hupert (2009) report on five states committed to fully implementing the best practices in adolescent literacy (Alabama, Florida, Kentucky, New Jersey, and Rhode Island). The report describes the policies crafted by different states and shares how the states (a) engaged key stakeholders, (b) set rigorous goals and standards, (c) aligned resources to support adolescent literacy goals, (d) used data to measure progress, and (e) built educator capacity (Bates et al., 2009). All five states focused on engaging key stakeholders as part of their plans. Four of the five states emphasized using stakeholder expertise and feedback to inform their policies. The fifth brought in family literacy programs to make literacy a greater priority for a wider base of people. All five
states reported that rigorous goals and standards were either in place or were continuing to be developed. When it comes to aligning resources in support of adolescent literacy goals, all five featured state policies take local context into account and had at least one education agency staff member in charge of adolescent literacy. Commitment to using data at the heart of decision making was unanimous; however, none of the five states were satisfied with the assessments at their disposal (Bates et al., 2009). Building educator capacity was a stated priority for all respondents:

All five states used a combination of direct training for teachers and training for coaches, usually with a focus on content-area literacy instruction and intervention with struggling readers. All used school-based coaches, and state-based coaches or literacy specialists were critical to professional development in all states but Florida. Yet the five states assigned different functions to such coaches and specialists, reflecting important differences among literacy improvement strategies. All states had systems for two-way communication between reading coaches or specialists and state-level staff (p. 18).

The five states studied share common challenges, and each state has interpreted the research findings to develop what they hope will be an effective adolescent literacy plan for their students. The report did not strive to compare the merits of each plan, so performance data are not included.

Another nationwide example is the Striving Readers initiative that features discretionary and competitive grants from the U.S. Department of Education with, according to the Department’s website, aims to “Raise middle and high school students’ literacy levels in Title I-eligible schools with significant numbers of students reading below grade-levels; and build a strong, scientific research base for identifying and replicating strategies that improve adolescent literacy skills” (U.S. Dept. of Education, 2009, para. 1-2). This initiative has been funding eight projects in states distributed across the country, incorporating many of the suggestions from adolescent literacy research. As
of spring of 2009, there is little concrete success to report, if success is gauged by reading test scores.

**Reading comprehension and the struggling reader**

The importance of academic literacy to both the individual and the greater society is evident through observation of the world around us. It is likewise common sense to reach the conclusion that school can be frustrating and self-defeating for those individuals who struggle to achieve the competence in reading and writing that seems to come naturally to many of their peers.

In their report evaluating the current state of reading comprehension instruction and making recommendations for future directions of research, the Rand Reading Study Group (RRSC, 2002) uses a simple heuristic for reading comprehension (reader, text, activity within a given context) that is useful to use when considering the adolescent struggling reader and how to most successfully intervene to ensure that each student has the dynamic reading comprehension skills needed for freedom and choice in adult life (Snow, 2002). By dividing reading comprehension into three focal points of analysis (the reader, the text, and the activity or purpose for reading) Snow’s (2002) heuristic was used in the development of the theoretical basis and implementation of the PARLI framework for this study. These three elements were integrated within the dynamic and iterative progression of comprehension across academic domains using the PARLI framework.

The reader is central to all consideration of reading comprehension. Each reader has his or her own array of cognitive capabilities, motivation, knowledge, and experience that he or she brings to bear on each reading task. Of course, readers vary on each of
these attributes in myriad ways that influence reading comprehension in general. Each person also performs variably based on the particular text or activity (Snow, 2002).

The next element to consider is the text itself. One need merely reflect on one’s own experiences of attempting to comprehend a particular text outside of one’s field of study as compared to reading within one’s area of expertise to illustrate this reality on a personal level. Seven common structures that writers of nonfiction use to organize their texts are: (a) web, (b) matrix, (c) list, (d) linear string, (e) cause-effect, (f) problem-solution, and (g) persuasion/argument (Dymock & Nicholson, 2007). These text structures can be grouped into descriptive and linear categories. Dymock and Nicholson conclude: “What many otherwise good readers lack is knowledge of text structure. They can read the words but they can’t see the design of the text. Many students will not develop text structure awareness without explicit teaching,” (p. 17). Vacca (1998) was an early supporter of this notion that many readers become struggling readers because they have never been explicitly taught the conceptual and contextual constructs that are part of all texts.

Current research (Snow, 2002) continues to highlight the reality that for novice and struggling readers in particular, the ability to construct the various representations of a text critical to comprehension embodies a tremendous challenge. Research on adolescent literacy across the content areas continues to make the point that struggling readers need text support knowledge (Alvermann, 2002; Beers, 2003; Daniels & Steineke, 2004; Harvey & Daniels, 2009; Harvey & Goudvis, 2007; Keene & Zimmerman, 1997; Robb, 2008; Schoenbach, Greenleaf, Cziko, & Hurwitz, 1999).
The third element, the reading activity itself, also plays a key role in reading comprehension. The reading activity is made up of purposes or tasks, operations to process the text, and the ultimate outcomes of the activity. Students need to be taught how to engage in the variety of reading comprehension activities that are part of academic literacy.

The literature provides many examples of specific strategy instruction (Alvermann, 2002; Beers, 2003; Daniels & Steineke, 2004; Fisher & Frey, 2009; Harvey & Daniels, 2009; Harvey & Goudvis, 2007; Keene & Zimmerman, 1997; Robb, 2008; Schoenbach et al., 1999; Smith & Wilhelm, 2002, 2006; Wilhelm, 1997, 2001, 2007; Wilhelm, Baker, & Dube, 2001). Among some of the general strategies with evidence-based success are: (a) comprehension monitoring, (b) cooperative learning, (c) using graphic and semantic organizers, (d) answering questions, (e) generating questions, (f) using text structure, and (g) summarizing (Alvermann, 2002). The effectiveness of these reading and writing strategies is supported by evidence from research on the effectiveness of instructional strategies in general, including the work of Marzano et al. (2001).

There are a number of ways to organize the variety of reading strategies available from the literature. Reading strategies can be grouped by engagement in the reading process: (a) before reading, (b) during reading, and (c) after reading (Beers, 2003; Fisher & Frey, 2009; Robb, 2008; Wilhelm et al., 2001). These strategies may be organized by what the reader is doing while reading: (a) visualizing, (b) connecting, (c) questioning, (d) inferring, (e) evaluating, (f) analyzing, (g) recalling, and (h) self-monitoring (Daniels & Steineke, 2004; Harvey & Goudvis, 2000; Keene & Zimmerman, 1997; Smith & Wilhelm, 2002, 2006; Wilhelm, 1997, 2001, 2007; Wilhelm et al., 2001). Furthermore,
reading strategies can be taught and practiced using a variety of models like literature circles (Daniels & Steineke, 2004), inquiry circles (Harvey & Daniels, 2009), and an apprentice approach (Schoenbach et al., 1999), to name but a few.

Cantrell and Carter (2009) define strategies as “the mental processes and procedures that readers purposefully evoke to enhance their comprehension of academic texts” (p. 199). The authors also make a distinction between skills in reading that are habitual, unconscious and automatic procedures, and strategies as being processes that are only used “when needed to gain greater meaning from text” (Cantrell & Carter, 2009, p. 199). Cantrell and Carter used the Metacognitive Awareness of Reading Strategies Inventory (Mokhtari & Reichard, 2002) in their study to examine the relationships among specific student characteristics and perceived reading strategy use. Cantrell and Carter (2009) found that better readers reported using global and problem-solving strategies more often than poor readers did, and poor readers favored support reading strategies. The authors suggest that research that examines how best to facilitate adolescents’ cognitive reading strategy development and addresses “context-and subject-specific reading strategy use and the impact of instruction in specific strategy types on students’ reading achievement are needed” (p. 217).

**Theoretical basis of PARLI**

PARLI finds support across a wide range of literature in Mind, Brain, and Education (MBE), (Fischer et al., 2007). MBE is a transdisciplinary field promoting “the integration of the diverse disciplines that investigate human learning and development—to bring together education, biology, and cognitive science to form the new field of mind, brain, and education” (p. 1). The key objectives of the field revolve around creating
dynamic reciprocal relationships between scientific research and practical knowledge, moving research into the complex, real-world educational settings that enable all three fields to be enriched by the exchange. The five primary categories of foundational support for PARLI are: (a) motivation and engagement research; (b) cognitive developmental research; (c) research-based instructional strategies, both specific to literacy and in general; (d) nature of reading skills research, and (e) neuroscience research as it pertains to the other categories. Figure 1 provides a heuristic for understanding the interaction among these five primary categories as being one that is dynamic and complex, like the act of comprehending written text itself. If one recognizes motivation and engagement as the necessary precursors to learning, then this body of research forms the leading gear in the heuristic of the theoretical system that informs the PARLI framework. The other gears in this system are cognitive development, instructional strategies, and the nature of reading. Neurology (and neuroscience) provides the unifying “belt” that both limits and enables the dynamic interaction among the others. The system of gears heuristic suggests the interactions between and among the various bodies of research that are dynamic and complex rather than linear. While in reality these five categories of research interact in complex, synergistic patterns to create understanding of the complex skill that is reading comprehension they are separated here in a linear fashion for discussion purposes.
Motivation and engagement research

Motivation and engagement are critical to the human ability to learn anything and everything, and provide a key component for consideration in any exploration of adolescent literacy. As such, motivation and engagement research is represented by the lead gear in the PARLI heuristic to represent its role in initiating action/learning. When the target population is struggling readers, it is really a discussion of re-motivation and re-engagement. Classroom instruction may be focused on one of three loci: the content, the teacher, or the learner. The American Psychological Association (APA), (1993) spells out the primacy of a learner-centered approach in effective instruction. These learner-centered principles include several that deal specifically with motivation and affective
factors. However, the key point of learner-centered instructional practices is that through their implementation, opportunities for learning are exponentially increased, largely because this implementation ultimately results in a positive shift in motivation and engagement.

In studies conducted with younger students in grades 3 and 5, Guthrie, McRae, and Klauda (2007) found that when it comes to struggling readers, this need for engagement cannot get any stronger. For the struggling reader, who has already experienced a significant amount of failure in school, getting him or her to engage in reading challenging text is no small feat. Guthrie, et al. (2007) discuss the power of autonomy-enhancing practices with adolescents, an effect that can be observed with adolescents when they read difficult text. When autonomy-enhancing practices are integrated into instruction in a positive fashion, they correlate highly with engagement and subsequent success (Guthrie et al., 2007). Guthrie evaluated an intervention called *Concept-Oriented Reading Instruction* (CORI) that is designed to enhance autonomy by increasing students’ reading comprehension and motivation for reading. CORI has a set of five motivational constructs documented in prior research by Brophy (1998), and Guthrie and Alao (1997), as well as Stipek (1996) (as cited in Guthrie et al., 2000). CORI is based on an engagement model of reading development that “suggests that reading comprehension is facilitated by reading engagement, which in this study consisted of the joint functioning of cognitive comprehension strategies and motivational processes” (Guthrie, Wigfield et al., 2004, p. 406). CORI features five practices that explicitly support student motivation and engagement: “Briefly, these instructional practices consist of (a) relevance--to foster intrinsic motivation, (b) student choice--to increase perceived
autonomy, (c) success--to build self-efficacy, (d) collaborative structures--to enhance social motivation, and (e) thematic units--to improve mastery goals” (Guthrie et al., 2007, p. 240).

In a meta-analysis of “11 studies consisting of quasi-experimental designs in which CORI was compared to one or more control groups that were initially comparable to the CORI group” (Guthrie et al., 2007, p. 244), 75 effect sizes (ESs) were computed to evaluate CORI’s impact on outcome variables. In these studies, CORI was found to have “positive effects, moderate in magnitude, on a range of internal motivations for reading” (p. 247). However, autonomy-suppressing behaviors have a strongly negative effect. “Students experiencing those autonomy-undermining practices stated that they preferred not to participate in class, did not attempt to understand material provided by that teacher and felt angry or bored in classes taught by that teacher” (p. 239). CORI has been researched as relevant to younger students than the PARLI framework, and has been studied with adolescent readers as well, but none of these studies meet the criteria of the What Works Clearinghouse (2010), making the results inconclusive. Although lacking clear demonstration of efficacy with the adolescent population, CORI provides a reference point for work with adolescent readers that focuses on engagement and motivation.

Bandura’s (2006) Social Learning Theory places emphasis on learning as a result of observation and modeling of behaviors, attitudes, and emotions. Bandura focuses specifically on issues of agency. “To be an agent is to influence intentionally one’s functioning and life circumstances” (p. 3). Agency is inextricably wrapped up in, and rather synonymous to, personal efficacy. The author continues, “Unless people believe
they can produce desired effects of their actions, they have little incentive to act or to persevere in the face of difficulties” (p. 3). It is not a leap to anticipate that the sense of agency of struggling adolescent readers regarding reading tasks may play a role in their school performance.

Ford’s (1992) *Motivation Systems Theory* (MST) provides an overall comprehensive taxonomy of motivational concepts. MST is based on the basic premise that “both precision and scope are needed to address complex, real-world problems effectively” (pp. 10-11). “In MST, motivation is defined as the organized patterning of three psychological functions that serve to direct, energize, and regulate goal directed activity: personal goals, emotional arousal processes, and personal agency beliefs” (p. 3). According to MST, self-concept is a blend of feelings about one’s abilities along with one’s assessment of the role of environmental factors, much like Bandura’s (2006) concept of agency. By considering this construct of agency as made up of these two component parts, MST (Ford, 1992) provides a frame that generates ten different patterns of motivation (*Robust, Modest, Fragile, Tenacious, Vulnerable, Self-doubting, Accepting, Antagonistic, Discouraged, and Hopeless*), as illustrated in Table 2, the patterns, in turn, dictate the approach to goals, with emotions playing a key role in the decision process of whether a given goal is attainable or not. Resilience is tied to one’s personal agency beliefs.
Table 1.

*The MST Taxonomy of Personal Agency Beliefs*

<table>
<thead>
<tr>
<th>CAPABILITY BELIEFS</th>
<th>Strong</th>
<th>Moderate or Variable</th>
<th>Weak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>R Pattern</td>
<td>Modest Pattern</td>
<td>Fragile Pattern</td>
</tr>
<tr>
<td>R Pattern</td>
<td>Robust — “strong and firm in purpose or outlook”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M Pattern</td>
<td>Modest — “placing a moderate estimate on one’s abilities”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F Pattern</td>
<td>Fragile — “intact but easily broken or damaged”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T Pattern</td>
<td>Tenacious — “suggests strength in dealing with challenges and obstacles”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V Pattern</td>
<td>Vulnerable — “functioning adequately but may be at risk under conditions of stress”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S Pattern</td>
<td>Self-Doubting — “having a lack of faith in one’s chances for success”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1 OR A2 Pattern</td>
<td>Accepting or Antagonistic — “to endure difficulties quietly and with courage”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D Pattern</td>
<td>Discouraged — “tending toward actively expressed annoyance or hostility”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H Pattern</td>
<td>Hopeless — “having no expectation of success”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Definitions (adapted from Webster’s *Seventh New Collegiate Dictionary*):

1. R Pattern: Robust — “strong and firm in purpose or outlook”
2. M Pattern: Modest — “placing a moderate estimate on one’s abilities”
3. F Pattern: Fragile — “intact but easily broken or damaged”
4. T Pattern: Tenacious — “suggests strength in dealing with challenges and obstacles”
5. V Pattern: Vulnerable — “functioning adequately but may be at risk under conditions of stress”
6. S Pattern: Self-Doubting — “having a lack of faith in one’s chances for success”
7. A1 Pattern: Accepting — “to endure difficulties quietly and with courage”
8. A2 Pattern: Antagonistic — “tending toward actively expressed annoyance or hostility”
9. D Pattern: Discouraged — “being deprived of but potentially maintaining some confidence or hope”

---

Applying MST, in a study of the comparison of self-concept and motivational patterns of 17 academically resilient and 19 non-resilient students drawn from a pool of 170 urban, Caucasian, high school sophomores, Gordon Rouse (2001) demonstrated “that more positive motivational patterns are associated with resilient students than with non-resilient ones. The patterns are associated with better academic achievement” (p. 470). Academic achievement was measured by GPA and goals, self-concept, and environmental support beliefs were measured by the Assessment of Personal Agency Beliefs and the High School Assessment of Academic Self-Concept. Both groups of students came from the same economically deprived, stressful environment, but the resilient students achieved a GPA of 2.75 or better on a 4.0 scale (Gordon Rouse, 2001).

In later research on MST (Gordon Rouse & Austin, 2002) presenting the findings of three separate studies of the relationship of GPA and gender to motivation, it was found that there were some significant differences in motivational patterns within ethnic groups based on GPA and gender. One study was conducted with African-Americans, one with Hispanic-Americans, and the third with Euro-Americans. The findings revealed significant within-group differences that also varied from one ethnic group to another. In total, Rouse’s work points out the merits of developing an understanding of motivational patterns among struggling adolescent readers.

**Cognitive developmental research**

Referring to the PARLI heuristic, cognitive developmental research is the next gear that moves in response to the movement generated by motivation and engagement. The development of the PARLI framework is based upon early work in Mind, Brain, and Education (MBE), (Fischer et al., 2007). Hillocks (1980; Hillocks & Ludlow, 1984) was
focusing specifically on reading comprehension during the 1980s. Fischer (1980) was considering the intersection of biology (neuroscience), cognitive science, development and education as he was developing *Dynamic Skill Theory* to explain how humans develop new skills. Fischer and Immordino-Yang (2002) were able to confirm and expand on earlier concepts, through the use of technological advancements of the 21st Century “to make possible a new cognitive science of education--one grounded in analysis of detailed data on learning in real-life settings such as schools,” (p. 3). Fischer’s Dynamic Skill Theory is oriented toward “a constructive web of multiple, parallel strands (domains) that sometimes intersect or divide,” (Fischer & Immordino-Yang, 2002, p. 7).

Figure 2 illustrates how, in this model, individual skills, which are represented by different strands of the web (Fischer & Biddell, 2006), develop independently within domains, “proceeding largely at their own pace,” (Fischer & Immordino-Yang 2002, p. 8). According to Fischer (2008), learning is variable based on one’s level of expertise. As shown in Figure 3, the growth curve of the novice in any particular skill is much less predictable and smooth than that of the expert.
Figure 2. The nature of skill development. This figure illustrates that the nature of skill development is represented as a web rather than a ladder.²

Figure 3 A comparison of growth curves of skill development. This figure illustrates growth curves for learning a task for novice, intermediate, and expert performers.³


Development of new technologies in the 21st Century and the emergence of a new trans-disciplinary field of MBE have resulted in the discoveries on brain activity. Brain activity demonstrates reorganizations that parallel the skill levels, suggesting a neurological basis for the cognitive discontinuities observed in optimal level performance (Fischer & Immordino-Yang, 2002, p. 19). Dynamic Skill Theory provides further theoretical and neurological support for Vygotsky’s (1978) social learning theory and posits the importance and power of scaffolding learning (Figure 3). One of the aspects of Dynamic Skill Theory that is most directly related to the development of the PARLI framework is the consideration of Fisher’s skill scale (Fischer & Bidell, 2006). The skill scale, developed as an outgrowth of Fischer’s (1980) Dynamic Skill Theory provides cognitive and neurological support for, as well as explanation of, the development of reading comprehension. For learners to move from one category to the next, instruction needs to feature scaffolding based on the skill levels from Representations through Principals, while planning for the reality that, through differentiated instruction:

- People differ in rate of development: Some move through the hierarchy of levels much faster than others. People differ in their profiles of cognitive skills--catalogues of which skills have attained which levels. And most interestingly, people differ in the paths through which they develop (p. 513).

Fischer’s (1980) Dynamic Skill Theory’s complex and dynamic scale spells out the progression of skill development:

- Representations (Rp1) - Person can represent simple properties of objects, events, and people independently of their own immediate actions.
- Representational Mapping (Rp2) - Person can relate variations in one representation to variations in another representation.
- Representational Systems (Rp3) - Person knows how two concepts relate to each other and still only concerned with overt characteristics.
- Abstractions (Rp4/Ab1) - Person can accomplish the coordination of two representational systems to get an intangible attribute that characterizes broad categories of objects, events, or people.
- Abstract Mappings (Ab2) - Person can coordinate two complex systems to achieve a strong conceptual understanding of the architecture of the systems.
- Abstract Systems (Ab3) - Person can flexibly differentiate the relationships between two concepts.
- Principles (Ab4) - Person can consider systems of abstract systems and relate groupings of multifaceted concepts to grouping of others across time and conditions.

*Figure 4.* Cyclical spurts of development. This figure illustrates the progression of skill development under optimal conditions.  

Fischer’s Dynamic Skill Scale (Fischer, 1980) allows for the slow and uneven process of human learning (Fischer, 2008). The anticipation of fluctuations in progress of

---

individual students, as depicted by the cyclical spurts of Figure 4, helps teachers set reasonable, developmentally appropriate goals and expectations by which to evaluate abstract reasoning within the context of reading comprehension. The “common scale for behavioral complexity that captures a central dimension of both long-term development and short-term change” (Fischer & Bidell, 2006, p. 323) is supported by research using various methods (Commons, Trudeau, Stein, Richards, & Krause, 1998; Dawson & Wilson, 2004; Fischer, 1980; Fischer & Bidell, 2006; Fischer & Immordino-Yang, 2002).

*Figure 5.* Skill scale illustration of tiers of Representations & Abstractions. These developmental cycles are the areas of development that pertain to middle school students.⁵

---

According to Fischer and Bidell (2006), this scale provides clear direction for curriculum designed to effectively scaffold students’ comprehension skills to ensure continued growth. Specifically, the scale illustrates that the dynamic developmental cycles of skill development move through ten levels and three tiers. Further, the tiers of Representations and Abstractions are the focus of development during the middle school years and represent the skill demands of inferential thinking required to meet the increasing reading demands, particularly in nonfiction, of this educational level. The skill scale developed as an outgrowth of Fischer’s Dynamic Skill Theory, places middle school students’ development in the range moving from the Representational tier, with most students building functional, independent skills with this type of cognition at the level of Representational Principles and Single Abstractions, equivalent to basic inferential thinking, around the age of the average sixth grade student. The dynamic development of skill (from the Representational tier to Abstract Mappings) along with the corresponding cortical development, are shown in Figures 5 and 6. Development through each level follows the same trajectory. The development of expertise at the level of Single Abstractions continues into young adulthood, with the developmental tier of Abstract Mappings beginning at the very end of middle school between ages 14-15 (Fischer & Bidell, 2006). This equates to the more complex inferential thinking expressed in Hillocks’(1980) level of Complex Implied Relationships. This type of thinking is not yet stable and independent. Table 1. matches the skill scale tiers and matches the skill scale tiers and levels with the corresponding PARLI levels. Nonfiction reading comprehension may also be hierarchical and taxonomic, but according to Fischer (2008), learning is characterized by cyclical spurts of development (Figure 6). Changes in the
nature of the text, be it the content domain, the complexity of the text, the demands of the context, or combinations of these factors results in iterative changes in ability to demonstrate reading comprehension skills, as predicted by Dynamic Skill Theory. The research reveals that no individual acts consistently at a specific level or strand; instead, the individual acts within a range of levels (Fischer, 2008; Fischer & Bidell, 2006; Fischer & Immordino-Yang, 2002), making some of the variability in scores that is seen in reading comprehension across contexts expected.

Figure 6. Cortical network cycle. This figure illustrates the cortical network cycle for the two consecutive cognitive levels of Representations & Abstractions is shown.6

---

Table 2.  
*Considering the Skill Scale and the PARLI Framework Together*

<table>
<thead>
<tr>
<th>Skill Scale Tier &amp; Level &amp; Description</th>
<th>PARLI Level &amp; Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rp1- Representations:</strong> Person can represent simple properties of objects, events, and people independently of their own immediate actions.</td>
<td><strong>PARLI Level One- Basic Stated Information:</strong> Reader can represent basic information stated in the text, including general information about what the subject of the piece is.</td>
</tr>
<tr>
<td><strong>Rp2-Representational Mapping:</strong> Person can relate variations in one representation to variations in another representation.</td>
<td><strong>PARLI Level Two- Key Details:</strong> Reader can determine those details that loom large through more than one paragraph and generally without which the cohesion of the piece is lost. These details fit together to form the overall concept of the piece.</td>
</tr>
<tr>
<td><strong>Rp3-Representational Systems:</strong> Person knows how two concepts relate to each other and still only concerned with overt characteristics</td>
<td><strong>PARLI Level Three- Stated Relationships:</strong> Reader can determine the relationship between two ideas, processes, events etc are stated in the text. The reader must have an understanding of each independently to be able to make sense of the stated connection between them.</td>
</tr>
<tr>
<td><strong>Rp4/Ab1-Abstraction:</strong> Person can accomplish the coordination of two representational systems to get an intangible attribute that characterizes broad categories of objects, events, or people.</td>
<td><strong>PARLI Level Four- Simple Inference:</strong> Reader must deal with connotative and denotative clues in the text (representational system 1) and relate them to her own experience and knowledge base (representational system 2) to infer the cued relationship.</td>
</tr>
<tr>
<td><strong>Ab2-Abstract Mappings:</strong> Person can coordinate two complex systems to achieve a strong conceptual understanding of the architecture of the systems.</td>
<td><strong>PARLI Level Five- Complex Implied Relationships:</strong> Reader can determine nonfiction text structures used to convey the message. Reader can distinguish between primary and secondary text structures (from among 6 types) and use these to cue the author’s intent and facilitate greater abstractions.</td>
</tr>
<tr>
<td><strong>Ab3-Abstract Systems:</strong> Person can flexibly differentiate the relationships between two concepts</td>
<td><strong>PARLI Level Six- Author’s Generalizations:</strong> Readers can utilize numerous connotative and denotative clues in either very detailed and complex text, or multiple texts, to derive causes of change and development of individuals, properties and/or processes over time.</td>
</tr>
<tr>
<td><strong>Ab4-Principles:</strong> Person can consider systems of abstract systems and relate groupings of multifaceted concepts to grouping of others across time and conditions.</td>
<td><strong>PARLI Level Seven- Nonfiction Structural Generalization:</strong> Reader can articulate and evaluate the author’s world view and what it implies about human nature or scientific principles (depending on discipline), as it exists outside the text.</td>
</tr>
</tbody>
</table>
Vygotsky’s (1978) *Theory of Social Cognitive Development* has done much to inform educational practice. His original work was done in the context of language learning and posits that the potential for cognitive development in an individual is limited to one’s *zone of proximal development* (ZPD). Particularly germane to the current work is that ZPD claims that students will gain the greatest benefit from instruction that takes place in a supportive environment mediated by tools. Among the tools referenced are anything that helps to organize and generate information to assist the learner, and includes mentors, models, and cognitive strategies. Within social cognitive theory, these tools are owned by the teacher first, and are called upon specifically to provide a scaffold that supports the learner at the upper limits of his or her ZPD and that can be systematically withdrawn as the learner gains competence. Social cognitive theory and the ZPD are complimentary to Dynamic Skill Theory in the explanation of the variability of independent performance often exhibited by learners. The ZPD is the area between the functional and optimal levels of performance depicted in Figure 3.

*Cognitive Flexibility Theory* (CFT) is a theory “designed for learning in ill-structured domains, where cases of knowledge application are characterized *individually* by complexity and *across cases* by considerable variability and irregularity in the conditions of knowledge use” (Spiro, 2004, p. 654). CFT is complimentary to Dynamic Skill Theory (Fischer, 1980), and has direct application to literacy instruction and reading instruction in particular. Reading and teaching reading are ill-structured domains of real-world practice (Spiro, 2004). Each reading task or event is unique and, as such, makes particular demands on the actors. As Spiro so aptly states: “. . . the only summary statement that applies to all of reading is this. *It all depends*” (p. 655). As discussed
earlier, several key elements combine in any reading situation, and each one contributes to the complexity that makes it an ill-structured domain. CFT provides a cognitively flexible stance from which one can analyze and remediate the complex domain of reading, while Dynamic Skill Theory provides the means for setting developmentally appropriate goals and assessing progress toward them in an equally complex and genuine manner that reflects the pluralism inherent in the process of becoming literate across multiple disciplines.

**Research-based instructional strategies**

In the dynamic teaching and learning environment represented by the PARLI heuristic, the interaction of the motivation and engagement gear with the cognition gear in turn drives the instructional strategies gear. According to Marzano et al. (2001) Midcontinent Research for Education and Learning (McRel) conducted a meta-analysis using selected research studies on instructional strategies that fit into a K-12 classroom environment to determine effect sizes of different strategies. This meta-analysis of instructional strategies concluded that there were nine categories that have a strong effect on student achievement: (a) identifying similarities and differences, (b) summarizing and note taking, (c) reinforcing effort and providing recognition, (d) homework and practice, (e) nonlinguistic representations, (e) cooperative learning, (f) setting objectives and providing feedback, (g) generating and testing hypotheses, and (h) questions, cues, and advance organizers. The nine instructional strategies found to have the greatest effect sizes were analyzed with a range of 21 to 1,251 educational studies in each meta-analysis conducted by McRel, with most of the meta-analyses featuring between 120 and 250
Meaningful differentiated instruction is the result of professional teachers using the knowledge base and their specific knowledge of their students and their curriculum to discern the best strategies to use in optimal circumstances to ensure that all students learn at high levels (Tomlinson & Allan, 2000). When considering the complexity of incorporating strategies (a) through (h) in an actual classroom setting populated by students who will be distributed across a spectrum of proficiency for each of these strategies, it is quickly clear that differentiating instruction is the means for creating a classroom in which all students learn. As such, differentiated instruction is shown to be preferable to models that centralize standardization and teacher proofing (Tomlinson, 1999, 2001; Tomlinson & Allan, 2000; Tomlinson & Eidson, 2003). The differentiated instruction research (Tomlinson, 1999, 2001; Tomlinson & Eidson, 2003) points to the potential merit of a flexible framework to differentiate reading instruction for struggling readers in the middle grades; these are students identified as between 6 months and 2 years below grade level placement, based on standardized reading scores. For students who are further behind, a different type of intervention is called for so that they can first focus on comprehension building exclusively: “... there remains a group of middle and high school students who have reading problems that result from not having mastered the alphabetic principle” (Kamil, 2003, p. 9). Students who struggle with reading comprehension in middle school do so for a host of distinct and complex reasons, making it naïve, at best, to anticipate that a one-size-fits-all approach will result in each student
advancing and closing the gap (Greenleaf & Hinchman, 2009; Robb, 2008; Tomlinson, 1999, 2001; Tomlinson & Eidson, 2003; Torgesen et al., 2007).

Research in the field of instructional strategies (Bennett & Rolheiser-Bennet, 2001; Bulgren, Deshler, & Schumaker, 2003; Marzano et al., 2001) supports the full complement of elements of differentiated instruction spelled out by Tomlinson (1999, 2001). In addition, there are many key features of differentiated reading instruction recommended by Robb (2008), like teacher-led reading groups and using writing to support comprehension for every student. All effective differentiated instruction has effective assessment practices as a key component; it is through effective assessment, that teachers know when and how to differentiate instruction. Student-friendly rubrics that support students’ receiving explicit instruction in how to use the rubrics to evaluate their work make for effective assessment practice (Bennett & Rolheiser-Bennet, 2001; Marzano et al., 2001; O’Connor, 2002; Stiggins, 2001).

Further, students who learn to justify their scores with specific examples of their work are ready to take the greatest advantage of learning opportunities presented by the use of rubrics. This combination of student ownership of the assessment process and consistent, ongoing feedback represents the type of powerful integration of assessment within the teaching and learning process that assessment experts are extolling (Bennett & Rolheiser, 2001; Marzano et al., 2001; O’Connor, 2002; Stiggins, 1997). While using rubrics, students must be taught the progressive levels of performance and what distinguishes them (Stiggins, 1997) to be able to obtain the complete benefits of these powerful rubrics.
One area of instructional strategy research that is underutilized in reading and comprehension is the work on concept attainment from the University of Kansas Center for Research on Learning (Bulgren et al., 1997). The original focus of this work was to improve the educational experiences of students with identified learning disabilities. Specifically, the Concept Mastery Routine, which is one of a series of routines from the Content Enhancement Series, offers assurances that students will understand the concepts being taught through the creation of a concrete, non-linguistic representation of a given concept.

The Content Enhancement series was developed as a set of “instructional procedures for teaching concepts in mainstream secondary courses in which students with learning disabilities are enrolled” (Bulgren, Schumaker, & Deshler, 1988, p. 4). They represent a way of teaching an academically diverse group of students that balances individual and group needs with the integrity of the content, by selecting critical features of that content and transforming it in ways that promote student learning. Using the Content Enhancement series, students gained an average of at least 10 to 20 percentage points on tests or tasks that required them to demonstrate learning (Bulgren et al., 1988). Similar to the expansion of differentiated instruction from gifted students to all students (Tomlinson, 1991, 2001; Tomlinson & Eidson, 2003), this Content Enhancement (Bulgren et al., 1997) work that was born of a need to serve students with learning disabilities is also effective practice for all students (Bulgren, et al., 1988). Regarding the application to a hierarchical reading comprehension model (Hillocks, 1980; Hillocks & Ludlow, 1984), the Concept Mastery Routine specifically (Bulgren et al., 1997) helps
students distinguish between levels in ways that further their skills in comprehending what they read.

**Nature of reading research**

Reading for understanding is a complex act and to suggest otherwise does a tremendous disservice to the community at large and to struggling adolescent readers in particular. The necessary complexity of the PARLI framework was addressed in the section addressing the cognitive developmental research, and will be addressed further in this section and. As the PARLI heuristic illustrates, the system of gears must work together to result in movement for the reading-specific gear. Research on the nature of reading is informed by the previous three categories. With the current interest in adolescent literacy, there is research specific to reading from which to draw in developing a middle school literacy instructional framework, none of which is simple.

The PARLI framework was initially inspired by the work of Hillocks. In the early 1980s, Hillocks (1980) focused on the development of an informal reading inventory to guide instruction in Literature and English classes and developed a hierarchical theoretical model of reading comprehension of narrative form. His basic premise began with the notion that before students can deeply analyze text inferentially, they must master literal and basic levels of understanding. Hillocks began his work by creating an informal reading inventory and developed a complete hypothesis about the nature of reading comprehension that incorporates a developmental model of reading with seven progressive levels of understanding, explained in Table 1. Hillocks developed a paradigm with the particular types of questions one would need to ask to discern whether or not an individual understood the text at hand at each particular level of depth. Hillocks and
Ludlow (1984) found that, in fact, the different question types are ordered hierarchically, and that their relationship to each other is taxonomic.

Use of inventories based upon the taxonomic skill levels in the interpretation of fiction articulated by Hillocks (1980), and further developed by Hillocks and Ludlow (1984) yields rich results regarding student comprehension levels that make clear suggestions on how to differentiate reading instruction to meet the needs of a diverse group of learners. These results provide the teacher with evidence from essay responses that show where each student is currently performing along the hierarchy of seven levels of comprehension. These responses enable the teacher to distinguish between novice or immature work, competent work, and masterful work at each level and organize instruction accordingly. For the PARLI framework, these hierarchical levels were specified as follows:

- **Level One-Basic Stated Information:** Reader can represent basic information stated in the text, including general information about what the subject of the piece is.

- **Level Two- Key Details:** Reader can determine those details that loom large through more than one paragraph and generally without which the cohesion of the piece is lost. These details fit together to form the overall concept of the piece.

- **Level Three-Stated Relationships:** Reader can determine the relationship between two ideas, processes, events etc are stated in the text. The reader must have an understanding of each independently to be able to make sense of the stated connection between them.
• Level Four-Simple Inference: Reader must deal with connotative and denotative clues in the text (representational system 1) and relate them to her own experience and knowledge base (representational system 2) to infer the cued relationship.

• Level Five-Complex Implied Relationships: Reader can determine nonfiction text structures used to convey the message. Reader can distinguish between primary and secondary text structures (from among 6 types) and use these to cue the author’s intent and facilitate greater abstractions.

• Level Six-Author’s Generalizations: Readers can utilize numerous connotative and denotative clues in either very detailed and complex text, or multiple texts, to derive causes of change and development of individuals, properties and/or processes over time.

• Level Seven-Nonfiction Structural Generalization: Reader can articulate and evaluate the author’s world view and what it implies about human nature or scientific principles (depending on discipline), as it exists outside the text.

In addition to Hillocks (1980), there are more recent research studies to consider that identify what works best with students at the secondary level. In evaluating what high performing schools did better than average schools, Langer (2001) cites the strategic nature of effective instruction. In Langer’s study, middle and high schools that had been trying to increase students’ learning and performance in English language arts were studied over a five year period.

The study focused on the workings of schools, teachers, and classrooms that strive to increase student performance and, despite obstacles and difficulties of serving the poor, beat the odds on standardized tests in reading and writing; that is, gain higher literacy, beyond comparable schools. (p. 844)
The five year study involved data gathering in successive cohorts in four states. “Each teacher and school was studied for 2 years, permitting extensive study of how patterns in curriculum and instruction played themselves out in schools and classes across time” (p. 844). Challenging tasks, explicit teaching of skills, interconnections among activities, and a careful matching of tasks and instruction to student competence levels were strongly in evidence in the high performing schools. Although Langer’s (2001) study exclusively addressed standard English literacy and fiction specifically, it is logical to conclude that these practices translate to nonfiction as well.

The adolescent reading literature supports an instructional framework with the attributes of the PARLI framework, as exemplified by Greenleaf and Hinchman (2009):

Rather than shielding students from the hard work of academic literacy until they demonstrate the capability to comprehend such texts on their own, actually engaging them in academic reading, with expert teacher support and a collaborative learning environment, is seen as the most important way to build young people’s capability. (p. 10)

The PARLI framework emphasizes the reciprocal nature of reading and writing as the means of developing complex high level academic literacy across content areas. The tendency to focus secondary remediation on the basics sets the stage for these students to remain behind. As Snow and Biancarosa (2003) assert, “A foundation doesn’t make a house, and basic skills don’t make for high-level competence” (p. 2). The authors go on to share that “Without ongoing literacy instruction, students who are behind in reading when they enter the middle grades likely will never catch up” (p. 2). These assertions are supported by the work in developmental psychology and its applications to reading development of Stanovich and others (Cunningham & Stanovich, 1997; Stanovich, 2000, 2008). Stanovich adopted the term the “Matthew Effect” in a 1986 article, recently re-
published in 2008, to draw an analogy to the realm of reading development from the reference in the Bible to the notion that the rich get richer. When applied to reading comprehension, this has been the trend explicated across the adolescent reading literature base, specifically the body of literature focused on struggling or striving readers. In general, when children fail at reading in the early grades, they develop a distaste for reading and disengage from it. As they read less than their peers who are stronger readers, they dig a bigger gulf between themselves and academic literacy competence. The gap that emerges in early literacy is robust and persistent across time, such that by the time these students reach the middle grades, the expression of the “Matthew Effect” is readily apparent.

Instruction that effectively develops literacy across the content areas requires a coordinated, systematic, research-based, dynamic literacy plan that operates with teachers of all content areas playing an integral role. It is through the professional development and use of multiple, coordinated reading comprehension strategies, along with expectations that students will read across the content areas, and time to do so within the daily schedule, that effective instruction will be implemented (Alexander, 2003; Alvermann, 2002; Beers, 2003; Biancarosa & Snow, 2006; Greenleaf & Hinchman, 2009; Harvey & Daniels, 2009; Torgesen et al., 2007; Vacca & Vacca, 2005). Academic literacy is a multi-dimensional domain of its own, with cognitive, motivational and sociocultural forces that interact to build literacy (Alexander, 2003).

In their work with think-aloud protocols, Coté and Goldman (2004) found that correlations between strategies that students reported using and reading comprehension scores suggest that individuals might be differentiated on the basis of the dynamic
interaction of their processing activities and their efforts at gaining a coherent understanding of the text. Individuals fell into one of four categories, based upon analyses of the individual protocols: (a) successful-knowledge building, (b) less-successful knowledge-building, (c) text-focused processing, and (d) minimalists. The data indicate some detail about aspects of monitoring that can inform reading intervention: “Our data indicate that although monitoring is important, unless readers actively apply strategies to resolve the problems they identify, they are likely to end up with fragmentary representations” (p. 678).

Neuroscience research

While the direct application of neuroscience to education is in the early stages, brain development is beginning to inform our knowledge of how the adolescent brain learns to effectively read increasingly complex material. While it is clear that there is no “piece” in the brain that is responsible for reading (Bunge, Klingberg, Jacobsen, & Gabrieli, 1999; Cooke et al., 2001; Cutting, Eason, Young, & Alberstadt, 2009; Dehaene, 2009; Della Chiesa, Davis, Miyamoto, & Momii, 2007, Schmalhofer & Perfetti, 2007; Shaywitz et al., 2004), communication is a quintessentially human thing that is governed by brain function. Accordingly, the theoretical basis for the PARLI framework considers neuroscience to be the belt that moves and coordinates the gears that represent the theoretical components of this instructional framework. Brain imaging clearly shows that the brain grows throughout adolescence until young adulthood (Giedd et al., 1999).

The work of Giedd et al. (1999) demonstrates that proliferation and pruning akin to what happens in toddlerhood, is at play with young adults. Della Chiesa et al., (2007) report that this second wave of pruning affects some of our highest mental functions and
occurs in the late teens. Some of the other key brain developments during adolescence that may contribute to how young adults perform in school are those that come into play with motivation and the choice to engage.

Della Chiesa et al. (2007) conclude that when specifically considering literacy and the teenage brain, it is important to hold three challenging truths in mind: “The brain is biologically printed to acquire language” (p. 85) and “in contrast to language, there are no brain structures designed by evolution to acquire literacy,” (p. 86). However, neuroscientists do know that “learning to read involves connecting two sets of brain regions that are already present in infancy: the object recognition system and the language circuit” (Dehaene, 2009, p. 195). Dehaene’s neuronal recycling hypothesis postulates gradual specialization of the visual system to make predictions at the brain level, but the technology is not at a point that we can see what is going on in the brain as reading develops. In addition, reading improves due to an increased activation of the left occipito-temporal region. This improvement correlates more neatly with reading scores than age, so it is a function of being a reader, not just regular development: “Literacy drastically changes the brain--literally!” (Dehaene, 2009, p. 208). Since literacy at the secondary school level is predominantly concerned with whole sentences, paragraphs, and essays (the exception being students profoundly behind in literacy skills), it is wise to note that limited neuroscientific work has been done at this paragraph level with young adults to date (Della Chiesa et al., 2007).

Among neurophysiological explorations that inform the reading comprehension field, the majority of studies consider reading at the beginning stages of decoding and the micro-level of phonemes and individual words or lists of words. Cooke et al., (2001)
articulated the neural bases for sentence comprehension. They found that a core region of the left posterior superior temporal cortex “plays a central role in sustaining comprehension that is common to all sentences,” (p. 80). Further, they found distinct activation patterns in the right temporal region associated with contrasts of different types of sentences. In addition, the left inferior frontal cortex demonstrated interaction effects being recruited under specific sentence contrasts and not others. The researchers were able to attribute this activation of left inferior frontal cortex to extra memory cost associated with syntactically more complex sentences. This evidence of variability in processing when only considering the sentence level is suggestive of the likely complexity (on a neurological level) of processing full-length texts in the complex and dynamic environment of a secondary classroom. Shaywitz et al., (2004) found that an evidence-based phonologically mediated reading intervention “brings about significant and durable changes in brain organization,” (p. 931). This intervention resulted in brain activation that resembles typical readers in the appearance of the left occipitotemporal area and improvement in reading fluency. The work of Shaywitz et al., (2004) lends further support to both neuroplasticity and the potential for research in neurological functioning to effectively inform instruction.

Reading comprehension is a complex process. As such, it makes significant demands on cognitive processing in general, and working memory in particular. Bunge et al. (1999) investigated two ways that the brain might recruit additional resources to do two things at once. This research supported a resource model in which “Resources may be recruited from new areas specialized for dual task-specific processes, such as task coordination, that are not invoked by either component task,” (p. 3573). Their findings
that dual-task performance is brought about through “increased activation in brain regions that subserve performance of the component tasks,” (p. 3577) points to potential sources of diminished comprehension in individuals with working memory deficits. Nevo and Breznitz (2011) explain how working memory’s four components (central executive, phonological loop, visuospatial sketchpad, and episodic buffer) represent a control system with limited capacity and processing capabilities. This study’s focus on determining “the effect of working memory components--singularly and in combination--on reading abilities (p. 75) is instructive in this discussion of how neuropsychological constructs can inform reading comprehension instruction and remediation. The study found a difference between the contributions of phonological memory and visuospatial memory in reading achievement in general and decoding specifically with the capacity measure of phonological complex memory showing the greatest contribution to variance.

Cutting et al. (2009) focus on the category of students with specific deficits in reading comprehension who do not exhibit deficits in decoding. Research on the specific reading comprehension deficit (S-RCD) individuals supports the suggestion that areas outside the word-level and language need to be considered when trying to develop understanding and action plans to address the problems of these readers. Cutting et al. cite a number of researchers beginning to explore the neurobiological correlates of text comprehension that makes up the daily work of secondary school. Ferstl, Rinck and von Cramon; Jobard et al.; Karunanayaka et al.; Schmithorst, Holland, and Plante; Virtue, Haberman, Clancy, Parrish, and Beeman; Wilson, Molnar-Szakacs, and Iacoboni; and Xu, Kemeny, Park, Frattali, and Braun (as cited in Cutting et al., 2009) all explored these neurobiological correlates of paragraph comprehension, yielding results that, “may prove
fruitful in terms of understanding abnormalities in not only sentence-level but also discourse-level processing in various readers types,” (p. 201).

Coordinating constructs from cognitive models of comprehension and neuropsychology can create linkage to specific neurophysiology that may be responsible for children struggling to comprehend text. As the access to neuroimaging expands with developments that allow more flexible and complex research settings, the potential to learn about the brain circuitry of impaired readers in ways that may point to actions for remediation also expand. While it is not possible to know what is not yet known, structuring future neuroimaging studies with struggling readers interacting with stimuli that represent different types of comprehension may be especially productive in growing an understanding of how readers process the variety of texts one must master to be highly literate.

Cutting et al., (2001) highlight that while standard thinking has long been that decoding problems consistently precede comprehension struggles, there is an alternative hypothesis that one could demonstrate difficulties with reading comprehension primarily resulting from weakness in processes outside of word-level difficulties. Many students who fall into the S-RCD category exhibit what is categorized as late-emerging reading disability. To date, this area has not been examined at great depth in the reading field. That said, Cutting et al., (2001) support the plausible hypothesis that these underlying processes were always faulty in these readers, but the weakness did not become apparent until the demands of deeper reading taxed the system to reveal it.

Knowledge gained about the neurological correlates of observed reading difficulties has the potential to lead to diagnostic tools that would enable identification of
individuals with a variety of deficits in executive function in preschool. Early identification, in turn, creates an increased opportunity to intervene before de-motivating patterns of failure are established. Additionally, early intervention is more likely to capitalize on the known plasticity of the brain. Encouraging this hope, Foorman, Francis, Shaywitz, Shaywitz, and Fletcher (1997) provided data in support of the wisdom of early intervention as the best strategy for remediating reading disability. They found that 82% of remedial children developed into successful readers when effective intervention is provided in the early grades. Unfortunately, this percentage plummets to 10-15% in the later grades.

While the specifics of neurological development are far from providing explicit direction, the one aspect of applications of neuroscience to education that has particularly encouraging early findings is that new neurobiological evidence points to the fundamental role of emotion in cognition (Immordino-Yang & Damasio, 2007). Neuroscience is providing the adolescent reading field with biological evidence for what cognitive and behavioral scientists, and remedial reading teachers have long thought to be the case: learning is an emotional endeavor and learning to read involves complex cognitive processes.

Ultimately, by blending educational models with cognitive models of reading comprehension within a neuropsychological framework that includes neurobiological measures, the prospect arises of discovering and understanding the brain-behavior connections that govern reading comprehension (Schmalhofer & Perfettiti, 2007). This transdisciplinary approach provides a lens for teasing out the origin of reading disability and the means to prevent and correct crippling reading comprehension deficits.
Summary of the Chapter

Chapter 2 has provided a review of the literature of the complex and dynamic research that focuses on struggling adolescent readers. By examining the issues the United States is facing regarding adolescent literacy, and some of the large-scale efforts being undertaken to address them, the areas in need of adolescent literacy research and focus are apparent. They include the need to understand how adolescents’ reading comprehension strategies apply to a wide range of comprehension tasks. These areas also encompass the development of reading assessments to guide teachers’ efforts to address the particular needs of students who struggle with reading comprehension. Another key area is linking measures of motivation and engagement effectively with reading comprehension instruction. Reading comprehension is a profoundly complex endeavor for novice and expert alike. Research into struggling adolescent learners and the reading task itself both inform the future direction of the field.

Chapter 3 explains the PARLI framework and the progression of learning and instruction through it. Chapter 4 features the detailed methods used to evaluate the PARLI framework’s effectiveness and make recommendations about how to continue to develop it for future use. Chapter 5 presents the findings of the research and, finally, Chapter 6 is a discussion of the findings.
Chapter 3: The PARLI Framework

The focus of this research was the evaluation of the PARLI framework. As such, this chapter provides a brief overall description of the implementation of the framework. In addition, the research basis for the development of the framework is explained.

Development of PARLI

The Pragmatic Analytical Reading Level Inventory (PARLI) was theoretically based on a blending of educational research in reading, effective instructional strategies, cognitive developmental research, engagement and motivation research, and the emerging support of a neurological basis for learning processes and the emotional components of learning. Specifically, Hillocks’ (1980) hierarchy of reading comprehension for fiction, Fischer’s (1980) Dynamic Skill Theory from the cognitive research domain, and Immordino-Yang and Damasio’s (2007) work with connecting affective and social neuroscience to education inform the development of the framework. The framework also reflects the findings of leading reading researchers who specialize in adolescent literacy (Alvermann, 2002; Biancarosa & Snow, 2004; Hillocks, 1980; Jacobs, 2008; Kamil, 2003; Langer, 2001; Pressley & Afflerbach, 1995, Schoenbach et al., 1999; Slavin et al., 2008; Snow, 2002; Spiro, 2004; Torgesen et al., 2007; Vacca & Vacca, 2005), cognitive development research (Spiro, 2004; Vygotsky, 1978), effective instructional strategies research (Bulgrren et al., 1988; Bulgren et al., 1997; Kamil, 2003; Marzano et al., 2001; Tomlinson, 1999, 2001; Torgesen et al., 2007; Stiggins, 2001), and motivation and engagement research (Bandura, 2006; Ford, 1992; Guthrie, 2004; Guthrie et al., 2007).
Implementation scheme for PARLI

The seven hierarchical levels of reading comprehension described by Hillocks (1980) closely parallel Fischer’s (1980) model for describing developmental and learning patterns observed over time and across disciplines. When they are taken together, they form a structure that makes the complex task of understanding what is read at increasingly deep and complex levels, that is, being highly literate, clear, and approachable. PARLI involved making the hierarchy concrete for students, then building comprehension by guiding them through the hierarchy using discussion and writing. It started with a diagnostic assessment to establish current levels of reading and writing. With this data regarding current levels of reading and writing shared with students, PARLI focused on the explicit teaching of the nature of each of the first several levels in the hierarchy. This teaching was differentiated by student need while using familiar and easily navigable picture books (Albright, 2002; Ammon & Sherman, 1996; Harvey, 1998; Moore, Alverman, & Hinchman, 2000; Rief, 1992; Robb, 2000; Yokota, 2001).

Instructionally, each PARLI level was broken into three stages in order to provide further scaffolding as students moved from being novice readers in comprehending and responding at a particular depth of understanding, through mastery reading at that level. Breaking the comprehension levels into these stages was based upon the Dynamic Skill Theory model (Fisher, 1980) that suggests how complex skill is built. The PARLI framework integrates Hillock’s levels (1980) with Fischer’s skill scale levels and tiers from Representations to Abstractions (Fischer & Bidell, 2006) with the establishment of Stages A through C at each of the levels of comprehension.
The PARLI framework incorporated many of the strategies that Marzano et al., (2001) found to be effective. The most effective literacy strategies are comparing and contrasting, summarizing, setting objectives and providing feedback, and generating and testing hypotheses (Marzano et al., 2001). All of these strategies were essential in structuring responses within the PARLI framework; the complexity of the academic task using each of these strategies builds progressively as students move through the hierarchy. For example, at Level 3, students set the objective to be able to demonstrate their understanding of stated relationships through their effective literature response, measured against a performance rubric for that level. One of the key types of relationships being explored involves the author performing the task of comparing and contrasting. At Level 3 and subsequent levels students were coached to first summarize the piece up to the point of the current response. Then they were to explain the nature of the stated relationship. At the inferential levels of the PARLI framework hierarchy, students were called upon to regularly develop hypotheses about the author’s meaning, and test these hypotheses through their connections between the text and their knowledge of the world outside of the text. At these higher levels in the hierarchy, students were coached to engage in substantive comparing and contrasting of character development and change over time, as well as between various characters and other story elements.

In the PARLI framework, practice moved from picture books to short stories and continued with full-length student-choice fictional texts, followed by nonfiction narratives. At each stage of instruction, cooperative learning structures that have been found to be successful in advancing adolescent literacy (Alvermann, 2002; Slavin et al., 2008) were implemented. The process was then repeated to address issues of nonfiction
EVALUATING THE PARLI FRAMEWORK

expository and descriptive texts, and moved back up in the same progressive fashion through texts of increasing length and complexity.

**Explaining the progression of the PARLI framework.** The PARLI featured 7 levels (1-7) with 3 stages (A, B, C) at each level. Students began responding at the level at which they were assessed, unless they were having difficulty. When a student was experiencing difficulty, he or she was coached to activate appropriate reading strategies and offered scaffolding as needed, including revisiting of earlier levels of the hierarchy as a prompt for deeper comprehension. The progress was as follows:

- **Stage A** involved students writing a response, including their opinion, at the level on which they were working. For example, Level 3, Stage A involved a student noticing and writing about two ideas or events that the author clearly stated were linked.

- **Stage B** involved the students providing direct evidence from the text to support their thinking. This direct evidence was in the form of a properly cited quote.

- When students reached Stage C, they were establishing proficiency at that level and were requested to generate a question that would elicit an appropriate Stage B response at this level, if they were to approach another reader with the question. A Stage C response still required the complete Stage B response as well.

- When students reached Level 4, and in all subsequent levels, they were expected to integrate their evidence at Stage B into their response, rather than being allowed to just tack it on at the end.

The process of using PARLI went through a series of simultaneous and distinct steps to build reading comprehension. However, because the framework used a model of differentiating instruction, these steps were recursive for some students, as needed to ensure the advancement of their academic literacy skills.

The natural flow of the learning process for various students in a given classroom is not always linear. Because student engagement and metacognition are non-negotiable
elements of PARLI, students were given the Metacognitive Awareness of Reading Strategies Inventory ([MARS] Mokhtari & Reichard, 2002) and the Middle School Assessment of Academic Self-Concept and Motivation ([MAASC] adapted with permission from Gordon Rouse & Cashin, 2000) to provide a baseline to work, and evaluate growth, for each student. Next, the PARLI framework used the first subtest of the Pragmatic Analytical Reading Level Assessments (PARLA-NARR) to assess current comprehension levels along the hierarchy for narrative works (fiction). When individual scores were available, instruction moved into teaching the hierarchy with scaffolded instruction (Vygotsky, 1978) using a variety of materials, including picture books (Albright & Alriail, 2005; Costello & Kolodziej, 2006; Dean & Grierson, 2005; Murphy, 2009). At this time, several things were happening simultaneously. Students were introduced to the Student Guide that scaffolded their work with reminders of the nature of each level and stage, including sentence starter prompts for responses, examples of complete responses, examples of rubrics, and examples of justification of scores. In addition, use of the *University of Kansas Concept Mastery Routine and Concept Diagram* (Bulgren et al., 1997) to teach differences between the various levels of reading comprehension was incorporated in a differentiated instructional model (Tomlinson, 1999, 2001).

The introduction to the framework involved familiarizing students with what is known about how the human brain learns. Through this knowledge, each individual began to take ownership of his or her own brain, and actively develop the executive function skills that contribute to reading comprehension. These students in middle school struggling to comprehend text had experience with a variety of instruction that was not
successful for them. By teaching them about how the brain learns and how the PARLI framework used this as its foundation, students were able to recognize PARLI as substantively different from earlier, unsuccessful instruction. This knowledge, in turn, gave them hope that they would now be successful. The reasonable response to this hope was engagement in the process that began to result in improved reading comprehension, creating improved confidence, which reinforced and increased engagement.

Once the initial introduction was complete, whole class instruction at the first several levels, based upon PARLA-NARR data, took place. This was followed by small group instruction, with groups formed by current levels, using picture books. At this point in Stage A, questions were provided for students at whichever level they began their work. The importance of effective and timely assessment and feedback is an essential element to learning. This model placed a priority on building autonomy. Instruction and practice using rubrics began with the initial work in responding. When students gained some comfort working collaboratively with picture books, the move was made to short stories to continue to gradually build their comprehension skills in a trajectory that matches what Dynamic Skill Theory (Fischer, 1980) conveys regarding the development of complex skills. This practice began with the mystery and detective genres because they are both of high interest to adolescents and place a premium on the ability to infer (Pollock & Chun, 2008). It is important to offer text that supports the need for the cognitive work at ever-deepening levels of understanding that the PARLI is designed to develop. Students chose from among a number of stories, as well as whether they preferred working in small groups, in pairs, or individually. While reading a series of these short stories, students were moving through the stages of the PARLI model.
When students were proficient at Stage B of the first level they were then coached to increase their metacognitive work and skills by justifying the scores they gave themselves on the rubric. From this point forward, students were expected to always score their own work and justify those scores in writing, using the language and concepts of the rubric, before submitting work for teacher or peer feedback. In addition, students were encouraged to explore developing responses based on what they noticed in their reading, rather than depending on the provision of particular questions at each level and stage by the instructor; this strategy was in line with the goal of reading independence.

After completing several short stories (this was differentiated based on student readiness), students formed literature circles (Daniels & Steineke, 2004) to select a novel to read together. Students worked in small groups with the teacher to develop a schedule of reading and responding in discussion and written format, and created a contract that specified the number and type of responses expected.

Throughout, instructional time was spent in whole-class, small group, or individually, with coaching for specific, research-based reading strategies on a differentiated basis (Beers, 2003, Harvey & Goudvis, 2007; Robb, 2008). Students were coached to reflect on the strategies used and their effectiveness in specific contexts, such as using context clues, read/pause/retell/read on or reread, predict and support, making connections, and so forth. The use of coaching students in before, during, and after reading strategies such as these worked to harness the potential of increasing metacognitive awareness (Abromitis, 1994; Alexander, 2003; Vacca & Vacca, 2005).

After completing a novel, students were ready to move on to nonfiction texts. Students began work with nonfiction by taking a formative Pragmatic Analytical Reading
Level Assessment for Narrative Nonfiction ([PARLA-NARR NF], Appendix A). When the current comprehension level for nonfiction narrative was established, students repeated the process, starting again with picture books in the memoir and biography genres. The instructional unit concluded with a summative assessment (Appendix A).

Upon successful completion of the narrative nonfiction texts, students moved to Social Studies and expository text. Again, they began with assessment (PARLA-EXPO), and started their comprehension building work with picture books featuring expository text. Continuing in the established pattern, based on time and interest, students worked with increasingly complex texts, and ended the unit with a summative assessment (Appendix A).

Summary of the Chapter

The PARLI framework synthesizes the work of neuroscientific research with cognitive research and studies of best practices in educational instructional strategies, while emphasizing the power and importance of engagement and motivation, particularly with struggling adolescent readers. The PARLI framework acknowledges the very complexity of the ill-structured domain that is adolescent literacy. This framework incorporates Spiro’s (2004) recommendation for the use of Cognitive Flexibility Theory to inform the work of developing highly literate adolescents.
Chapter 4: Methods

Methodology overview

This chapter reports the research design and methodology used in the case study evaluation of the Pragmatic Analytical Reading Level Instruction (PARLI) framework. In this study, a pragmatic approach using mixed methods of both quantitative and qualitative research in a naturalistic setting was used. Pragmatism provides a paradigm that sets aside the purist paradigms of both qualitative and quantitative research traditions (Johnson & Onwuegbuzie, 2004; Tashakkori & Teddlie, 2003). Purists in both of these paradigmatic traditions hold to a thesis that the two are inherently incompatible and their associated methods cannot and should not be mixed (Howe, 1998). Mixed methods research represents a third research paradigm. To be pragmatic is to be practical. A pragmatic lens leads the researcher to consider whatever methods, in whatever combination, will yield the strongest understanding. The pragmatic research paradigm posits that “both quantitative and qualitative research are important and useful” (Johnson & Onwuegbuzie, 2004, p. 14). Tashakkori and Teddlie (2003) define pragmatism in the mixed methods paradigm: “Pragmatism rejects the either/or choices associated with the paradigm wars, advocates for the use of mixed methods in research, and acknowledges that the values of the researcher play a large role in interpretation of results (p. 713).”

Merriam (2009) discusses seven common types of qualitative research strategies that are commonly used in education: basic qualitative research, phenomenology, grounded theory, ethnography, narrative analysis, critical qualitative research, and case study.
The employment of a case study research design, incorporating elements of grounded theory in data analysis, was a natural fit for this research that sought to “retain the holistic and meaningful characteristics of real-life events” (Yin, 2003, p. 2). To be a case study, this study needed to have a distinct, bounded unit of analysis. Specifically, in the current study, the unit of analysis needed to be either the individual students over the course of the implementation of the framework, or the entire remedial reading class within this same bounded timeframe.

Stake’s (2005) illustration of what is and is not a case is helpful in understanding the current study’s design. Stake makes the distinction between a doctor and his or her doctoring: “A doctor may be a case. But, *his or her doctoring* probably lacks the specificity, the boundedness, to be called a case,” (p. 444). The two units of analysis of the current study were the particular students and the particular classroom. Furthermore, in case study research, the researchers gather a variety of data to help them fully understand the case. Stake (2005) shares that the data in case studies include observations and artifacts of functioning and interviewing others to gather information about what the researcher cannot observe for him or her self. The current study was designed to gather a variety of data, both qualitative and quantitative in nature, about the key participants. This study sought greater understanding of the complex factors that contribute to the development of reading comprehension among struggling adolescent readers. These complex skills and the stages of development of the participant group combined to result in the strongest opportunity for understanding to be accomplished through an interpretive task utilizing case study methodology as the research framework. This study incorporates
the rigor and trustworthiness of mixed methods research design, its implementation, and
the processes thereof.

Merriam (2009) makes a clear distinction between basic and applied research,
stating that the latter “is undertaken to improve the quality of practice of a particular
discipline” (p. 3). Merriam (2009) continues elaborating that a particular form of applied
research is evaluation studies. Further, she explains that the difference between
evaluation and research resides more in the questions asked, and not necessarily the
methods used. The present study was applied research in reading comprehension
designed to evaluate an instructional framework for remedial readers. Merriam quotes
Patton: “When one examines and judges accomplishments and effectiveness, one is
engaged in evaluation. When this examination of effectiveness is conducted
systematically and empirically through careful data collection and thoughtful analysis,
one is engaged in evaluation research” (as cited in Merriam, 2009, p. 4). The current
empirical and systematic study was, by this criterion, evaluation research.

Hypothesis and Research Questions

The purpose of this study was to conduct a research and development evaluation of
the Pragmatic Analytical Reading Level Instruction (PARLI) curriculum framework with
struggling adolescent readers.

**Hypothesis.** The hypothesis of this study were that participation in the PARLI
curriculum framework for one academic semester would result in growth of reading
comprehension among struggling readers.

**Research question.** The research question is: *Is the Pragmatic Analytical
Reading Level Instruction (PARLI) framework effective with struggling middle school*
readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

**Sub-questions for the PARLI evaluation case study.**

1. *Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*
2. *Do students participating in PARLI report a shift in agency and motivation?*
3. *Do students participating in PARLI demonstrate improved metacognition?*
4. *Does student performance on the assessments form a pattern of development?*
5. *How can measurement tools, including observations, used with struggling readers result in better understanding of these students and how to optimize their learning opportunities?*

**Research Design**

**Mixed methods evaluation case study.** This study combined multiple measures over an intervention period of one school semester to determine change in reading comprehension among the participants. The researcher used a mixed method case study design, using both quantitative and qualitative methods (Stake, 2005; Teddlie & Tashakkori, 2009; Yin, 2003), with key elements of expansion evaluation design incorporated as well.

**Case study.** A qualitative case study is defined by Merriam (2009) as “an intensive, holistic description and analysis of a bounded phenomenon such as a program,
an institution, a person, a process, or a social unit” (p. x). Stake (2000) further delineates the nature of case study research:

... neither new nor essentially qualitative. Case study is not a methodological choice but a choice of what is to be studied. ... By whatever methods, we choose to study the case. We could study it analytically or holistically, entirely by repeated measures or hermeneutically, organically or culturally, and by mixed methods. (p. 443)

Yin (2003) provides more specific boundaries for case study as a form of inquiry used to understand social phenomena and “retain the holistic and meaningful characteristics of real-life events” (p. 2). The case study draws its strength from its ability to deal with a tremendous variety of evidence, and is “generalizable to theoretical propositions and not to populations or universes” (Yin, 2003, p.10). Thus the goal of case study is to expand and generalize theories rather than to enumerate frequencies. Information can be gained from case studies by comparing how this case is both similar to and different from other cases (Stake, 2000). However, Stake also cautions that while case studies can certainly provide insight into the human condition, readers often are too quick to accept this insight: “The case researcher needs to provide grounds for validating both the observation and the generalization” (p. 456). Yin (2003) defines case study:

A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. The case study inquiry copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result benefits from the prior development of theoretical propositions to guide data collection and analysis. (pp. 13-14)

Finally, qualitative case studies utilize six sources of evidence: (a) documentation, (b) archival records, (c) interviews, (d) direct observations, (e) participant-observation, and (f) physical artifacts. The current study incorporated five of these types, with
traditional interviews not being part of the evidence. This case study was bounded by several contexts: the eighth graders identified as struggling readers themselves, their experiences in a remedial reading class, and their experiences in their Communication Arts class.

The current study can be considered a multicase study because as Stake (2006) explains:

In multicase study research, the single case is of interest because it belongs to a particular collection of cases. The individual cases share a common characteristic or condition. The cases in the collection are somehow categorically bound together. They may be members of a group or examples of a phenomenon. (pp. 5-6)

Multicase describes this study because there were 10 struggling eighth grade readers participating in the research, each representing a case, and belonging to a collection of cases of adolescent remedial readers. Because of the complex characteristics of the subject of this research (adolescent literacy development and reading comprehension), the pragmatic nature of this study was grounded in the field of mixed methods research.

Reading comprehension is a profoundly complex construct. When the issues of struggling adolescent readers are contemplated, and a desire to evaluate and understand the relationships and resulting interactions between these contexts, hierarchical models of skill development and reading, metacognition, and self-concept and motivation, in the natural complex environment of a middle school remedial reading classroom throughout the course of a school semester, the richly complex understanding possible through case study was a good fit. A more comprehensive understanding of these adolescent readers
and how to develop their reading comprehension skills was sought through this methodology.

The current study was not exploring the culture of a particular society (ethnography). According to Fetterman (1998), “Ethnography is the art and science of describing a group or culture” (p. 1). The ethnographer’s goal is to understand and describe “a social and cultural scene from the emic, or insider’s perspective,” (p. 2). Ethnography is a descriptive approach that Fetterman characterizes as using a theoretical model to guide the work and involves the researcher spending significant time immersed in the culture. While the current study borrowed some methodological strategies from ethnography, it was not guided by a theoretical model or delving into a group or culture as its primary aim.

While the researcher borrowed some methodological strategies from Grounded Theory (Corbin & Strauss, 1990), it was not the aim of this work to build substantive theory about how individuals become struggling adolescent readers. This study utilized the constant comparative analysis strategy from Corbin and Strauss’ Grounded Theory. Constant comparative analysis involves taking one piece of data and comparing it with all others. The research compared both similar and different data. The objective was to develop conceptualizations of the possible relations between various pieces of data. However, to truly be Grounded Theory, the goal of the research must be focused on the development of substantive theory. Further, in Grounded Theory (Corbin & Strauss, 1998), the researcher approaches the work with an interest in the subject of study and without conducting an extensive review of literature that may bias the development of theory. This study began with an extensive review of the literature to analyze what is
known and not-yet-known about effective instruction with struggling adolescent readers and construct and evaluate a framework to address these needs.

Furthermore, according to Merriam (2009), qualitative research that does not fit neatly into one of the other types (phenomenological, grounded theory, narrative analysis, critical, or ethnographic study) and whose aim is to understand “(1) how people interpret their experiences, (2) how they construct their worlds, and (3) what meaning they attribute to their experiences” (p. 23), is Basic or Generic qualitative research. “The overall purpose is to understand how people make sense of their lives and their experiences” (p. 23). This present study’s aim was to evaluate an instructional framework for struggling adolescent readers by delving into the complexity of several individual cases, as well as considering the overall case of this remedial reading classroom, such that it did not fit with a Basic qualitative research study design.

**Mixed methods research.** According to Greene, Caracelli, and Graham (1989), expansion mixed-methods evaluation designs seek to “extend the breadth and range of inquiry by using different methods for different inquiry components” (p. 259) and are generally executed by using qualitative methods to assess the program processes and quantitative methods to assess the outcomes. The current study used an expansion mixed methods design; Figure 7 is a graphic representation of this design. A discussion follows of the mixed methods research approach and how it was employed in this “one-phase design in which researchers implement the quantitative and qualitative methods during the same timeframe and with equal weight” (Creswell & Plano Clark, 2007, pp. 63-64). Creswell and Plano Clark define mixed methods research:
Figure 7: Study diagram. This figure illustrates the study model based on a blend of the Triangulation Design: Convergence Model (Creswell & Plano Clark, 2007, p.63), and the Graphic Illustration of Fully Integrated Mixed Design (Teddle & Tashakkori, 2009, p. 157).
**Mixed methods research** is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone. (p. 5)

Creswell and Plano Clark (2007) reference three stances or world views, each with philosophical assumptions, that are discussed in the mixed methods literature. The first is that pragmatism is the one best paradigm or worldview that fits mixed methods research. The second is a dialectical perspective that posits that “multiple paradigms may be used in mixed methods research; researchers must simply be explicit in their use” (p. 27). The third philosophical stance from the mixed methods literature is that “worldviews relate to the type of mixed methods design and may vary depending on the type of design,” (p. 27).

Greene et al. (1989), through their empirical review of 57 mixed-methods evaluation studies found that “authors’ stated primary or secondary purpose for using a mixed methods design was often triangulation (23%) or expansion (26%)” (p. 260). They found that while the motivation for these evaluations that include both process and product components was a desire to produce a more comprehensive evaluation, “there was a paramedic quality to the qualitative component” (p. 269). “Paramedic quality” is explained as a study wherein the qualitative data was brought in to “resuscitate what was either a failed program or a failed evaluation” (p. 269). In contrast, this research strives to realize the potential of mixed-method expansion studies by employing a more integrated
combination of qualitative methods to assess both implementation and outcomes, incorporating elements of triangulation design. Greene et al. (1989) envision this “higher order expansion design” (p. 269) as one that blends methods creatively in a design that assesses “conceptual strands that span or link program implementation and outcomes” (p. 269), resulting in strengthened inferences.

The hypothesis as described above was tested, and the primary research question answered using an integrated analysis of the mixed data. While some of the research sub-questions have been answered using primarily either qualitative or quantitative data, the majority of the questions are addressed using mixed data:

The sub-questions were:

1) *Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*

   a) Quantitative data, collected by means of standardized reading tests and teacher-made assessments and surveys, and qualitative observations were all used.

2) *Do students participating in PARLI report a shift in agency and motivation?*

   a) Quantitative data collected by means of surveys were used. Qualitative data in the form of observations, qualitative think-aloud protocols, and qualitative learning reflection logs were employed.

3) *Do students participating in PARLI demonstrate improved metacognition?*

   a) Quantitative data, collected by means of researcher-developed reading assessments, and qualitative data in the form of qualitative observations,
qualitative think-aloud protocols, and qualitative learning reflection logs were employed. General student work throughout the course of the study was analyzed using mixed methods.

4) *Does student performance on the assessments form a pattern of development?*

   a) Quantitative data, collected by means of standardized reading tests and teacher-made assessments, surveys, qualitative data from qualitative observations, and qualitative learning reflection logs, and mixed data from student artifacts were all used.

5) *How can measurement tools, including observations, used with struggling readers result in better understanding of these students and how to optimize their learning opportunities?*

   a) Qualitative data from observations, artifacts, think-aloud protocols, and learning reflection logs were used.

In collecting, analyzing and interpreting both quantitative and qualitative data to answer the research questions, this case study evaluation provides results and insights on both the products and the processes of implementation of the PARLI framework with struggling adolescent readers.

Teddlie and Tashakkori (2009) put forth a model of a fully integrated mixed methods design. This model represents the continuing evolution of mixed methods research. In this model, there are four basic stages of the research process: (a) the conceptualization stage, (b) the experiential stage for methodological work, (c) the experiential stage for analytical work, and (d) the inferential stage. Integration may take place during any of them. At the conceptualization stage, the questions are formulated.
According to Teddlie and Tashakkori, integration at this stage (represented by the dashed arrows in Figure 7) involves the formulation of the Quantitative (QUAN)-oriented question informing the formulation of the Qualitative (QUAL)-oriented questions and vice versa. In the current study, the questions were developed in an iterative fashion. The experiential stage was made up of both methodological and analytical sub-stages. To be integrated at the methodological stage is to have both types of data being collected in a way that is not purely sequential or simultaneous. In the current study, there was a blending of sequential and simultaneous data collection of both qualitative and quantitative forms. When it comes to the analytical part of the experiential stage, integration involves quantitizing qualitative data and analyzing it statistically as well as qualitizing quantitative data and generating narrative profiles. For the current study, this integrated analysis is explained in the Data Analysis section. Finally, Teddlie and Tashakkori (2009) use an example to provide insight into how the integration at the inferential stage results in meta-inferences: “the two major QUAN and QUAL strands, and their crossover analyses, directly influenced the formulation of the meta-inferences, which resulted in a dozen or so major conclusions, each involving triangulated data” (pp. 157-158).

Convergent triangulation design. The current study endeavored to engage in this level of integration, and made use of elements of convergent triangulation design in the study and data analysis. Creswell and Plano Clark (2007) assert that this convergence variant of the triangulation design model involves the researcher collecting and analyzing quantitative and qualitative data separately on the same phenomenon. Because this study blended the Teddlie and Tashakkori (2009) and the Creswell and Plano Clark (2007)
models, for clarity sake, it was divided into two methods and results sections for each of the cases, one for the quantitative analysis and one for the qualitative analysis. The two sets of findings were synthesized into a single discussion section (Creswell & Plano Clark, 2007) that contains the meta-inferences made possible by the integration of collection and analysis through the experiential and inferential stages of the study (Teddlie & Tashakkori, 2009). Findings were represented in the form of statistical results for the quantitative data, and in the form of narratives of the cases and interpretive commentaries for the qualitative data. Data was integrated to develop a more comprehensive description of each of these struggling readers (cases) and their growth across the course of the intervention: “The purpose of this model is to end up with valid and well-substantiated conclusions about a single phenomenon” (Creswell & Plano Clark, 2007, p. 65). Figure 7 shows this blended diagram for the current study.

Evaluation research. Evaluation research differs from other forms of research in three key ways: (a) it is usually initiated by a need to make a decision; (b) it is typically conducted for a specific purpose; and (c) its intent is to yield data about the merits of a particular academic phenomenon, rather than unearthing the essential characteristics of something (Gall, Gall, & Borg, 2007). Throughout the educational evaluation literature, there is support for mixed-methods designs. Luo and Dappen (2005) conclude that the multi-dimensional nature of evaluating an instructional model or framework results in the logical conclusion of “multiple ways of knowing and acting in evaluation because educational problems are increasingly complex and intractable” (p. 110).

Further, specifically applying a mixed-methods approach to program evaluation studies allows the researcher to validate program implementation, investigate program
outcomes, and provide a level of complexity and detail that improves the trustworthiness and validity of the findings (Greene et al., 1989). In evaluating a complex educational program, it is seldom adequate to know just that it was effective; understanding the interplay of process and product and replication of positive results are a key goal. Therefore, the incorporation of mixed-methods provides the process evaluation that enhances construct validity regarding the treatment’s effect.

Greene et al. (1989) conducted their theoretical and empirical reviews of mixed-method evaluations, and discovered seven characteristics of these designs: (a) methods, (b) phenomena, (c) paradigms, (d) status, (e) implementation independence, (f) implementation timing, and (g) study. The authors’ descriptions of each of these characteristics, along with the application to the empirical results to the expansion mixed-methods purpose, inform this design. Briefly, methods are the degree to which the qualitative and quantitative measures selected differ regarding form, assumptions, strengths, limitations, and biases. These differences are found in the Instruments Used in Data Collection section of this chapter. The degree to which each method type is intended to measure something different or the exact same thing is considered within the characteristic of the phenomenon. The measures that evaluate reading comprehension were meant to measure the same phenomenon for both the qualitative and quantitative measures. The other phenomena being examined in this study were each measured using the method that best fits the phenomenon as specified in the Instruments and Analysis sections of this chapter. “The design characteristic labeled paradigms refers to the degree to which the different method types are implemented within the same or different
paradigms” (Greene et al., 1989, p. 264). In this study, the single paradigm was pragmatic.

Status is one of the more straightforward characteristics and represents the degree to which one method is more important in the analysis than another. Qualitative methods were emphasized over quantitative in this study, but both were essential. Implementation independence is how much the two types of methods are integrated or independent, while implementation timing is specific to whether they are implemented concurrently or sequentially. In this study, the methods were predominantly independent during data collection, became integrated during analysis, and were integrated during the inferential stage and the timing was predominantly concurrent. Finally, the characteristic of the study describes whether the research encompasses more than one, or just a single study. The current work was a single research study. The empirical review showed that expansion studies using qualitative methods to evaluate processes and quantitative methods to assess outcomes were the norm within a single study. Greene et al. (1989) express their belief that these expansion studies “have not yet tested the limits of their potential” (p. 269) and extol the potential virtues of a more integrated use of methods by combining qualitative and quantitative methods to assess both implementation and outcomes, as in the current study.

**Participants and context.** This study was carried out using purposeful sampling and, specifically, typical case sampling methods (Patton, 2002) to evaluate students in a specific Midwestern suburban middle school undergoing a demographic shift, and in a typical group of struggling adolescent readers in one remedial reading classroom. Student participants selected their own pseudonyms in place of their names in the study. In
addition, the students were enrolled in a course taught by the researcher. Patton explains typical case sampling: “the purpose of a qualitative profile of one or more typical cases is to describe and illustrate what is typical to those unfamiliar with the setting--not to make generalized statements about the experiences of all participants. The sample is illustrative not definitive” (p. 236).

There were two levels of participants in this study as well as 10 individual cases that were explored both as individual cases and collectively as a multicase (Stake, 2006). The first level of participants consisted of all eighth grade students in the school and was used as a comparison group to answer research sub-question 1: *Does the PARLI framework contribute to reducing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?* Only archival data was collected, analyzed, and interpreted at this level of participation. The second level was the case study level involving 10, eighth grade participants enrolled in a remedial reading course as individual cases, as well as collectively as a multicase. The school district used an established procedure in which the Missouri Assessment Program (MAP), the Gates-MacGinitie Reading Test (GMRT-4), and the Scholastic Reading Inventory (SRI) were used to identify students who were not reading on grade level. To be in this course, students needed to be identified as being between 6 months and 2 years below grade level in reading comprehension performance. The 10 students in this case study were purposefully selected based on their categorization as typical struggling readers as specified in the district protocol. The district’s protocol at the time of the study included evaluating data on all three reading assessment measures as well as teacher
reports to qualify students from among the almost 40% of the student body identified as not meeting grade level expectations on the GMRT-4 reading assessment.

Participation in the study was voluntary, and students (and their parents) had the choice to elect to not participate in this study by not having their data included, without any consequence to the student. Each remedial reading class at this level has a maximum number of 10 students. There are a total of approximately 850 students at the middle school, with the eighth grade constituting approximately 33% of the student body, or 280 students.

As with many suburban schools across the country the middle school featured in the present study underwent a shift in socioeconomic status (SES) and cultural/ethnic diversity over the most recent 5 years preceding the study (Missouri, Department of Elementary and Secondary Education, 2010). Even with these shifts, suburban schools in America are still predominantly white (Alt, Choy, & Hammer, 2000; Aritomi, Coopersmith, & Gruber, 2009; Balfanz, 2009; Baum-Snow & Lutz, 2008; Fry, 2009). Specifically, if free and reduced lunch (FRL) is taken as the measure of low SES, in the 5 years between the 2004-2005 and 2009-2010, this school has seen an increase of 5.9% from 14.6% overall to 20.5% in eligible recipients. Even more telling, the incoming class of sixth grade students in the 2009-2010 school year increased from the previous year’s level of 17.6% to 26.3% eligible for food assistance based on income criteria. According to the Missouri Department of Elementary and Secondary Education (MODESE), during the same time period, this school experienced a 6.6% increase in cultural/ethnic diversity. Mobility rates increased as well. Based on mobility rates for the most recent 2 school years before the study (2008-2009 and 2009-2010), this school building saw an average
increase in mobility to 10.3% (up from 6%) since the 2004-2005 school year (Missouri Department of Elementary and Secondary Education, 2010).

The ethnic makeup of the school was approximately 81% White, 4.1% Asian, 10.4% Black, 4.3% Hispanic, and .3% Native American (Missouri Department of Elementary and Secondary Education, 2010). The ethnic makeup of the participant class reflected a greater proportion of students of color than that of the school overall. Asians and Native Americans were not represented. Specifically, of the 10 students in this class, 7 students (70%) were white, 2 were Hispanic (20%), and 1 was Black (10%). Seven of the 10 students were female (70%), when 47% of the school population overall was female.

Of the group of approximately 230 struggling readers as identified by GMRT-4 and MAP data, the school was able to serve a total of 96 remedial reading students. Specifically, between 30 and 36 students considered to be dramatically below grade level (more than 2 years below) and 60 who were categorized as below grade level (6 months to 2 years below), for a total of 11% of the student population receiving remedial reading services. The study participants were students who were in the below grade level category only. The sample of 10 remedial readers in the case studies represents 17% of those in this below grade level category who were receiving remedial reading intervention services at this school.

Protection of human subjects. This study involved adolescent struggling readers in a remedial reading classroom. Permission to conduct this study was granted at the school and district level. All University of Missouri-Saint Louis Institutional Review Board policies were adhered to in the protection of these students. Both parental consent
and student assent were secured accordingly. Students were participating in regular classroom instruction for remedial readers, as mandated by the participating school district.

**Instruments used in data collection.** As a mixed methods study, both qualitative and quantitative data collection instruments were used. The determination of which instruments was based upon the appropriateness of their use for gathering data to answer a given question. This section shares information about each instrument that allows the reader to evaluate this appropriateness.

**Quantitative Instruments.**

*Gates MacGinitie Reading Test (GMRT-4).* The Gates MacGinitie Reading Test (GMRT-4) is a reading assessment that was originally developed in 1926 and is regularly given in U.S. schools. It is a paper and pencil measure designed for teachers and schools to know the general level of reading achievement of individual students. Students took the Gates MacGinitie Fourth Edition (MacGinitie, et al., 2000). This edition of the norm-referenced reading achievement test features subtests that assess essential literacy skills as highlighted by recent research (National Reading Panel, 2000). It is a multiple-choice test featuring short excerpts administered as a group test. Eighth graders in the study district took form T during the study year. According to the GMRT-4 technical report (MacGinitie, et al., 2002), the reliability estimates indicate strong total test and subtest internal consistency levels with Chronbach’s Alpha coefficient values at or above .90 for the total tests and the subtests at all Levels except AR (Adult Reading): Form S-Vocabulary .88 and Comprehension .89, Form T-Vocabulary .89 and Comprehension .89. Alternate form correlations for the total test scores were at or above .90 for most Levels,
including Grade 8. Alternate form correlations for the subtests ranged from .74 to .92. According to the publisher, stability correlations were calculated for several thousand students who participated in testing with Form S for both fall and spring standardization administrations. The total test coefficient values were at or above .88 for most levels, again including Grade 8.

Content validity was documented through the use of a thorough process of test development to identify the scope of the subtests and identify effective items within subtests (using conventional and item response methods). Item-bias studies were used to eliminate problematic items. Construct validity is supported by the strong intercorrelations between the subtests and their respective total test scores; however, no specific discussion of construct validity is included in the technical report. Similarly, no specific concurrent validity data were presented. (Johnson, 2004, para.10)

The GMRT-4 is administered by the district (of the study school) each March; the 2010 assessment served as a pre-intervention measure. The GMRT-4 was administered to study participants again in December of 2010 as an immediate post-intervention evaluation. The results from the annual administration in March of 2011 were also used as a point of comparison to the previous year and between groups of students (cases, non-struggling peers, and peers in reading intervention in other buildings). In addition, archival data of the GMRT-4 from the sixth grade was also collected.

*Scholastic Reading Inventory (SRI)*. The Scholastic Reading Inventory (SRI) Interactive (Scholastic, 2000) is a computer-adaptive assessment multiple-choice test featuring short excerpts. It is designed to measure how well readers read literature and expository texts of varying difficulties. Scholastic’s Technical Guide (Scholastic, 2001) outlines the measures’ validity and reliability. Its construct validity with the reading comprehension construct is evaluated relative to other reading assessments. The SRI correlates with the print reading assessments it was compared to at an overall level of .83,
and .62 for seventh grade (the highest grade level reported). According to Scholastic (2001), when it comes to reliability, the SRI is strongest when the test is well-targeted (grade level and prior reading level of the student are known). Scholastic reports Standard Error of Measurement (SEM) based on the number of questions the student responds to in this dynamic, interactive assessment. SEMs range from 84 to 104 Lexile points when only the student’s grade level is known, and a much smaller range of 54 to 58 when both the grade and reading level are known (at eighth grade, approximately each 70 Lexile points distinguish the bottom of this grade level with the bottom of the seventh grade level). As a point of reference, for an eighth grade student, a student reading at a measured Lexile level of 600 to 900 is considered Basic, from 900 to 1150 as Proficient, and above 1150 as Advanced Proficient.

Overall, levels of validity and reliability for the SRI are within acceptable ranges (Scholastic, 2001, 1999). The SRI was administered three times during the study, as a pre-intervention measure in August, as a post-intervention measure in December, and finally, as an end-of year assessment in March.

*Metacognitive Awareness of Reading Strategies Inventory (Marsi).* The Marsi (Mokhtari & Reichard, 2002) is a structured questionnaire “designed to assess sixth through 12th grade students’ awareness and perceived use of reading strategies while reading academic or school-related materials” (p. 251). The instrument underwent rigorous development and the reliability and factorial validity of the scale were demonstrated. Specifically, analysis of the Marsi yielded three subscales. The correlations between factors and Chronbach’s alpha reliabilities for each subscale were .89 for the total sample, and .86 for the eighth graders in particular. Relationships
between self-reported reading ability and strategy usage were also investigated, providing evidence of construct validity by conducting post hoc comparisons using the Ryan-Einot-Gabriel-Welch multiple-range test with $\alpha = .05$ for both Global Reading Strategies and Problem-Solving Strategies. “Overall, the psychometric data demonstrate that the instrument is a reliable and valid measure for assessing students’ Metacognitive awareness and perceived use of reading strategies while reading for academic purposes” (Mokhtari & Reichard, 2002, p. 254). The MARSI (Appendix B) is designed for students and teachers to be able to collaboratively complete the scoring to develop a profile of the reader placed on a continuum of high, medium, or low regarding each of the three subscales: Global Reading Strategies, Problem-Solving Strategies, and Support Reading Strategies. It was administered as a pre-post intervention assessment measure.

*Middle School Assessment of Academic Self-Concept and Motivation (MAASCM).*

The MAASCM (Appendix C) represents a minor adaptation of the Assessment of Academic Self-Concept and Motivation Scale ([AASCM]; Gordon Rouse & Cashin, 2000), a structured questionnaire based on motivational systems theory ([MST]; Ford, 1992, 1995). The AASCM focuses specifically on school experiences and consists of four academic domains as well as four aspects of self-concept. The aspects of self-concept measured are ability beliefs, environmental support beliefs, control beliefs, and value/importance beliefs. The academic domains are cognitive, social, personal, and extracurricular. Specifically: “Cognitive refers to learning and doing homework. Social refers to making friends and the quality of friendships. Personal refers to obtaining relevant information and support. Extracurricular refers to participation in extracurricular
activities” (Gordon Rouse & Austin, 2002, p. 300). The AASCM has 80 items and has internal consistency reliabilities on the subscales ranging from .87 to .90.

The AASCM was modified slightly to more specifically target the goals and contexts of specific interest and to simplify the scales to improve its suitability with middle school students. The extra-curricular goals were removed, since the current study middle school has very limited offerings, and was replaced with reading-specific cognitive goals, making it more context-specific. Some of the scale labels from the Assessment of Personal Goals (Ford, 2002) were incorporated to create simpler, more uniform labels. These modifications were based upon guidance from the authors and met with their approval (M. E. Ford, personal communication, March 2, 2010; K. Gordon Rouse-Biddle, personal communication, February 27, 2010). In addition, the original 7 point scale was adapted to remove the midpoint and replace it with “Don’t Know” to address the tendency of the middle school population to select the midpoint when uncertain. Finally, the layout and the categories were slightly simplified to improve suitability with the test population. It was administered at the start of the study as a pre and post intervention measure.

Pragmatic Analytical Reading Level Assessment (PARLA). The PARLA assessment, developed by the study’s author (Bradarich, 2008), is an essay response measure with four subtests. The four subtests are: (a) PARLA-NARR (narrative fiction), (b) PARLA-NARR NF (narrative nonfiction/ Communication Arts); (c) and PARLA-EXPO (expository/ Social Studies). Each assessment features a short story/feature length article complete with photos, illustrations, graphs, and charts appropriate to each genre.
The nonfiction pieces were written by the study’s author. The fiction pieces, written by others, were used with permission and are not reproduced here.

The PARLA measures are paper and pencil essay response measures featuring a complete short story (Appendix A) and were developed in keeping with the guidelines described in Hillocks (1980), and revised with input from the author (G. H. Hillocks, Jr., personal communication, February 16, 2009). Hillocks developed a taxonomy of skills in the interpretation of fiction that is comprised of seven different levels. His work examining the characteristics of several question sets featured Rasch Rating Scale model analysis to confirm experimentally the hierarchical and taxonomic nature of the item types he proposed (Hillocks & Ludlow, 1984). In accordance with the seven levels of the hierarchy, seven item types, “which can be discriminated from each other and organized taxonomically through logical analysis” (Hillocks & Ludlow, 1984, p. 7) are used in the PARLA assessments. The questions in the assessments start at the most basic level and move through the progression of the hierarchy. The levels can be categorized into two overall categories: those requiring a literal level of comprehension and those requiring an inferential level of comprehension. Specifically, the seven question levels are:

- **Literal Level of Comprehension**

- **Level 1: Basic Stated Information.** These questions are directed at the literal information that is prominent and repeated in the text.

- **Level 2: Key Details.** Questions here refer to information that is important in defining key plot twists and turns.
• **Level 3: Stated Relationships.** Questioning here refers to a relationship that is explicitly stated to exist between at least two pieces of information in the text.

• Inferential Level of Comprehension

• **Level 4: Simple Implied Relationship.** Questions at this level mark the move out of the concrete realm. This is similar to Level Three, but the reader must make a single inference to understand the relationship between two pieces of information in the text.

• **Level 5: Complex Implied Relationships.** To assess the reader at this depth of understanding, the questions must ask the reader to make inferences based on many pieces of information throughout the text. The reader must discern a pattern among a variety of inference and draw reasonable conclusions.

• **Level 6: Author’s Generalization.** Questioning at this level represents a significant challenge wherein the reader must abstract generalizations about the nature of the human condition. It differs primarily from Level Five in that it deals with ideas implied in the world beyond the text.

• **Level 7: Structural Generalization.** The final questions in the hierarchy consider how parts of the literary work operate together to achieve the effects desired by the author (Hillocks & Ludlow, 1984).

The PARLI framework begins with fiction and progresses into nonfiction genres. Accordingly, a PARLA-NARR was administered at the start of the narrative nonfiction unit of study, as a formative assessment (Appendix A). At the conclusion of that unit of
study, an additional PARLA-NARR subtest was administered to provide a summative
evaluation of student progress (Appendix A). This pattern was repeated for each of the
remaining two subtest disciplines (see Table 4 for a detailed timeline). The PARLA
nonfiction stimulus pieces were constructed to reflect content-area classroom reading
tasks (Alvermann, 2002; Beers, 2003; Daniels & Steineke, 2004; Harvey & Daniels,
2009; Harvey & Goudvis, 2007; Keene & Zimmerman, 1997; Robb, 2008; Schoenbach,
et al., Vacca & Vacca, 2005) and were Lexiled at the eighth grade level (MetaMetrics,
1984), to provide a pragmatic measure of student comprehension in each respective
content area.

These hierarchical assessments were based on Hillocks’ (1980) hierarchical
levels, featuring constructed-response and essay questions at progressively more
challenging levels of comprehension. This essay format closely parallels that of the
Hillocks’ inventories (1980, Hillocks & Ludlow, 1984) and follows the guidelines
presented by Hillocks and Ludlow, making it reasonable to anticipate that these
assessments share the reliability and validity that those measures were found to possess.

The author conducted a pilot of versions of these assessments among 281 eighth
grade students (Bradarich, 2008) that featured scaled multiple-choice questions. Because
the proficiency level of each student was determined by his or her cumulative
performance on the questions at that level, individual response items were combined to
create a level score. Spearman’s correlation coefficients (R) for nonparametric data were
calculated between students’ scores on each parallel level across subtests, with the result
that the three subtests were found to have statistically significant correlations at the literal
levels of understanding, with Spearman’s correlation coefficients ranging from .182 to
.344 ($p<.01$, $N=281$). As predicted, at the higher levels of comprehension the PARLA subtests did not consistently correlate with each other, with one notable exception. There was only one correlation from among Levels Four through Six with a Spearman’s correlation coefficient of .128 ($p<.05$, $N=281$), supporting the theoretical stance that reading comprehension varies for individual students across disciplines.

The exception occurred at Level Seven, where the responses did correlate for two of the three subtests (.172 and .186, $p<.01$, $N=281$). It was postulated that this may have been a result of the scaled multiple choice format and its unsuitability as a measure of the complex level of understanding reflected by Structural Generalization in the evaluated version of the assessment (G.H. Hillocks, Jr., personal communication, February 16, 2009). It should be noted that the assessments used in this study return to the format of the work completed by Hillocks (1980) with the original enumeration and validation of the hierarchy, by requiring students to provide exclusively written responses that offer the opportunity for complex responses appropriate to the complex questions asked at this level.

Qualitative data collection.

Classroom observations and field notes. Teddlie and Stringfield (1993) explain the value of classroom observation in the study of school effects and indicate that variations in teacher behavior have dramatic influence over results. Teddlie and Stringfield indicate that “these data are necessary, however, if the field is to move beyond collection of archived and survey data into the actual study of class and school processes” (p. 227). They also made a distinction between observation for research purposes and that used for personnel evaluation. Observations for personnel evaluation follow strict
guidelines, as agreed upon by employment contracts. On the other hand, observation for research purposes has as its aim developing an understanding of the complex processes of teachers and students engaged in teaching and learning to allow for the development and implementation of the most effective practices.

While this research was not ethnography, it borrowed from this methodology in the observational collection of data. The specific (ethnographic) focus of these participant-observations was similar to strategic ethnography because this evaluation is in the service of the human need (Spradley, 1980) and to improve reading comprehension instruction among struggling adolescent readers. Spradley’s (1980) twelve steps, and how they were, or were not followed in this research:

(a) Locating a social situation: The social situation was the eighth grade reading class.
(b) Doing participant observation: These observations were conducted twice per month.
(c) Making an ethnographic record: “An ethnographic record consists of fieldnotes, tape recordings, pictures, artifacts, and anything else that documents the social situation under study” (Spradley, 1980, p. 63). This study included field notes, recordings, and artifacts, but not for the purpose of ethnography.
(d) Making descriptive observations: These were done daily in the form of field notes, as well as by analyzing the video recordings of the observations and using Spradley’s Descriptive Question Matrix (Spradley, 1980, pp. 82-83).
(e) Making a domain analysis: This involves a search for patterns in the social setting and was done for understanding the case, not as ethnography.
(f) Making focused observations: Ethnographic focus refers to a single cultural domain. Because this study was not a study of culture, this step did not strictly apply; however, this study’s observational focus was reading comprehension and the attendant behaviors and interactions with which comprehension was connected.

(g) Making a taxonomic analysis: Through the process of axial coding, the researcher conducted a taxonomic analysis, but not for the purpose of understanding the structure of a culture; the purpose here was to understand the case.

(h) Making selected observations: Formal observations took place every ten days.

(i) Making a componential analysis: Selective coding was this study’s correlate to Spradley’s componential analysis that is a “systematic search for the attributes (components of meaning) associated with cultural categories” (Spradley, 1980, p.131).

(j) Discovering cultural themes, (k) taking a cultural inventory, and (l) and writing an ethnography: were not done.

As part of observing in the field, many ethnographers frequently jotted down notes of their observations. In addition to the recorded classroom observations that were conducted approximately two times each month throughout the course of the study from August through December, the researcher took regular, daily field notes. As Emerson, Fretz, and Shaw (1995) share, “jottings translate to-be-remembered observations into writing on paper as quickly rendered scribbles about actions and dialogue” (p. 20). These jottings are meant to refresh the researcher’s memory later in the day and to “enable the
fieldworker to catch significant actions and construct evocative descriptions of the scene” (p. 20). Emerson et al. (1995) caution researchers about open jotting and its potential disruption to the events being observed. In the present study, the teacher/researcher making notes during class was a usual occurrence and did not provide any disruption; however, the majority of field notes were written at the conclusion of each day. Emerson et al. (1995) “strongly encourage researchers to sit down and write full field notes as soon as possible after day’s (or night’s) research is done” (p. 40), an assertion echoed by Merriam (2009). This was how the field notes were collected in this study.

Merriam (2009) describes field notes as a “written account of the observation” (p. 128) and shares that in many instances the participant observer “will jot down notes during an observation and wait until afterward to record in detail what has been observed” (p. 128). As advocated by Merriam, field notes in this study were written in a format that allowed the researcher to find information easily. Time, place, and purpose of the observations, along with participants and diagrams, where warranted, were consistent elements of field notes. They were written daily to maximize accuracy and, as Merriam (2009) directs, be highly descriptive. Merriam defines highly descriptive as “enough detail should be given that readers feel as if they are there, seeing what the observer sees” (p. 130). On days that observations were video recorded as well, field notes were compared to the recordings as a check for both accuracy and researcher bias.

*Think-aloud protocols.* Think-aloud (T-A) protocols (Pressley & Afflerbach, 1995) are a one-to-one data collection method that involves the reader speaking aloud about the moves and strategies he or she is using while reading a piece of text. Students participated in a T-A twice for each of three genres, for a total of six times during the
course of the study, with two instances each being from narrative text, nonfiction narrative, and the final one from expository text.

For each instance, students read from two texts, one that was about more familiar content, and one that was less familiar. This was to allow for comparison between processing text that is less likely to cause struggle with one that will be more revealing of “the problems and processes involved in children’s construction of a coherent representation” (Coté & Goldman, 2004, p. 662).

Furthermore, none of these texts was unreasonably difficult; rather, they were chosen for their similarity to texts the students were likely to encounter in their content area classes. The direction given was open-ended and asked the reader to speak about his or her reading and thinking processes while reading. A detailed transcript of each protocol was generated. “‘Moves’ are responses reflecting what the reader is doing at a particular point in time to understand what he is reading, and ‘strategies’ are the patterns of moves utilized to solve a particular problem with comprehension” (Meyers, Gelzheiser, & Pruzek, 1989, p. 4). The think-aloud protocol (T-A) is qualitative in nature, in that it results in a narrative transcript to be evaluated. Think-aloud protocols have been investigated and analyzed to allow the researcher to build on the work of others using deductive analysis (Coté & Goldman, 2004; Ericsson & Simon, 1984; Meyers et al., 1989; Pressley & Afflerbach, 1995).

In a meta-analysis of think-aloud studies of reading, Pressley and Afflerbach (1995) found three overarching types of activities that readers engage in: (a) constructing meaning, (b) monitoring, and (c) evaluating (Pressley & Afflerbach, 1995). The author’s analysis of these studies yielded a detailed classification scheme with an extensive list of
categories of specific moves that readers make in each of these types. Table 4 lists the
categories of focus for this study. While the list created by Pressley and Afflerbach is
rather exhaustive, the authors do not make a claim to saturation: “The fact that relatively
few adjustments were made for the last few studies integrated into the analysis permits
confidence in the classification structure” (p. 82). Overall, this classification scheme
highlights the valuable contribution of think-aloud protocols (T-A) to shed light on the
processing in which readers engage when trying to understand text. The classification
scheme allows for a systematic quantitizing of the data generated from think-aloud
protocols.

Pressley and Afflerbach (1995) recommend that think-aloud protocols are
strengthened by including the following seven components: (a) detailed characteristics of
the subjects; (b) specific characteristics of the texts read; (c) clear and specific directions
given to subjects with verbatim presentations, modeling, and reminders given to subjects
during the T-A; (d) directions given and the nature of the practice, including feedback
and coaching, provided for the T-A before and during data collection; (e) details of the
formal check of reader understanding of directions for the T-A that includes details of
any intervention implemented if directions were not understood; (f) a complete report of
analysis of the transcripts; and (g) a detailed account of how the example transcripts are
chosen for inclusion in the report that includes their representativeness of the total
transcript base.

A panel of literacy experts consisting of three secondary reading specialists, two
employed by the study district, and one from a neighboring district, was incorporated at
the outset of the gathering of this data to offer control for potential researcher bias and to
inform the remainder of the data collection and analysis. This panel was likewise included at the end of the study as a means for controlling potential bias.
Table 3.

*Think-aloud Protocol Categories*

<table>
<thead>
<tr>
<th>Before Reading</th>
<th>Processes Specific to Monitoring Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal setting</td>
<td><strong>During Reading Monitoring</strong> Text characteristics like difficulty, author style</td>
</tr>
<tr>
<td>Overviewing/skimming</td>
<td><strong>Meaningful processing of text</strong> like purpose for reading, own behaviors, awareness of success or lack thereof</td>
</tr>
<tr>
<td>Activating prior knowledge</td>
<td><strong>Recognizing problems</strong> with self or Text</td>
</tr>
<tr>
<td>Predicting</td>
<td><strong>Monitoring stimulation of cognitive processing</strong> and making decisions about how to proceed</td>
</tr>
<tr>
<td></td>
<td><strong>Activating processing</strong> due to awareness of difficulties at word or phrase level</td>
</tr>
<tr>
<td></td>
<td><strong>Activating processing</strong> due to difficulties beyond word or phrase level</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>During Reading Specific Actions</th>
<th>After Reading Specific Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading front to back</td>
<td><strong>After Reading Monitoring</strong> Monitoring decisions to process additionally</td>
</tr>
<tr>
<td>Skimming</td>
<td>Recitation</td>
</tr>
<tr>
<td>Repeating/restating</td>
<td>Summarizing</td>
</tr>
<tr>
<td>Making notes</td>
<td>Questioning</td>
</tr>
<tr>
<td>Pausing &amp; reflecting</td>
<td>Reflecting</td>
</tr>
<tr>
<td>Paraphrasing &amp;/or interpreting</td>
<td>Integrating different parts of text</td>
</tr>
<tr>
<td>Predicting</td>
<td></td>
</tr>
<tr>
<td>Identifying important information</td>
<td></td>
</tr>
<tr>
<td>Conscious inference making</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Think-aloud protocol categories adapted from Pressley & Aflerbach, 1995. These were the categories that initial coding of the protocols was based upon.*
Learning reflection logs and artifacts of reading comprehension. The learning log is a student’s commentary upon a particular course of study, in this instance, with the PARLI framework. In using a learning log, the PARLI framework was tapping into the power of this thinking and learning tool for students, while providing an explicit window into the students’ processes for the teacher/researcher. Among the research studies that support writing to learn, learning logs have been shown to help students process information they are reading, particularly in science (Santa & Havens, 1991; Shepardson & Britsch, 1997).

Members of the quintain were encouraged to document the details of what they did, jot down why they did it, their initial reactions, questions that occurred to them at the time (so they could be followed up later), and tentative conclusions they reached throughout their course of study. Entries were made frequently, and were dated. These entries resulted from two processes: The first was each student keeping a log of what he or she did when using the PARLI framework while reading. The second was a periodic exercise of reflecting on one’s log either weekly, or bi-weekly. To maintain their intended roles as a flexible and responsive learning tool, learning logs were kept in three-ring binders that allowed students to add and subtract documents and forms as needed. As part of these learning logs, the artifacts of reading comprehension that were addressed in the framework and reflected upon in the logs were also collected and analyzed.

Procedures

Approval process. The approval of the research site was obtained from the superintendent of the school district. An application for approval was first submitted to the Institutional Review Board of the University of Missouri, Saint Louis Institutional
Review Board. Following approval, consent from the parents of the participants, and assent from the participants was secured. The forms for the approval process are in Appendix D. Once all the approvals were obtained, the study proceeded according to the study timeline in Table 4. In addition, this study employed several tools that are the intellectual property of other researchers. Accordingly, explicit permission to use them was secured (Appendix D). Over one semester, this was the intervention strategy:

Process of evaluative/assessment events.

- The Gates MacGinitie Reading Test (GMRT-4) and the Missouri Assessment Program (MAP) were administered by the school district. The spring 2010 GMRT-4 was the pre-intervention score for this measure.
- Students took the post-intervention GMRT-4 in December of 2010.
- The pre-intervention score for the Scholastic Reading Inventory (SRI) came from the first administration in August of 2010, and the post-intervention score from the December 2010 administration of this instrument.
- Initial measures for pre-post comparisons began with the Metacognitive Awareness of Reading Strategies Inventory (MARSI) and the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) pre-intervention measures. Classroom observations began at the outset of the study. Think-aloud protocols (T-A) were conducted after the initial measures of the pre-post survey tools. Observations and think-aloud protocols continued from August through December.
• Literature circle discussions were video recorded three times during the Narrative Nonfiction unit as this was the unit that featured a full-length piece read by the entire quintain.

• Both the MARSI and the MAASCM were given again in December of 2010 as post-intervention measures.

• The process for the PARLI assessments started with pre-tests, followed by in-depth differentiated instruction in the framework, and concluding with the post-test. The order of implementation was fiction (assessed with the PARLA-NARR), narrative nonfiction (assessed with the PARLA-NARR NF), and expository (assessed with the PARLA-EXPO).
Table 4

Study timeline from implementation to analysis.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GMRT-4**</td>
<td>Committee and IRB Approval process</td>
<td>Research begins-</td>
<td>9/13-9/14 PARLA NARR-Post</td>
<td>Fall Break 10/4-10/8</td>
<td>11/2-11/3 T-A Protocol NARR NF</td>
<td>12/6-12/7 T-A Protocol EXPO</td>
<td>Analysis of Data</td>
</tr>
<tr>
<td>MAP test**</td>
<td>8/12-school year and PARLI instruction begin</td>
<td>9/20 NARR NF-Pre</td>
<td>10/11 SRI 2</td>
<td>11/8-11/9 PARLA NARR NF -Post</td>
<td>12/ 8 SRI 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8/16 SRI 1</td>
<td></td>
<td></td>
<td></td>
<td>12/16-12/17 GMRT Post</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8/16 NARR-Pre</td>
<td></td>
<td></td>
<td></td>
<td>12/15 Marsi &amp; MAASCN Post</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8/20-8/23</td>
<td>T-A Protocol NARR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data transcription and management. The audio recorded think-aloud protocols were transcribed. During transcription, the researcher employed the common rules of transcription (Corbin & Strauss, 1990) to ensure consistency of data recordings. The lines of each transcript were numbered for easy access and data reference, which was necessary when developing a code book. Atlas.ti software was used to facilitate and manage the consistent system of coding, categorizing, and memos that is essential for rigorous analysis and reliable record keeping. In addition, Atlas.ti was used to both maintain a chain of evidence and create a case study database of the various data types collected for analysis, as recommended by Yin (2003).

Reliability and validity. This case study addressed two steps pertaining to construct validity. The first step was specifically selecting changes in reading comprehension processes and outcomes to study. This study looked at metacognitive processes and outcomes of improved agency, motivation, and greater reading comprehension. The second step regarding construct validity was demonstrating that the selected measures, for example those of reading comprehension (GMRT-4, SRI, PARLA), actually reflected what was selected for study. The validity of each measure was addressed within the description of the measure. Each measure utilized herein has demonstrated validity independently prior to its selection for inclusion. When focusing on internal validity, the concern was for those aspects of the case study that were looking for explanatory relationships among the different types of evidence gathered. External validity was about whether or not the findings of this case study are generalizable beyond this particular community of struggling eighth grade readers. Merriam (2009) suggests several strategies to address external validity concerns that include thick description,
multisite designs, modal comparison, and random sampling. Merriam (2009) shares that the researcher has “an obligation to provide enough detailed description of the study’s context to enable readers to compare the ‘fit’ with their situations” (p. 226). Ultimately, it is the reader of the research who determines whether or not, and to what degree, study results are generalizable or applicable to a particular situation.

The researcher used thick description as a key means of establishing a level of detail that allowed for this generalizability. The study also incorporated modal comparison in describing how typical the selected sample was as described earlier in this chapter. As stated previously, the appropriate level of generalization for case study is at the theoretical level. As such, the potential for analytical generalization to the broader theoretical frameworks on which the PARLI framework was built was the measure of the external validity of this case study. However, even if this case study does provide analytical generalization, further work to replicate these findings through additional case studies would be needed to provide strong support for this broader theoretical framework (Yin, 2003). The rigor and transparency of the mixed methods analysis employed, including a detailed audit trail, helped to secure the validity of this case study as described.

Finally, reliability entails ensuring that if a later investigator followed the same procedures, and conducted the same case study again, similar findings and conclusions would result. As with validity, the measures selected for inclusion in this study were all previously demonstrated to be reliable, as noted in each individual descriptive section. By making as many steps as operational as possible and conducting the “research as if someone were always looking over your shoulder” (Yin, 2003, p. 38), and the strict
maintenance of an audit trail, the present research hopes to withstand the test of reliability.

**Data analysis.** Analysis for this study included both inductive and deductive analytical processes. The quantitative analysis in this study was deductive. The qualitative analysis incorporated both inductive and deductive analysis.

**Quantitative analysis procedures.** Because the GMRT-4 was the primary quantitative measure in this study and was used to compare growth in reading comprehension across three groups, it would be customary to use one-way between-groups ANOVA (Tabachnick & Fidell, 2007) to explore the relationships between them. However, according to Gall, Gall, and Borg (2007), because these data are nonparametric, appropriate nonparametric alternatives must be used. The Kruskal-Wallis Test is the non-parametric alternative to a one-way analysis of variance between groups. It allows for the comparison of scores on a continuous variable for three or more groups. Scores are converted to ranks and the mean rank for each group is compared.

There were common components to the statistical design of the analysis of all of the quantitative measures. First, the small sample size of this case study and the sampling methods employed limited the power of statistical measures; since the assumption of a large, probabilistic normally distributed sample is not possible, only nonparametric statistics may be used (Gall et al., 2007). All of these measures share the procedures of group comparisons (between the case study students and their non-struggling peers) and individual pre and post comparisons exclusive to the case study participants. The products of these analyses were both measures of growth and measures of the gap between struggling readers and their non-struggling counterparts. In this study, one of the
key questions to be answered was to what degree the students in the case study close the
gap between themselves and their non-struggling peers in standardized reading scores.
Usually, when there are two samples that the researcher wants to compare concerning
their mean value for some variable of interest, the $t$-test for independent samples is used
(Mendenhall & Sincich, 2003). In this case, a nonparametric alternative was required.

For this study, to answer the question: Does the PARLI framework contribute to
reducing both the fiction and nonfiction reading comprehension gap between struggling
readers and non-struggling, grade level readers? the Mann-Whitney U test (Hill &
Lewicki, 2010) was used. This test was conducted for the GMRT-4, SRI, and appropriate
common quantitative assessments. To answer the research questions that involve
comparing pre and post scores for each case, median was the measure of central tendency
that was used because the median is less affected by the tails of the distribution (Hays,
1994; Hill & Lewicki, 2010) and thus, one is not restricted by assumption of normality.
The small sample size of the study did not conform to statistical normality assumptions.

*Pragmatic Analytical Reading Level Assessments (PARLA).* Rubrics were created
to score the results for each measure and distinguished between “right” answers, “partly
right” answers, and “wrong” answers. For literal level questions, Level One through
Level Three, the right answers had the key information sought and allowed for a small
amount of variability. Partial credit was given to answers that were correct, but
incomplete responses. Given that higher level questions, by definition, cannot be drawn
directly from the text, more variability in responses was anticipated. To receive full
credit, higher level responses had to include evidence from the text; correct
generalizations without evidence received partial credit.
In addition to the scoring based on the rubrics developed, the responses were evaluated using content analysis and compared with expected levels of abstract thinking based on Dynamic Skill Theory (Fischer, 1980), to generate a profile of the reader’s current inference skill level when reading each form of text. Finally, the assessments asked students to rate their confidence in their responses on a 4 point scale. These responses were analyzed to add depth to the assessment of each student’s current independent level of reading comprehension and evaluation of their Personal Agency Beliefs.

**Qualitative analysis procedures.** As with the quantitative measures in this mixed methods research, it was anticipated that there would be significant overlap in the analysis process across all qualitative measures. For this reason, the aspects of analysis that were consistent are explained here. Both inductive and deductive coding was employed to achieve reduction of data into categories that provided explanation of findings to answer the research questions. As such, both began with the research question. Mayring (2000) offers helpful step models of both inductive and deductive analysis that help clarify these processes and were used to guide analysis for this research in conjunction with a Grounded Theory approach.

Inductive processes have, as their goal, finding patterns in the data. According to Mayring (2000), inductive analysis begins with determination of category definition and levels of abstraction for inductive categories. Analysis then proceeds with step-by-step creation of categories out of the data in an iterative fashion that subsumes old categories and creates new ones; the specific process for this study incorporated Mayring’s (2000) procedural suggestions in following a Grounded Theory approach. When 10% to 50% of
the material has been analyzed, Mayring (2000) recommends a formative check of reliability that takes the researcher iteratively back to the research question and the steps of category formulation. Ultimately, the final working through of the data, a summative evaluation of reliability and interpretation of results may include quantitative analysis aspects such as frequencies of coded categories.

On the other hand, deductive processes have as their goal confirmation of hypothesized solutions to a research question. In Mayring’s (2000) step model of deductive category application the researcher also begins with the research question. However, in this circumstance, analysis starts from prior formulated, theoretically based, main categories and subcategories. Definitions, examples, and coding rules for each deductive category are theoretically based and determine the precise conditions under which data is coded with a given category, using a coding agenda. The process is iterative, including revision of categories and the coding agenda and formative checks of reliability by revisiting the research question and the earlier steps in the analysis process until the final working through of the data with a summative check of the reliability.

“Category definitions, prototypical text passages, and rules for distinguishing different categories were formulated in respect to theory and material, are completed step by step, and are revised with the process of analysis” (Mayring, 2000, para. 17). As it does with inductive analysis, the process concludes with interpretation of results that may include quantitative analysis, such as frequencies of coded categories.

The qualitative data collected in this study were either textual information gathered through observations and think-aloud protocols, or artifacts generated by students as a work product. The researcher used the analysis system of Corbin and
Strauss’s (1998) Grounded Theory, which involves several progressive, and usually overlapping, coding steps including open coding, axial coding, and selective coding. Gibbs’ (2007) system of developing coding hierarchies in data analysis, as part of the move from open to axial coding were also incorporated.

The researcher maintained a complete record of her thought processes and decisions and used a combination of electronic and hand-written notes. Specifically, a journal notebook was maintained for general field notes and observations throughout the course of the study. In addition, Microsoft OneNote® was used to gather and store progress and anecdotal information on case study participants to facilitate sharing relevant information with students and parents throughout the course of the school semester. Analysis was facilitated by the use of Atlas.ti, a qualitative data analysis (QDA) software package. An extensive audit trail comprised of the various notes or memos became an integral part of the analysis process. While this study was not Grounded Theory, it borrowed heavily from its methodology in the coding process.

Open coding. To be more specific, qualitative coding was an on-going process from the start of the research project. Coding began with open coding of each transcript line-by-line, and in some places word-by-word, conducting a thorough microanalysis (Strauss & Corbin, 1998) and engaging in initial memo writing in long hand. At this point, not surprisingly, the number of codes generated was totally unwieldy, at 208 codes in seven code families (Appendix E). Subsequently, the second stage in open coding to place, as explicated by Strauss and Corbin (1998), wherein a constant comparative analytic process was used to break down the data into incidents, ideas, and events. Next, coding was revisited recursively, using Atlas.ti to revisit the coding of all transcripts and
documents and engaged in memo writing, particularly where I was trying to look at similarities, differences, and specificity of the initial codes.

Microsoft’s OneNote® was used in conjunction with Atlas.ti to create memos and comments with codes and maintain a detailed audit trail, as well as consistency across all 10 cases. This constant comparison process of axial coding, with a focus on further reflection on the concepts previously coded and how they might come together to define categories generated a more precise list of 41 codes in four code families representing concepts. Strauss and Corbin (1998) define concepts as “an abstract representation of an event, object, or action/interaction that a researcher identifies as being significant in the data” (p. 103).

**Axial coding.** Axial coding has a focus of further reflection on the concepts already coded and how they might come together to define categories. Strauss and Corbin (1998) define concepts as “an abstract representation of an event, object, or action/interaction that a researcher identifies as being significant in the data” (p.103). Axial coding is hierarchical coding and, for this study, building webs to graphically represent the data and the relationships between categories, properties, and dimensions was done using *Inspiration*. Gibbs (2007) lists the benefits of organizing codes in a hierarchy: 1) it keeps data neat, 2) it can be a form of data analysis itself, 3) it helps eliminate duplication of codes, 4) it helps the researcher see dimensions, and 5) it helps the researcher ask some analytic questions that can lead to understanding of patterns both within and between participants. Gibbs does caution that the building of a hierarchy requires going back into the data in detail; I did this and found it an effective choice. This
hierarchy building flowed naturally with the creation of a code book facilitated by Atlas.ti (Appendix F).

Axial coding involves reassembling data in new ways by making new connections between categories and subcategories (Strauss and Corbin, 1998). As analysis proceeds, each subsequent datum will be identified and analyzed for its fit with the category under construction. If it is conceptually similar to this one, it will be grouped in this category, if it is substantially different, into new categories.

Axial coding does not take place in a single phase, and the next phase of axial coding began with utilization of a strategies articulated by Saldaña (2009). Saldaña explains that the axis from which axial coding gets its name is a category that was derived from the data in open coding and can be understood through the metaphor of a wooden wheel with extended spokes. Saldaña continues to explain, citing Charmaz: “This method ‘relates categories to subcategories [and] specifies the properties and dimensions of a category’ (as cited in Saldaña, 2009, p.159). Properties (i.e., characteristics or attributes) and dimensions (the location of a property along a continuum or range) of a category refer to such components as the conditions, causes, and consequences of a process--actions that let the researchers know if, when, how, and why something happens (Gibbs, 2007). The researcher created digital webs (using Inspiration) to facilitate this further axial coding, using a wheel format. Categories were linked to their sub-categories and associated concepts to allow the researcher to begin to develop an explanation of the development of metacognition, personal agency beliefs, motivation, and ultimately reading comprehension for each case and the multicase.
Through this process the researcher strived to discover if any categories are congruent categories, defined by Merriam (2009) as categories with the same level of abstraction. At this point, the researcher began to consider the possibility of a central phenomenon, its contexts, and the causal and intervening conditions which seem to be underlying the processes and outcomes that are the focus of this study. According to Strauss and Corbin (1998), selective coding is “the process of integrating and refining the theory,” (p. 143). They share that the first step of selective coding is deciding on a central category. The authors explain the central category is one with analytic power that, “What gives it that power is its ability to pull the other categories together to form an explanatory whole,” (Strauss & Corbin, 1998, p. 146). If this central category becomes apparent, what the researcher thought might be other, independent categories prior to this moment, will become actually properties, dimensions, and subcategories, resulting in a small number of major categories, and the central category. As these categories are dimensionalized, the researcher checked to be sure this development was in keeping with the data.

Axial coding resulted in four main categories: Context, Product, Personal Agency Beliefs, and Aspect. Explanations of the interrelations of these categories, sub-categories, and numerous associated concepts are described. These four categories help explain students’ complex reading comprehension development across three text forms: Narrative, Narrative Nonfiction, and Expository.

Context was the first category, as it sets the tone for each entire literacy interaction. The properties of assessment (yes or no), social context (individual [solo] and collaborative), output (talk or write), and the actors (quintain) all define the nature of the
*space* in which the reading comprehension work took place. *Context* factored heavily in the evaluation of pattern (sub-question 5).

*Product* was the second largest category with sub-codes of *Levels* and *Quality*. *Levels* was a deductive code category, and has five dimensions representing five of the levels of the hierarchy of reading comprehension that underlies the PARLI framework. Levels reflected the reading comprehension being demonstrated. Quality was a dimensional code that identifies the quality or strength of a given response or thought process along a continuum from *missing* or *wrong* to *strong*. This code identified data related to metacognition.

Quality sub-codes define the dimensions of the quality of *Products*, and provide depth to the understanding of the nature of the *Product* produced within any *Context*. They are also used to describe the depth/nature of the reflections about *Personal Agency Beliefs* and overall performance.

The Quality dimension of *Product* addresses the nature of a particular product along a continuum and further delineates the characteristics of responses at different levels where both dimensions are applicable, and also describes the *Aspects* such as reflections or words per minute. The dimension of “*strong*” was for *Products* or *Aspects* that leave no question about the student’s ability to think at this level, communicate this characteristic clearly, demonstrate support for their thinking from the text, and express confident and positive opinions and beliefs; in a word, mastery. Likewise, the dimension “*good*” shows positive characteristics, but just a slightly lesser level than does “*strong*”; in a word, proficiency. Novice-level, less sophisticated thinking and products are coded with “*ok*” to indicate that evidence of the students’ mastery was not unequivocal, but was
certainly suggested. The last three dimensions, “weak”, “miss”, and “wrong” all describe the lower end of the spectrum of the quality of work produced. “Miss” reflects the absence of an expected Aspect, while “wrong” labels a present, but incorrect response, and “weak” was one that begins to hint at quality, or was of a very basic nature.

Aspect was closely related to Products, but specifically looks at the action/interaction taking place within a Context and producing a particular Product. Aspect has 13 subcategories to delineate specific actions/interactions across literacy events and tasks and identified the thinking exhibited (metacognition) by the case in articulating the student’s understanding of the text.

Aspects coded broadly across Products and Contexts included annotation of text, be it a written or verbal aside or comment (ASP anno), and answer (ASP answer), which was for the action/interaction of the student providing an answer/response that just covers the basic response, providing no explanation or “because” to support it. Explain (ASP explain) identified the action/interaction part of the response in which the student explained his or her thinking about the response and/or the importance of the segment of text to which he or she was responding. It may also have been used to code the entire response if it did not make sense to separate out the answer and the explanation, be it verbal or written.

Less frequently found dimensions of Aspect were ASP ques, for the creation of questions that demonstrate mastery of any particular level of comprehension, ASP quote, for the inclusion of a specific excerpt from the text that supports the thinking of the reader with solid evidence, and ASP title, wherein the reader engages in an action/interaction of creating a title for the response that captures the essence of the response.
Each of these codes categorized thinking that was indicative of metacognition and the reader/thinker needs awareness of his or her own thought process to produce them.

Finally, ASP WPM and ASP Ref were more specialized and applied only to particular Contexts and Products. ASP Ref was for the reflection of the student about his or her performance. They may have reflected about their anticipated performance overall, their performance on a particular question, or their performance after they have looked at their scores. If it was a reflection about a particular response at an identified level, the level of the response was coded also.

All Reflection codes also have PAB codes. ASP WPM was a coding of words per minute that the student spends reading the piece. The Quality categories here related to accepted grade level reading rates: a) QUAL strong = >eighth grade, b) QUAL good = eighth grade, c) QUAL ok = sixth-seventh grade, d) QUAL weak = fourth-fifth grade, and e) QUAL wrong = < fourth grade.

The profound theoretical and practical importance of motivation and engagement to learning in general, and advancement in complex literacy in particular, factor into the last category, Personal Agency Beliefs. The taxonomy used herein represents deductive coding based on Ford’s (1992) MST theory and Taxonomy. Personal Agency Belief, or PAB codes classified the action/interaction of the students’ beliefs about the task at hand and his or her likelihood of success that this task and context reflect, based on the MST Taxonomy (see Table 4).

PAB codes were used for student’s articulated reflections about their performance, either with a specific product or in general, and may be written or spoken. Additionally, students were asked to rate their confidence in their responses on the proximal measures
(PARLAs) and these were correlated with the rating of their response based on the standard rubrics used, to derive a PAB code that represented the intersection between their self-perception of efficacy (self) and the objective evaluation of their performance (context).

**Classroom observations and field notes.** Classroom observations included fieldnotes from the participant-observer perspective and videotaping of book discussions on three occasions. In addition, the researcher analyzed these notes, using elements of Grounded Theory, chiefly the constant comparative method of inductive coding on an ongoing, iterative basis such that codes were developed using earlier observations for incorporation into later observations and their analyses.

Montgomery and Bailey (2007) address the relationship between field notes and theoretical memos in Grounded Theory. While this study was not Grounded Theory, it did borrow heavily from it. In addition to using a constant-comparative method of coding through the three levels of coding in Grounded Theory, the researcher regularly engaged in constructing memos, including theoretical memos, as part of the write up of field notes as well as the analysis of the notes. Montgomery and Bailey assert a natural connection between field notes and memos because as field notes help the researcher make sense of the perspectives and actions of participants, capturing the meaning and ideas related to them is the work of memos. These extensive field notes and the attendant memos represented a key element of an extensive audit trail for this study.

**Think-aloud protocols.** The think-aloud protocols generated transcripts for qualitative content analysis. This analysis was conducted using deductive coding processes, starting with the categories listed in Table 4. A panel of adolescent literacy
experts (three secondary reading specialists) underwent common training in analyzing think-aloud protocols using these categories. Training continued until inter-rater reliability was achieved, and categories were modified, as appropriate, based upon the input of these experts. These categories were used as structural codes for the processes readers use and were subjected to a multistage analysis process of interpretation similar to inductive analysis and the coding as described earlier. This cluster analysis groups clusters of observations in a data set, producing a fairly easy to read, graphical output “based on the relationship between codes as they are applied to raw data and on the frequency with which they co-occur” (Guest & McLellan, 2003, p. 195). This cluster analysis facilitated the integration of analysis of the development of reading comprehension processes across both the quantitative and qualitative measures in this case study evaluation.

The multistage analysis process resulted in the following outputs: (a) code frequencies--number of times a code was applied to a transcript segment for each unique respondent, along with the number of times it was applied within the whole classroom case; (b) code co-occurrences--isolated pairs of codes applied to each transcript or a particular stimulus piece used in the protocol and associated with each case; and (c) saliency--number of times that a code was within a combination of codes associated with either a particular stimulus text or a particular case (Guest & McLellan, 2003).

Learning reflection logs and artifacts of reading comprehension. Analysis of learning reflection logs and reading comprehension artifacts was conducted in much the same manner as the analysis of classroom observations, with these collective artifacts as the stimulus rather than video recording or field notes.
Mixed analysis procedures. Parallel mixed data analysis was conducted by analyzing both the quantitative and qualitative data analyses as described, then linking, combining, and integrating them into meta-inferences (Teddlie & Tashakkori, 2009) through an iterative process. Teddlie and Tashakkori define a meta-inference as “a conclusion generated through an integration of the inferences that have been obtained from the results of the QUAL and QUAN strands of an MM study” (p. 152). In addition, conversion mixed data analysis was explored to determine if this form of analysis improves understanding of the patterns of developing reading comprehension skills within this case study.

Qualitative data was quantitized by assigning numeric codes to category codes, while quantitative data was qualitized by using the categories of the standardized reading measures and subjecting them to qualitative analysis. For example, when the analysis yielded 12 potentially key characteristics of metacognition during reading as revealed by think-aloud protocols, the researcher quantitized the data by assigning binary values to each case for each characteristic. Thus, if Case A set goals each time he or she reads, he or she received a score of 1 for that characteristic. If Case B did not, he or she received a score of 0 for that characteristic. A series of binary (1,0) codes was then assigned to each case. These quantitized data can be analyzed to statistically associate each characteristic with demographic, agency, motivation, and/or reading comprehension variables.

By subjecting one type of data to qualitative and quantitative analysis, meta-inferences were possible using both sets of results simultaneously (Teddlie & Tashakkori, 2009). In addition, quantitative data like socioeconomic data and reading comprehension data, were qualitized to create narrative profiles of each of the cases.
Spearman $R$ correlation coefficients were calculated for data transformed and integrated in this way. For example, data from the T-A protocols was quantitized and correlated with the results from the MARSI using the Spearman $R$ correlation coefficient (for nonparametric data). A matrix of all of the measures in this study illustrates the many possible relationships to be explored and analyzed. Those measures for which relationships were explored through this process are indicated in Table 5.

Table 5.

*Matrix of measures used for mixed methods analysis.*

<table>
<thead>
<tr>
<th></th>
<th>GMRT pre &amp; post</th>
<th>SRI 3 times</th>
<th>MARI pre &amp; post</th>
<th>MAASC pre &amp; post</th>
<th>T-A 6 times</th>
<th>Class Obs</th>
<th>PARLA pre &amp; post</th>
<th>Learning Logs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Logs</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARLA pre &amp; post</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class Obs</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-A 6 times</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAASC pre &amp; post</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MARS pre &amp; post</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRI 3 times</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>GMRT pre &amp; post</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Summary of the Chapter

Chapter 4 has provided an explanation of the PARLI framework and how it was implemented in this evaluation study. The progression through three units of instruction, in a predictable and consistent manner, was flexible based on student need. The progress was designed to begin with the simplest form of text (fiction) and build students’ competence at increasing complexity levels in this form before moving on to nonfiction texts.

Similarly, the first of the nonfiction texts was narrative nonfiction, the least complex of the nonfiction forms. This was done to take advantage of how students learn by moving them hierarchically through ever-increasing text complexity. In this way, each unit, in turn, provided a scaffold for the one to follow, with the end result of strengthening students’ capacities to comprehend complex, nonfiction texts across the content areas.

Chapter 5 will present the results of the research. Overall results as well as case specific detailed results are presented. Chapter 6 will feature a discussion of the results, including cross-case analysis.
Chapter 5: Results

This chapter presents the findings of the study. The study conducted a formal evaluation of the Pragmatic Analytical Reading Level Instruction (PARLI) framework. The chapter is organized in terms of the Hypothesis and research questions posed in Chapter 1, for the quintain, and then by case. The study included quantitative, qualitative, and mixed data for 10 cases. These cases together make up what Stake calls a quintain: “This quintain is the arena or holding company or umbrella for the cases we will study” (Stake, 2006, p. 6). This quintain was categorically bound together as struggling eighth grade readers in a reading intervention class, and has been referred to throughout as the quintain. As delineated by Stake (2006), “Multicase research starts with the quintain” (p. 6) and, while individual cases are studied, including making comparisons of what is similar and different between and among them, it is the goal of understanding the quintain that drives this research.

To provide context, overall quantitative results are briefly discussed for the research question and sub-question 1, followed by individual case results. The results for each case begin with background information for the students from the preceding school years. Both quantitative and qualitative results for each case are presented relative to each of the research questions. Cross-case analysis of both quantitative and qualitative results follows the individual analyses. Consideration of how both the individual case data and the cross-case analysis inform the researcher’s understanding of the quintain is presented in the discussion section. The data collection for the study took place between August 2010 and March 2011.
Overall Results

Research question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

When all data was analyzed, the PARLI framework was effective with nine of the 10 cases of the quintain. On two of the three reading comprehension measures 70% of the students moved into the proficient category (GMRT-4, SRI, and PARLI). On both the GMRT-4 and the SRI 30% of the students moved into the proficient category. Taken in light of the meta-analysis that suggest that 10-15% of remedial readers at this age achieving competence is the best that can be anticipated with this age group (Foorman, et al., 1997), the result is noteworthy.

The formative and summative assessment of agency and motivation for the quintain show mixed results, with slightly more declining than either improving or staying the same. It is useful to look at the data from throughout the study to understand these results, as all but one member of the quintain reflected higher levels of agency and motivation over the course of the study while engaging in regular coursework. Evidence from both the students’ artifacts and through observation and dialogue with them as the study progressed showed positive change for all but Case 2. When comparing pre and post measures of the awareness of metacognitive strategies used while reading (MARSI) not much change is evident. Three cases shared an improvement, while the remainder showed no appreciable difference.
**PARLA.** For the proximal measures of the Pragmatic Analytical Reading Level Assessments (PARLA) in each of three content areas of Narrative, Narrative Nonfiction, and Expository texts, the average growth for the quintain was: 1) 1.5 stages on the Narrative, 2) 3.9 stages on the Narrative Nonfiction, and 3) 2.7 stages on the Expository. Based on earlier research in the development of these assessments (Bradarich, 2008), it is expected that how much average readers will progress is connected to where they start, with less movement up the hierarchical ladder per year once the reader achieves the inferential level (PARLI Level 4). This is supported by Fischer’s Dynamic Skill Theory (Fischer, 2008). Accordingly, students who begin the eighth grade year on grade level are expected to improve by approximately two stages, moving from Stage B in Level 4, to Stage A in Level 5. On average, the students in the quintain exceeded this performance expectation.

**Sub-question.** 1) *Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*

**GMRT-4 and SRI.** In this district in general, students are expected to demonstrate growth of approximately 1 year on standardized reading assessments annually. For the SRI, this anticipated growth translates to approximately 64 Lexile points (Williamson, 2006); and for the GMRT-4, approximately 12 months. While discussing GMRT-4 scores in terms of months or Grade Level Equivalents (GLEs) is common practice, the GMRT-4 scoring manual recommends the use of Normal Curve Equivalents (NCEs) when evaluating growth; statistical analyses using the GMRT-4 utilized these NCEs. Students who maintain about the same NCE from one year to the next, are progressing at an
expected developmental rate. The GMRT-4 manual further states: “if a student’s Total score is less than 3 NCEs above or below his or her Total score from the year before, there is at least a 20% chance that the student’s relative achievement has not actually changed” (MacGinitie, MacGinitie, Maria, & Dreyer, 2000). The average scores across all groups (cases, other struggling readers, and non-struggling eighth graders) were below this number, suggesting no changes in relative performance.

Since the focus of this study, and the PARLI framework, is reading comprehension, the GMRT-4 subtest of Reading Comprehension will be the central data used in evaluation. Overall, the quintain had an average change in reading scores of 5.5 NCEs on the GMRT-4 and 51 Lexile points on the SRI. In addition, it was planned to use the teacher-made Common Assessments for Communication Arts as part of this analysis; however, with average scores for the eighth grade overall of 60% and below across all eighth graders in all five middle schools of the research site district on these teacher-mademade Common Assessments, this data was determined to not be illustrative of closing the gap and were not included in the analysis.

The GMRT-4 is the single quantitative measure comparing the quintain with their non-struggling peers; these peers averaged 16.68 months, or 4.55 NCEs of growth on the GMRT-4. To better understand the results for struggling readers, having more than one data point to consider, archival data was evaluated for likewise identified struggling eighth grade readers receiving remediation in the other four middle schools of the test district. These students showed an average gain of 16.95 months, or 4.9 NCEs, on the GMRT-4 (Comprehension subtest) over the course of the school year. Comparative Lexile data were only available at the end of the first semester, showing that the quintain
averaged growth of 37 Lexile points, and the four other middle schools combined
averaging 46.75 Lexiles in growth on the SRI. Further, since moving students in
remediation at least to grade level reading comprehension is the goal of this remediation,
it is worthwhile to compare the percentages of students moving into proficiency between
the quintain and the other struggling adolescent readers. While the other struggling
adolescent readers receiving remediation showed greater growth on both the SRI and the
GMRT-4, 30% of the students in the quintain moved into the proficient category on both
the SRI and the GMRT-4, while only 10.5% of the other struggling readers did so on the
SRI, and 13% did so on the GMRT-4, suggesting that growth relative to the goal need
also be considered.

A Kruskal-Wallis Test revealed that the difference in growth in NCEs across the
three groups was not significant (Gp0, n = 226: eighth graders not receiving reading
remediation, Gp1, n = 58: eighth grade struggling readers receiving remediation, Gp2, n
= 10: quintain), $\chi^2 (2, n = 294) = .378, p = .828$. The quintain received a higher median
score ($Md = 5$) than the other two groups, which both had recorded median values of 4.

A Mann-Whitney U test revealed the maintenance of the significant difference
(the gap) between the quintain ($Md = 43, n = 10$) and their non-struggling peers’ ($Md =
61, n = 226$) GMRT-4 scores on reading comprehension for the post test, $U = 369, z = -
3.61, p = .00, r = .02$.

Cases

As highlighted in Chapter 4, each case selected his or her own pseudonym for the
study. As a basis for deeper understanding of the quality of a given response, where
appropriate, the original question is bracketed and, where warranted, includes a model response for comparison. The student response immediately follows the brackets.

**Case study: Shenala.**

**Previous years of schooling.** Shenala (Case 1) attended a private parochial school for most of her elementary years and records were not available before fifth grade. During fifth grade, she was provided with reading tutoring, but no specific records were available.

According to school district records, for fifth through seventh grades, Shenala, has performed at a Basic level on the MAP test in Communication Arts. Her performance has changed very little since fifth grade, with her greatest increase being 14 points between sixth and seventh grades. Her classroom performance in core subjects was consistently mediocre for sixth and seventh grades, with Cs and Ds dominating with the occasional B. Her GPA remained steady at 2.6 across both years in large part due to her strong performance in elective courses.

During this study Shenala did not receive reading services in middle school prior to eighth grade. During the study year, her initial Lexile of 886 was on the edge of the Proficient/Has Met Standard range. This SRI growth was a bit stronger than her relative improvement in MAP score. Shenala’s Comprehension scores on the GMRT-4 in the 3 years prior to the study varied a bit, going from at grade level in fifth grade (5.8 in the spring), to below grade level in the following 2 years, with scores of 5 and 5.4 respectively. Shenala initiated contact with the teacher/researcher at the end of the seventh grade year, requesting intervention for the following year, as she felt herself slipping below proficiency and was concerned. While her scores resulted in flagging for
further watch in both sixth and seventh grades, her teachers did not express any concerns formally.

**Research questions.**

Research question. *Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?*

The PARLI framework was effective for Case 1. Appendix G features the Event Flow diagram for Case 1 that provides a graphic representation for the data.

**Narrative formative assessment.** On the Narrative formative assessment, Shenala scored in the Basic range, at Level 3, Stage C. She demonstrated *Personal Agency Beliefs* in the *Robust* category through this level of understanding, and moved down to *Accepting* for the higher levels.

Metacognitively, Shenala’s quality of thought and articulation of her understanding at the concrete levels of comprehension were *OK*, while those at the inferential levels were *Weak*:

> [What is different about how Harry is with his father in public than when they are home? Model--When Harry is with his father in public, he is embarrassed because of the parrot his father talks to. At home, though, they eat together, tease each other, and have a good time.] Harry doesn’t really talk to his dad in public but when they get home they have a lot to say to each other.

As an example of the *Weak* metacognition exhibited at inferential levels, the following example required a paragraph-long answer, containing evidence from the beginning, middle, and end of the piece. It is presented with all grammar, punctuation and
spelling as in the original: “came to the shop every day to buy candy To embarresed [sic] to go because of Rocky. Finds out his dad was always wondering where Harry was.”

*Narrative developmental work.* During think-alouds, Shenala made almost the same number of moves at Level 1 and Level 4, with a few moves at Level 2 as well. At Level 1, her moves were split between simple comments about the text and miscues that went uncorrected. At Level 4, her strongest thinking was shown when she reflected on the story at the end:

So, it will probably be sad for her and she might want to go with him and then they’ll be happy together. But, since he already left, she won’t be able to see him again, which I kind of think is sad. But, I get why it’s called “I See You Never” now.

Throughout the time devoted to class work during the Narrative unit, Shenala consistently produced Level 3 and Level 4 responses. Her *Personal Agency Beliefs* were *Robust*, with self-evaluations that were accurate and she exhibited confidence in her work. While she did little thinking aloud during the think-aloud protocols in this unit, her reflections at the end of the process demonstrated both understanding and confidence in that understanding at the inferential level.

In contrast, through the majority of the work she did not display evidence of strong metacognition. With limited breadth of thinking, her outcomes were split between poor quality (*Weak, Wrong, Miss*) and *OK*, and very few occurrences of *Strong* thinking. An example of *OK* metacognition at Level 3 includes an answer and a brief explanation, but none of the other aspects that reflect complex thinking: “*Answer*--When Mrs. Grave is near a plant every flower turns brown. *Explanation*--I say this because when Sara and Seth gave Mrs. Grave some flowers and they went brown and acted like it never happened.” An example of *Poor* metacognition was when she did not read the entire line
of text in one of the think alouds, skipping over all of the following: “mouths inadvertently merged in a kiss. It happened somehow inadvertently. Another kiss followed the.”

**Narrative summative assessment.** On the Narrative summative assessment, Shenala scored again in the Basic range, improving to the top of the range with a score of Level 4, Stage A. She also attempted higher levels during this assessment, but was not successful. Her self-reflection of agency at the end of the assessment was *Modest*, but evidence of agency throughout the assessment was variable, being on the positive end of the scale for the concrete levels of understanding, and shifting to the neutral range of the scale, with some evidence of negative feelings of agency, at the inferential levels. An example of Shenala’s Level 4 work on this assessment shows her appropriate understanding but lack of supporting details that demonstrate metacognition. It should be noted that each question at the inferential level reminded students to provide evidence or support for their thinking.

> [Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story?] Roger doesn’t take Mrs. Jones purse when he has the opportunity because he started to learn his lesson and she was treating him nicely.

Shenala’s *Modest* reflection at the conclusion of this summative assessment is illustrative of her progress in this unit. At inferential levels, thinking is not detailed and she is aware of this, as evidenced in her reflection. She is lacking quality and detail throughout when this was apparent in class work by the end of the unit:

> I think this assessment was appropriately challenging because it made me start thinking more and understanding a lot more and each level was a little harder but still was kind of easy. I thought it was kind of hard but I’m getting better, and understanding it.

Evidence of Shenala’s confidence in her understanding can be found in her
attempts to respond at the higher levels. However, these same examples demonstrate the complexity of developing metacognition as they illustrate both weaknesses in metacognition at this point in the study, while at the same time marking some beginning attempts to justify her thinking. The first is an example of a Weak answer at Level 5: “Takes Roger in her place. Makes him get cleaned up. Asks certain questions about stuff connected to why he tried taking her purse.” This can be compared to an example of progress, with OK thinking: “I think when she asks the questions he learned a lot because he started feeling bad about what he did.”

*Narrative nonfiction formative assessment.* On the Narrative Nonfiction formative assessment, Shenala scored in the Proficient range at Level 5, Stage A. She started off strong regarding *Personal Agency Beliefs* with evidence of a Robust sense of agency through the first half of Level 4, then dropping to Modest at Level 5. Interestingly, her personal reflection at the conclusion of the assessment was lower than the evidence would make one anticipate, being at the Vulnerable category:

I think this assessment was appropriately challenging because . . . the questions were hard but still easy enough to find the answer. I am really surprised cause when I wrote some of the ones my feeling didn’t give me much of a positive feeling about the question, but I’m really happy about how much I approved [sic] with everything Mrs. B’s teaching me.

A specific example of a Level 5 response, shows developing skill at this inferential level:

[Although Eleanor was a shy, self-conscious girl, she was one of the great women of history. How do you think she was able to do this? Use evidence from the essay to support your opinion.] I think Eleanor Roosevelt was able to do what she did because she always believed in herself and it seemed as she never gave up and always tried harder. Like when she encouraged people to write to her, she new [sic] what was happening and supported them in everway [sic] she could. When she tried harder she always achieved it and joined something else.
Metacognitively, Shenala’s quality of thought and articulation of her understanding at the inferential levels of comprehension was Good, as shown in this Level 4 response:

[How does Eleanor’s essay on ambition predict the way she would spend the rest of her life:] Eleanor’s essay on ambition predicts the way she would spend the rest of her life because she’s trying to tell people to do more and not to just be hiddin [sic] out of all the grand and great things shes [sic] trying to get people to do something and be known rather than having no ambition.

_Narrative nonfiction developmental work._ During think-alouds in Narrative Nonfiction, Shenala continued to divide her moves largely between Level 1 and Level 4. At Level 1, she continued to be split between simple connections or comments and miscues that went uncorrected. At Level 4, her strongest thinking was apparent when she reflected at the end of the story:

> I kinda was like, wow, they got out of there so good and they made it alive. I think it would have been hard for them to do that because they had like no food, they didn’t have water or anything. They didn’t have any other clothes.

Continuing her growth from the Narrative unit, throughout the time devoted to class work during the Narrative Nonfiction unit, Shenala consistently produced Level 4 responses. Her _Personal Agency Beliefs_ were Vulnerable, reflecting the circumstance that she did more work at home than at school. Observation notes indicate how she worked very slowly, taking 2-3 times longer than the other students did to complete a task. During observations of Level 4 work, evidence for work predominantly in the positive range of quality thinking (_OK, Good, or Strong_) was present. Most often she asked questions, indicating an awareness of her own thinking, enough to warrant asking others for clarification. For example, during one of the discussions of the full-length memoir
read during this unit, she asked her classmates, “But, why would he choose a young one?”

*Narrative nonfiction summative assessment.* On the Narrative Nonfiction summative assessment, Shenala scored again in the Basic range, dropping to the top of the Basic range with a score of Level 4, Stage A. Her *Personal Agency Beliefs* were categorized as *Robust* for the first 3 levels. She then moved from neutral *Acceptance* part way through Level 3 through half way through 4, then to negative from that point on. Her overall assessment was *Vulnerable*:

I think this assessment was appropriately challenging because . . . It was easy enough questions that I could answer but I just didn’t get enough time. I think this is okay but not great because we are about to be learning more on Level 5 but I think I should improve on Level 3 and 4 still because I’m not that great at it but I hope I can get really good at Level 5 and still be good at Levels 1, 2, 3, 4.

Her confidence at Level 4 coincided with answers of mostly *Good* quality, showing stable understanding at this first level of inference, even though she did not use direct examples from the text to support her thinking

*Expository formative assessment.* On the Expository formative assessment, Shenala scored in the Basic range, at Level 4, Stage A. She expressed negative *Personal Agency Beliefs* with evidence of a *Vulnerable* sense of agency throughout the assessment, not attempting to answer any questions after Level 4, sharing “I did not answer this because the class ended and I didn’t have enough time,” for the first question she did not answer. Her personal reflection in the *Vulnerable* category at the conclusion of the assessment matches her work:

I think that this assessment was difficult because . . . some of the questions were hard and I couldn’t understand what some of them were asking but the rest I think I answered pretty good. I think this is good but I hope I can do better and maybe higher.
Shenala’s thinking throughout the assessment is predominantly *Good* quality, but with some weaknesses at Level 3. At Level 3, it seems that she was answering based on ideas outside of the text, showing less command of metacognition. An example response from Level 3 illustrates this:

The Shang discovered Mars, this tells us that the Chinease [sic] people were really smart, especially in astronomy. I know this because they kept records of solar and lunar eclipses, the stars, and other events like comets. Even today Chinease people are still smart, like at making things, because almost everything you see says the words “Made in China,” on about everything.

*Expository developmental work.* During this unit, Shenala was absent frequently, and was always behind in her work. During the think-alouds, her moves were consistent with the previous two units, being predominantly split between Level 1 and Level 4, but she made Level 2 moves as well. Her Level 2 moves were most often miscues that went uncorrected. The higher quality of thinking moves at Level 4 involved predictions: “I think that this is going to be about like something historical from what I see on the page. But I never really heard about Sequoyah and the Cherokee.”

During the remainder of the unit coursework, Shenala collaborated with Alice to focus exclusively on Level 3 and Level 4 work. Her frequent absences affected the work production and the benefits of the collaboration, as her partner carried on in her absence, resulting in each of them doing more independent work and less collaborative work than the other teams. The quality of her thinking was most often *OK*, while the number of quotes used to support thinking was the highest of the study for this case, indicating a developing awareness and understanding of the need to support thinking with text. An example from Level 4 work during this unit is illustrative of Shenala’s exponential shift
to providing evidence of 5 times her earlier rate: “. . . like Confucius did when he said, ‘Alas! Heaven is destroying me!’ he cried, ‘Heaven is destroying me!”

*Expository summative assessment.* On the Expository summative assessment, Shenala scored again in the Basic range, dropping to a score of Level 3, Stage C. Her *Personal Agency Beliefs* were again categorized as *Vulnerable* overall:

I think that this assessment was difficult because the questions kind of got harder after words on Level 4 but the beginning was kind of easy. I think this is okay for me but I think I could’ve gotten better and wish I did get better.

Answers up through Level 3 all show *Good* quality and include support from the text in the form of quotes, showing Shenala’s awareness of this requirement and ability to meet it. After Level 3, she made no attempt to share any inferential thinking or support her thinking from the text.

*Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*

*GMRT-4 and SRI.* Shenala’s performances of Advanced on the GMRT-4 and Proficient on the SRI are evidence of her closing the gap. Specifically, she demonstrated a growth of 46 months or 18 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Shenala’s growth in her SRI score through the end of the year was 159 Lexile points, or more than three times the average of 51 for her counterparts in the quintain. She showed most of this Lexile growth at the semester, supporting the growth seen in regular class activities with performance on standardized measures that
was not always displayed in assessments at the end of the study framework implementation.

**PARLA.** Shenala exhibited fluctuating performance consistent with Dynamic Skill Theory’s development of complex skills (Fischer & Bidell, 2006) across the three content areas, as measured by the PARLA proximal assessments, and did not succeed in closing the gap to grade level performance on them. Specifically, on the PARLA-NARR she improved by 3 stages from pre to post assessment, doubling the average growth of the members of the quintain. On the PARLA-NARR NF, she dropped by 3 stages, as compared to the average of just under 4 stages growth for the rest of the quintain. Finally, on the PARLA-EXPO she again underperformed relative to her immediate peers, falling back 1 stage relative to the average of near 3 stages of growth. Her attendance averaged 2 days per week from just after the PARLA NARR NF pre-assessment through the final PARLA assessment in December (PARLA EXPO).

*Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?*

Shenala’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed positively by 2 categories, going from the Vulnerable category to the Tenacious (see Table 1. The MST Taxonomy of Personal Agency Beliefs). The most prevalent categories for her throughout the work across the three content areas were: Robust, Modest, Accepting, and Vulnerable, in that order, as shown in the pie graph of Figure 8.
Figure 8. Distribution of most prevalent PABs in Case 1.

Shenala’s reflections on the hierarchical proximal measures (PARLA-NARR, PARLA-NARR NF, and PARLA-EXPO), provide a more detailed glimpse of Shenala’s stance relative to the framework than her quantitative performance demonstrated. She was one of the few students to not complain about being asked to reflect, and was consistently forthcoming, which matches with her tendency to take a great deal of time and care with her work, often choosing to turn work in late or not at all if she did not have the time to complete it to her satisfaction.
Across the three content texts—Narrative, Narrative Nonfiction, and Expository—the most prevalent categories of Personal Agency Beliefs varied: The Robust rankings moved from 31% during Narrative, up to a high of 56% during Narrative Nonfiction, and back to 33% during Expository. The Personal Agency Belief of Accepting was at the greatest level of 47% at the start of the study, dropped during Narrative down to 16%, and rose to 21% in the final unit. The Modest belief was least vulnerable to change across contents, as shown in Figure 9. The lowest category of PAB for this case was the Vulnerable category which was low at 9% during Narrative work, but rose with the challenges of Nonfiction texts in Narrative Nonfiction and Expository work. Robust was a Personal Agency Belief expressed six times during Narrative, 13 times during Narrative Nonfiction and four times during Expository. There was only one instance of Robust PAB at the inferential levels, and that took place during Narrative Nonfiction. Evidence of

Figure 9. Case 1: tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.
inferential thinking did not accompany any of the PAB ratings of Accepting, while when it came to Vulnerable, there were two instances at the literal levels of text, and one at the inferential.

*Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?*

*MARSI.* On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Shenala’s self-reported changes in metacognition from the MARSI were greatest in the areas of Global Reading Strategies and Support Reading Strategies, with changes from a rating of 2.5 to 2.9 and 2.6 to 3.0 respectively from the start of the study to its conclusion. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. Support Reading Strategies encompass a variety of actions that include note taking, paraphrasing, and discussing the material, among others. Shenala’s self-rating improvement for these strategies moved her into a high rating. Finally, Shenala reported an increase from 3.5 to 3.75 over the study period in Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting and similar strategies.

Considering the places on the data matrix where assessments meet the MARSI, the growth in self-reporting of use of reading strategies support Shenala’s improved reading scores but are not particularly dramatic. When considering her reflections on the pre-study measure shown in Table 6, the reasons become clear. The most noteworthy is her comment regarding Support Reading Strategies, as those are the most time-
consuming of the metacognitive reading strategies evaluated with the MARSII and, given her issues with time, this makes sense.

Table 6.  
**Reflections on MARSII survey by reading strategy for Case 1.**

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I use about ¾ of these all the time.”</td>
<td>“I use about all of these all the time.”</td>
<td>“I don’t really use these methods.”</td>
</tr>
</tbody>
</table>

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. Level 1 and Level 4 were almost equally represented among the moves in Shenala’s think-alouds, as shown in Table 7.

Table 7. **Frequencies of Levels of think-aloud protocols for Case 1**

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.53</td>
<td>.16</td>
<td>0</td>
<td>.31</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.36</td>
<td>.28</td>
<td>0</td>
<td>.36</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>0</td>
<td>.25</td>
<td>0</td>
<td>.37</td>
<td>.02</td>
</tr>
</tbody>
</table>

*PARLA/PARLI* During the Narrative segment of the PARLI framework, Shenala attempted responses across all of the levels, 51% being at Level 3 and Level 4, and the junction of literal and inferential comprehension, as shown in Table 8. The *Quality of*
response (Table 9.) included 22% considered to be *Good*, but was predominantly *OK* at 43%.

Table 8.
*Case 1: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>114</td>
<td>.18</td>
<td>.16</td>
<td>.28</td>
<td>.23</td>
<td>.10</td>
<td>.01</td>
<td>.04</td>
</tr>
<tr>
<td>Narrative</td>
<td>143</td>
<td>.28</td>
<td>.30</td>
<td>.15</td>
<td>.26</td>
<td>.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>158</td>
<td>.26</td>
<td>.24</td>
<td>.28</td>
<td>.21</td>
<td>.01</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were fairly well distributed across Level 1 through Level 4, with 15% at Level 3, and 1% each at Level 5 as shown below in Table 7. The *Quality* of response (Table 9.) was at the midpoint or above (*OK* and up) 80% of the time.

Table 9.
*Case 1: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>90</td>
<td>.10</td>
<td>.12</td>
<td>.43</td>
<td>.21</td>
<td>.14</td>
<td>0</td>
</tr>
<tr>
<td>Narrative</td>
<td>117</td>
<td>.01</td>
<td>.22</td>
<td>.57</td>
<td>.14</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td>Expository</td>
<td>140</td>
<td>.01</td>
<td>.17</td>
<td>.60</td>
<td>.17</td>
<td>.02</td>
<td>.03</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the Levels of 78% of responses were at the concrete levels of comprehension shown in Table 8. (Level 1-Level 3), while the *Quality* of responses were split 18% positive (*Strong* and *Good*), 60% in the mid-range, or *OK*, and 22% in the lower tiers of *Quality* shown in Table 9.
Shenala’s classroom performance showed the greatest growth throughout the Narrative Nonfiction unit when she worked within the Context of solitary work. Her performances at Level 1 through Level 4 in the framework were fairly consistent, but attempts at Level 5 were dependent upon her taking them home for completion.

Starting with this unit, Shenala was absent no less than twice a week for the remainder of the study, putting her in a position of always being a bit behind. The recurring theme of her academic performance was one of a lack of time. Her unusual mix of average verbal fluency, with profoundly slow silent reading and processing time, and exceptional quality of work product when no time limits were placed was confounding to her core teachers as well. During the Narrative Nonfiction unit, along with her accelerating absences, her U.S. History, Communication Arts, and Science teachers all sought the researcher out to try to gain understanding of this student. They were generous in their willingness to give her the time she needed, and her resulting grades were mostly in the B range.

Sub-question. 4) Does student performance on the assessments form a pattern of development?

Of the quantitative measures, Shenala’s standardized reading assessment scores (GMRT-4 and SRI) improved consistently, and considerably. The PARLA performances form a predictable pattern, with improvement following instructional intervention, but are volatile, in a pattern that fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a variable process.

Summary. Shenala performed at the Proficient level on both standardized reading assessments, although only on one of the proximal measures (PARLA-Narrative
Nonfiction, formative assessment). Her formative and summative assessments of agency and motivation show growth from a starting place near the middle of the scale of *Vulnerable*, and ending place near the top of *Tenacious*. Her metacognitive scores on the MARSI survey changed only slightly, but all in a positive direction. Consideration of the evidence of her development regarding reading comprehension, agency and motivation, and metacognition across the course of the study, one can make a case for the PARLI framework being effective with Shenala.

**Case study: Rafael.**

*Previous years of schooling*. Rafael (Case 2) was on an Individualized Education Plan (IEP) for speech in early elementary and was exited at the start of third grade. He participated in Reading Recovery in third grade, and progressed at a pace meeting grade level expectations. His record was incomplete in fourth grade, but based on his GMRT-4 reading score at the end of that year, he was placed on a Reading Improvement Plan in fifth grade.

Rafael was one of the students who started with the researcher in sixth grade. He arrived to my classroom a polite and enthusiastic young man, who informed me that he sometimes liked reading stories, was not good at reading for core classes, and preferred playing tennis, which he did on a competitive level, to doing all else. Rafael was an optimistic young man who embraced each new opportunity. He regularly sought clarification, but often did not maintain new understandings over time until he had multiple opportunities to regain the insight. He got off track with some frequency, but redirected promptly.
There was a big difference in his work production and the quality of his work based on his work partners; he rose to the level of students who were performing well. To his credit, he tried to step up when he was partnered with someone who was struggling a bit, but often faltered. He made significant progress in his abstract reasoning and reading comprehension since sixth grade, but his classroom performance remained a bit of a roller coaster. Rafael wanted very much to please adults and frequently seemed to get distracted. He seemed genuinely surprised when his performance dropped, and was overly optimistic when he was performing well that it would remain thus. He did not read directions carefully, and performed best after several explicit modeling opportunities.

According to district records, at the end of both fifth and sixth grades, Rafael performed at a Basic level on the MAP test in Communication Arts. At the end of seventh grade, he dropped to Below Basic, but his actual score was virtually identical to previous years (the standard was raised). Rafael’s SRI score was at the Proficient/Has Met Standard range at the end of seventh grade, but dropped back to just below this at the start of his eighth grade year, putting him in the Basic/Partially met range; this was a bit stronger for a relative performance than his MAP score. Rafael’s Comprehension scores on the GMRT-4 in the last three years rose steadily, going from 3 at the end of fifth grade, to 4.4 in sixth, and 5.6 in seventh grade. Rafael gained more than 1 year in each school year, but remained behind.

Research questions.

Research question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension,
agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was not effective for Rafael. Appendix G features the Event Flow diagram for Rafael that provides a graphic representation for the data.

**Narrative formative assessment.** On the Narrative formative assessment, Rafael scored in the Basic range at Level 3, Stage A. He demonstrated *Personal Agency Beliefs* predominantly in the *Robust* and *Modest* categories (with one example of *Accepting*) up through the basic inferential level of understanding, and moved down to *Discouraged* at Level 5. Rafael was incorrect on one of the Level 1 questions, did not attempt one of the Level 3 questions, and demonstrated a basic (*OK*) understanding on the Level 4 questions before being wrong at Level 5. Specific examples of his Level 1 and Level 5 responses, along with his evaluation of them demonstrate this. At Level 1: [Why does Harry go to the store at the end of the story? ] “To watch the bird.” He earned an *Accepting* rating for the mismatch between his actual score and his confidence in that score. At Level 5: [How does Mr. Tillian’s behavior affect Harry throughout the story?] “It effects harry [sic] because he’s embarrassed because his dad talks to a parrot all day.” This response earned a *Discouraged* rating because he did not rate his confidence.

Metacognitively, Rafael’s *Quality* of thought and articulation of his understanding at the concrete levels of comprehension were a mix of *Good* and *OK*, while those at the inferential levels were a mix of *OK* and *Wrong*. An example of metacognition at the concrete level that is *Good*: [When does Rocky arrive? Model--Rocky arrives the year Harry turned 12.] “When hary [sic] turns 12 and doesn’t come to the store anymore.” An example of metacognition at the concrete level that is *OK*: [What does Harry do when he
turns 12? Model--When Harry turns 12 he stops coming to his father’s shop because the parrot embarrasses him. He and his friends did other things instead.] “Not visit the store anymore.” An example of metacognition at the inferential level that is OK shows similar brevity:

[Why does Harry keep walking past the shop after Rocky arrives? Model--Harry keeps walking past the shop after Rocky arrives because he is embarrassed that his father is talking to the parrot instead of people.] To check on what his dad is doing.

Narrative developmental work. During the unit devoted to Narrative texts, Rafael worked most at Level 4 followed by Level 1. Level 4 was almost exclusively done through annotation. The exceptions were the responses on the Concept Diagram for Level 4. It should be noted that students were allowed to collaborate on this, but observations support that Rafael seemed to be able to articulate the nature of thinking at this Level conversationally as well. His Personal Agency Beliefs were Robust, Modest, and Self-Doubting, with self-evaluations that were as inconsistent as his variable performance overall. An example of his Self-Doubting PAB can be drawn from think-alouds as well. He did not reflect at all after two of the think-alouds. His reflection for the second one: “That was a confusing story. I really didn’t get it that much” rather effectively sums up much of his experience.

Similarly, through the majority of the work he did not display evidence of strong metacognition, with just over 10% more positive quality ratings (Good, Strong, OK) than negative (Weak, Wrong). His Weak responses were mostly at Level 1, due to miscues during think-alouds

Narrative summative assessment. On the Narrative summative assessment, Rafael scored again in the Basic range, with no correct responses beyond Level 4, Stage A. His
self-reflection of agency at the end of the assessment was *Fragile*: “I think that this assessment was difficult because the assessment got harder as it went on.”

Evidence of agency throughout the assessment was variable, being on the positive end of the scale (*Robust and Tenacious*) to half way through Level 3, then neutral to negative (*Modest, Accepting, and Discourage*) for the second Level 3 and through the inferential levels. An example at Level 4 shows the prevalent nature of Rafael’s work; even when he is accurate he generally lacks detail: [Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story?] “He didn’t want to because she would even get more mad at Roger.”

At the lowest concrete levels of comprehension, Rafael demonstrated solid metacognitive processes. At concrete Levels 1 and 2, quality of thinking was *Good* or *OK*, and predominantly *OK*. An example of Level 2 considered *OK* illustrates:

[How does Mrs. Jones find out if Roger has an adult at home to care for him? Model--Mrs. Jones finds out that Roger has no one to care for him by asking whether he has anyone at home to tell him to wash his face.] Because he had a dirty face and the lady said where’s your parents and Roger said there’s nobody at home. (2:Q11:12)

During Level 3, Rafael’s reflection on one of the questions is indicative of the breakdown of his metacognition: [At the end of the story, what reason does Mrs. Jones give for understanding Roger’s behavior when they first meet?] He did not respond, reflecting: “I didn’t understand the question.”

Rafael’s reflections were characterized by the absence of metacognition, with him mentioning a detail from the scoring guide, without any demonstration of understanding of how his answer differed and without making any connections. A reflection at Level 4 illustrates: “He really just didn’t want to be mistreated.”
Narrative Nonfiction formative assessment. On the Narrative Nonfiction formative assessment, Rafael scored in the Below Basic range, at Level 2, Stage C. Little detail was present in any of his answers, and his reflection regarding Personal Agency Beliefs mirrors this, with an Accepting sense of agency overall. He did not answer one of the Level 2 questions correctly: [What was the first group that Eleanor joined that led her to so many other things? Model--Eleanor decided to join the League of Women Voters.]

“She helped pass laws to improve educational living and working.” Upon reviewing the correct response, he reflected: “Didn’t get the question, now I do.”

Throughout the study, Rafael’s PAB was variable, but the instances of Discouraged (two times) are noteworthy. An example from Level 3 illustrates:

[In an essay written when she was 14, why did Eleanor say that it is easier to have no ambition? Model--She said it was easier to not have ambition because you won’t have to face difficulty of disappointment.] Because she said to not try hard.

Rafael rated his confidence as a 2 and did not reflect on this response, for which he earned no points, showing that his state at that time was Discouraged. Observations support this Discouraged PAB, as he did not reflect in anticipation of the assessment, was finished quickly, and then put his head down while waiting for peers to finish.

Rafael’s personal reflection at the conclusion of the assessment largely matched the evidence throughout, being at the Accepting category: “I think that this assessment was difficult because it was nonfiction. I think my score is right because nonfiction was a lot harder.”

Again, consistent with Rafael’s poor performance overall, his thinking is predominantly of Weak and Wrong quality. An example of Weak from Level 4 illustrates:

[How does Eleanor’s essay on ambition predict the way she would spend the rest of her life? Model--The opinions expressed in her essay on ambition showed that
she thought it was not acceptable to not try to do good work and make a
difference. She talked about how important that was, then she did a great deal of
work to help a lot of people. She is remembered for many things, including many
things that she was the first woman to do.] She tried to go by that.

_Narrative nonfiction developmental work._ Throughout the time devoted to class
work during the Narrative Nonfiction unit, Rafael consistently focused on Level 4
responses. One example of this comes from one of the discussions:

It kind of seems like in Yemen men are on top of the whole thing, (gesturing at a
level with his hand) and in the United States, we’re all kind of equal, men,
women. Over there, it seems like men are on top and the women have to do
everything.

Rafael’s _Personal Agency Beliefs_ were most frequently _Robust_ at Levels 1 and 2,
followed by _Modest_ at Level 4. During the work with the memoir was the one time
during the study that he maintained his work on schedule. During observations of Level 4
work, evidence for work predominantly in the _OK_ range of quality thinking was present.
Most often he provided answers to others and gave explanations for his thinking. He did
not exhibit a level of engagement that included asking questions. He predominantly
worked at responding to others, and not at developed questions to seek clarity or push his
understanding level up. Observational field notes reveal that Rafael did a great job of
staying up with the reading and annotating during _I Am Nujood, Age 10 and Divorced_
(Ali & Minoui, 2010) and was a significant participant in all of the discussions. All of his
discussion was at the level of inference, but his responses were sparse and shallow.

_Narrative nonfiction summative assessment._ On the Narrative Nonfiction
summative assessment, Rafael scored again in the Basic range, improving to the top of
the Basic range with a score of Level 4, Stage A. His _Personal Agency Beliefs_ were
highly variable, spanning all ten categories and often widely distributed among them. For
example, at Level 3 his _PAB_ ratings included seven different categories and his Level 4
PAB ratings included Robust, Hopeless, and Modest. Rafael’s overall assessment was Vulnerable: “I think that this assessment was difficult because I think that this assessment was difficult because the story was a little confusing.” When reflecting on individual questions, he shared his frustration: “The question did not make any sense and I couldn’t find any detail.”

The dramatic variation in Rafael’s PAB indicates inconsistency in thought and lack of control over cognition. His answers were of variable quality through all four levels, indicating a lack of stability of quality of thinking. He did not use direct support from the text, suggesting that correct answers and why they are correct were not under his metacognitive control.

Expository formative assessment. On the Expository formative assessment, Rafael scored in the Below Basic range, at Level 1, Stage C. He expressed negative Personal Agency Beliefs answering all questions at Level 1, skipping one at Level 2, and stopping at Level 3. His personal reflection in the Hopeless category at the conclusion of the assessment matches his work: “I think that this assessment was difficult because this assessment was very difficult because the story was long and I didn’t get the questions.”

The Quality of Rafael’s responses was likewise highly variable, from Wrong through Good, with no progression to the pattern of quality that matches the increased depth of the academic task. His pre-assessment reflection may be most illuminating about where his thinking was: “Before I take this assessment, I am feeling nervous because it is about Ancient World stuff.”

Expository developmental work. Level 4 was dominant, but that was a result of the instructor setting that requirement, rather than Rafael choosing to work at the
inferential level. His PAB based on observational field notes was variable. Much of this unit was done collaboratively with Dominique. He was operating in the negative range regarding PAB throughout the unit. As a pair, they were consistently behind on this work.

The Quality of Rafael’s responses was predominantly OK, but a large number of examples of lower Quality levels (Weak, Missing) were also present. Pattern of thought is parallel to Narrative, with the poor quality responses coming largely from the independent think-aloud protocols and these being the lower level miscues and omissions. It is noteworthy that there were no Good responses at Level 4, showing his lack of readiness to engage consistently in inferential work.

Expository summative assessment. On the Expository summative assessment, Rafael scored in the Basic range, improving to a score of Level 3, Stage A. His initial, pre-assessment PAB was Self-Doubting: “Before I take this assessment, I am feeling scared because I have not done great on all the other ones.” His post-assessment reflection was Hopeless: “I think that this assessment was difficult because the answers did not make sense and the questions did not make sense.”

Interestingly, the Quality of his answers was predominantly Good but he did not attempt past the first question of Level 3. This illustrates the power of PAB in his thinking process, as he lacked the motivation to try beyond where he was absolutely confident in his thinking skills.

Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?
GMRT-4 and SRI. Rafael’s persistent performance at the Basic level on both the GMRT-4 and the SRI are evidence of his failure to close the gap. He demonstrated a growth of 1 month or a drop of 4 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Rafael’s decrease in his SRI score through the end of the year was 34 Lexile points, placing him well below the average performance of 51 for his counterparts in the quintain.

PARLA. Rafael made gains in all three content areas, as measured by the PARLA proximal assessments, but remained below grade level expectations in all areas. Specifically, on the PARLA-NARR he improved by 1 stage from pre to post assessment, placing him just below the average growth of the members of the quintain. On the PARLA-NARR NF, Rafael improved by 3 stages, as compared to the average of just under 4 stages for the rest of the quintain. On both the Narrative and Narrative Nonfiction areas his final performance was at Level 4, Stage A, just beginning to understand basic inferences. This performance was consistent with Rafael’s written class work, but lower than the thinking he demonstrates in discussion. Finally, on the PARLA-EXPO, he showed his greatest growth relative to his immediate peers, achieving 6 stages of growth relative to the average of near 3 stages. His final performance categorized him as emergent at the last level of literal understanding. This achievement level is anticipated by the fall of the sixth grade year, placing him just over 2 years behind grade level.

Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?
Rafael’s self-reporting about his motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Tenacious to the Vulnerable category (see Table 4, *The MST Taxonomy of Personal Agency Beliefs*). For Rafael, Robust was the most prevalent PAB rating by a sizeable margin, with Modest, and Tenacious being the other categories of some depth, as shown in the graph in Figure 10.

*Figure 10. Distribution of the most prevalent PABs across the study for Case 2.*

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied: The Robust rankings moved from 29% during Narrative, up to a high of 31% during Narrative Nonfiction, and down to 16% during Expository. The Personal Agency Belief of Modest was at the median level of 28% at the start of the study, rose during Narrative to 39%, and came back down to 21% in the final unit. The Tenacious belief was stable for the first two units, at 18%, then doubled during the final (Expository) unit to 36%, as shown in the graph in Figure 11. Robust represented quotations without Quality ratings and was expressed a total of 15 times at the inferential levels of comprehension. At the literal levels of comprehension, the Modest category was accompanied by Quality ratings on the negative end of the scale seven times, spread across all three contents, in the neutral
range two times in Expository, and two more times in Expository without a *Quality* rating. In addition, three instances of *Modest* with neutral *Quality* were present at the inferential level of comprehension. Finally, *Tenacious Beliefs* were present with *Quality* ratings in the negative three times at the literal levels of comprehension, two times in Narrative and one in Expository.

![Case 2: Robust Over Time](image)

**Figure 11.** Case 2: tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

Rafael’s reflections on the hierarchical proximal measures (PARLA-NARR, PARLA-NARR NF, and PARLA-EXPO), provide elaboration of Rafael’s performance within the framework beyond his quantitative performance. During the Narrative unit, Rafael’s reflection before the summative assessment in Narrative reflected a *Modest* PAB: “I think it might be challenging but all the work it might be easier.” After completing the first Narrative Nonfiction assessment, his reflection that it was difficult because “the assessment got harder as it went on” was considered *Fragile* and reflected his general struggle with depth of reflection throughout the study. His general strategy
was to try to glean the gist of responses from his peers and camouflage his lack of strong written expression with articulate discussions during class time.

There were several occasions when Rafael would check in that he was on the right track, walk back to a table and sit down to stare at the work and make no progress. His classroom performance remained largely unchanged across the study, which was echoed by his lack of demonstrable growth on the GMRT-4 or SRI. Rafael showed improvement on the PARLA assessments in all three units, but not enough to attain grade level on any of them. His area of greatest growth was the Expository unit, but he began with an assessment at the early elementary level. Rafael’s reflections were classified as *Hopeless* on the Expository formative assessment both before and after the test.

Even though his overall achievement was well below grade level, Rafael’s growth of 6 stages was significant. Rafael does not have a history of success with reading and responding to nonfiction, and referenced the difficulty of nonfiction on numerous occasions.

*Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?*

*MARSI.* On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Rafael’s self-reporting indicated high ratings across all three categories of reading strategies for both pre and post measures. His greatest reported changes in metacognition from the MARSI were in the area of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting, and similar strategies. Rafael reported a decrease
from 4.75 to 3.5 over the study period. For the subscale of Global Reading Strategies, he indicated a slight change from a rating of 3.5 at the start of the study to 3.8 at its conclusion. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. Similarly, Rafael’s self-rating improved from 3.5 to 3.8 for Support Reading Strategies; these include a variety of actions such as note taking, paraphrasing, and discussing the material, among others.

Rafael’s lack of growth across reading comprehension measures was consistent. His stagnant performance on the GMRT-4 and his decline on the SRI, match the lack of higher levels of cognition on his think-alouds. Rafael’s only positive standing on any measures came from his overly optimistic ratings on the MARSI. His reflections on the pre-study MARSI assessment illustrate his challenges with processing and following directions. Despite an explanation of what constitutes the categories of High, Medium, and Low on the MARSI, he gave himself high marks on all three strategy sub-types, in contrast to his reflections as shown in Table 10.

Table 10.
Reflections on MARSI survey by reading strategy for Case 2.

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>He shared that, “I use about half those strategies,” while giving himself high marks.</td>
<td>“the ones I used I use well”</td>
<td>“use once in a while”</td>
</tr>
</tbody>
</table>

Rafael’s generally positive outlook and his strong verbal skills make his level of thinking during collaborative discussion misleading. Under observation, he is found to be distracted and off task more often than not. He is the only member of the quintain about
whom others complained, sharing that he was not doing his part and asking to not work with him in dyads in particular. Rafael is an affable young man, and his peers largely expressed reluctance and dismay about complaining. However, they hit their thresholds during the Narrative Nonfiction unit.

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, the vast majority of the moves in Rafael’s think-alouds were at concrete levels, as shown in Table 11.

Table 11. *Frequencies of Levels of think-aloud protocols for Case 2.*

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.66</td>
<td>.17</td>
<td>0</td>
<td>.17</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>.65</td>
<td>0</td>
<td>.04</td>
<td>.30</td>
<td>0</td>
</tr>
</tbody>
</table>

*PARLA/PARLI.* During the Narrative segment of the PARLI framework, the levels of responses were distributed across Level 1 through 5, as shown below in Table 11. The Quality of response was predominantly *OK* or *Weak* (Table 12.).
Table 12.
*C4se 2: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>122</td>
<td>0.32</td>
<td>0.13</td>
<td>0.3</td>
<td>0.2</td>
<td>0.05</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>107</td>
<td>0.34</td>
<td>0.37</td>
<td>0.08</td>
<td>0.21</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>80</td>
<td>0.35</td>
<td>0.3</td>
<td>0.21</td>
<td>0.14</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 4, with the majority being at Level 1 and 2, together combining for 71% of the responses, as shown in Table 12. The Quality of response was at the midpoint or below (OK and lower) 91% of the time, with the remaining 9% at Good (Table 13.).

Table 13.
*C4se 2: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>107</td>
<td>0.01</td>
<td>0.08</td>
<td>0.47</td>
<td>0.36</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>94</td>
<td>0</td>
<td>0.09</td>
<td>0.38</td>
<td>0.33</td>
<td>0.06</td>
<td>0.14</td>
</tr>
<tr>
<td>Expository</td>
<td>65</td>
<td>0</td>
<td>0.12</td>
<td>0.42</td>
<td>0.36</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of 86% of responses were at the levels of comprehension (Table 12.), while the Quality of responses was only rated as Good 12% of the time, with the majority of responses at or below the median score of OK (Table 13.).

An excellent example from observations clearly demonstrates Rafael’s issues with focus and work completion. He had lagged behind on turning in work prior to leaving on the previous day and assured that he would bring in all six assignments that were past
due. It should be noted that each and every day he seemed to be working. He had paper
on which he was writing and was interacting with peers. He does have a history of
misplacing his work frequently. November 9, 2010 observation entry:

Rafael brought in four responses, none of which resembled the requirements. None were scored on a rubric either. His assignments were wholly inadequate. I met with him personally. We discussed the squint test of leaning back to physically compare his work with the model. He said “Uh, oh” We then re-read the Level 4 models. I reminded him that his annotation and work in discussion has been solid, but the written work has lagged terribly. When I asked which questions he was answering (this was not clear from his responses) he said “There are questions?” This was discussed in class, written on the board, they were passed out, and extra copies were available in the room. I also brought it up daily for the past week. Rafael’s attentional issues (he spaces out a bit) are really causing some difficulty. We discussed what he needed to do, he wrote it down and agreed that he would focus on the Level Four questions, as he is not yet ready for Level 5.

Sub-question. 4) Does student performance on the assessments form a pattern of development?

Rafael’s scores have not been consistent on quantitative measures, but his PARLA scores show the most variability, in a variable pattern that fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process.

Summary. Rafael performed at the basic level on both standardized reading assessments, and most of the proximal measures, with two instances of performance below the basic level for both PARLA-Narrative-Nonfiction and pre-assessments. His formative and summative assessments of agency and motivation show decline from a starting place on the high end of the scale of Tenacious, to an ending place near the bottom of Vulnerable. His metacognitive scores on the MARSI survey started and ended in the high range on all three subscales (Global, Support, and Problem-Solving), along
with his basic to below basic performance in almost all instances, suggests ineffective metacognition about his reading comprehension throughout the course of the study. By considering the absence of evidence of his growth regarding reading comprehension, agency and motivation, and metacognition across the course of the study, one cannot make a case for the PARLI framework being effective for Rafael.

**Case study: Edward.**

*Previous years of schooling.* During an assessment of his reading ability in the third grade, Edward (Case 3) was determined to be at the third grade instructional (teaching) level for reading; however, his teacher noted that he struggled with comprehension and needed extensive scaffolding with her comment “to prompt too much with the comprehension questions;” excessive prompting indicates a weakness in understanding the written text. Continuing with fourth grade, Edward’s record indicates that he received tutoring for reading and scored in his fourth grade level. During the fifth grade, his record indicates that he performed on the fourth grade level; however, at the end of fifth grade he was determined to be reading on level but still had difficulty with comprehension as noted in the teacher’s note “he has a great deal of trouble with comprehension.”

Edward was one of the students who started with the researcher in sixth grade. He arrived to my classroom a polite and considerate young man, who informed me that he did not like reading and preferred being outdoors to being in school. He consistently performed on the lower end of the class, virtually always being in the group that needed additional small-group instruction. During the fifth and sixth grades, he scored at the fourth grade level for reading on both the Scholastic Reading Inventory (SRI) and the Gates MacGinitie Reading Test (GMRT-4). He improved one grade level on the GMRT-
4 in seventh grade, but did not show any gains on the SRI. Throughout his seventh grade year, he remained in almost constant need of additional instruction and expressed continued and significant frustration with his lack of progress. He shared that he thought that he was stupid and would never get better at reading. He had largely disengaged from class activities across all content areas by the end of the seventh grade and was just doing what he needed to do to maintain his C average. His cumulative GPA dropped from 3.0 in sixth grade to 2.86 in seventh, but it is illustrative to note that he earned straight Cs in his core courses at the end of seventh grade, where he had earned all Bs, with the exception of math, at the same time in sixth grade.

**Research Questions.**

Research Question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Edward. Appendix G features the Event Flow diagram for Edward that provides a graphic representation for the data.

Narrative formative assessment. On the Narrative formative assessment, Edward scored in the Below Basic range, at Level 1, Stage B. He demonstrated Personal Agency Beliefs in the Robust category for Level 1 only, then dropped to Accepting and Discouraged categories for the remainder of this assessment. Specific examples of his Level 2 and Level 4 responses, along with his evaluation demonstrate his performance and negative sense of agency. At Level 2:

[What does Harry do when he turns twelve? Model--When Harry turns twelve he stops coming to his father’s shop because the parrot embarrasses him. He and his
friends did other things instead]. He starts liking candy and stops by the shop every day after school.

Edward anticipated his response being inadequate by indicating a low confidence level of 2 and earning an Accepting rank for PAB. With Level 4, Edward did not rate his confidence or reflect after the fact, earning him a Discouraged rank for a question that most of the quintain answered correctly: [Why does Mr. Tillian bring Rocky to the shop?] “For people to see and because Mr. Tillian is the only one in the shop”

Metacognitively, Edward’s work showed a quality of thinking that was predominantly negative (Miss, Weak, Wrong). At the concrete levels of comprehension (Levels 1-3) it was split between Wrong and OK. An example from Level 2 highlights his best (OK) responses:

[Why does Harry go to the store at the end of the story? Model--Harry goes to the store at the end of the story because his dad is in the hospital and there is no one to work in the store. He takes care of Rocky and the store.] To help out because Mr. Tillian had gone to the hospital.

As an example of the weak metacognition exhibited at inferential levels, he did not attempt to answer the simplest of inferential questions: [Why does Harry keep walking past the shop after Rocky arrives?]

Narrative developmental work. During think-aloud work in the Narrative unit, Edward worked mostly at the inferential level, connecting what was said to his experience to make meaning/connect to the story. His PAB ratings were split between Tenacious and Discouraged. An illustrative example of Tenacious comes from the second think-aloud protocol:

Now that I finished the story, I think it was harder than the other one, because it had like different words that I haven’t really known, and harder to pronounce like the names and the words. And it was a little bit longer. And, yeah. It was pretty hard.
In a later think-aloud protocol, Edward was Discouraged: “This story kind of didn’t make sense [sic] to me, but um, its ok, but like I said its kinda confusing and long. I didn’t like this story.”

*Narrative summative assessment.* On the Narrative summative assessment, Edward scored again in the Basic range, improving to Level 4, Stage A. His self-reflections of agency prior to taking the assessment, during the assessment, and following the assessment are all indicative of lower levels of motivation and personal agency. He began with a Vulnerable self-rating: “Before I take this assessment, I am feeling kinda nervous, and thinking I can do better than the one I took before.” During the assessment, his ratings and reflections varied between Vulnerable and Accepting, with his concluding reflection and rating at Discouraged: “I think that this assessment was difficult because I couldn’t focus today.”

It is not surprising with these levels of personal agency that Edward’s Quality of thinking was variable, with more poor examples. At Level 2 and 3, examples of Wrong responses are short and not responsive to the question asked. Specifically, at Level 2: [What is the first thing that Roger does? Model--The first thing that Roger does is to try to steal Mrs. Jones’ purse.] “Wash off his face.” At Level 3, he structured a complete sentence response, but it did not answer the question:

[At the end of the story, what reason does Mrs. Jones give for understanding Roger’s behavior when they first meet? Model--She says that she also wanted things she could not get and she did things she doesn’t want to talk about.] She gives him $10 for suede shoes.

*Narrative nonfiction formative assessment.* On the Narrative Nonfiction formative assessment, Edward scored in the Basic range, at Level 3, Stage B. His initial reflection
was *Modest*, shifting to *Robust* through Level 3, *Accepting* at Level 4, and back to *Modest* overall at the conclusion of the assessment: “I think that this assessment was easy because I was easier to find the answers to the questions. I think this is a fair enough grade.” An example of a Level 3 response with a *Robust* rating showed his confidence in his accurate understanding: [Who were the two groups of people that President Roosevelt helped because his wife influenced him to do so?] “Minorities and women.” At Level 4, the ratings that could be either *Accepting* or *Antagonistic* were rated *Accepting* as noted in an observational memo: “This is based on knowledge of Edward--he does not get angry, he accepts with minimal frustration when he does not do well.” An example from Level 4 shows a confidence rating of 3 (he was pretty sure he was correct) for a response that was not incorrect, but was missing much of the detail required for a complete score.

[Why was the way Eleanor chose to be First Lady so noteworthy? Model--Eleanor was the first First Lady to travel in an official capacity and to speak out for justice. She went all over the place talking to people. She would tell her husband, the President what people told her. She did many things that helped a great number of people and was very outspoken about it.] Because she traveled around the country talking to all kinds of Americans who were still not being helped by relief programs.

At this time, Edward’s metacognitive development did not support the understanding of how much information was needed at the inferential levels to demonstrate clear proficiency. His *Quality* of thinking was predominantly *OK*, with the most prevalent aspect of thinking he engaged in being answering questions. He chose to reflect on his thinking some of the time, but not always. He did not offer explanatory details for his answers, or ask questions of the text.

*Narrative nonfiction developmental work.* Continuing his growth from the Narrative unit, throughout the time devoted to class work during the Narrative Nonfiction
unit, Edward consistently produced Level 4 responses. Observations yield Personal Agency Beliefs of Robust, reflecting both his high level of engagement and his significant success at thinking and sharing his thinking at the inferential level. Observational field notes discuss his shift from one of the members of the quintain in frequent need of re-teaching in a smaller group, to a class leader who was able to work independently and offer assistance to others. The transformation was noticeable, and he commented to his peers during reading class, his parents (a meeting and an email communication with Edward’s mother regarding this are noted), and his teacher about his new competence and how he was enjoying it. He was engaged, enthusiastic, and ready to take on any new challenge daily. He worked collaboratively, spending equal time giving and receiving feedback. Edward was quick to go back into previous work for examples to support his thinking about the concept of complex inferential thinking.

During the coursework of the Narrative Nonfiction unit, Edward demonstrated predominantly Good thinking across the varied Aspects of thought, demonstrating the thinking of explaining his answers as well as providing direct support from the text most of the time. In addition, during discussion he engaged in the higher order thinking process of posing questions 13 times. An example of the breadth and depth of thinking comes from a Level 5 written response. The parts of this response were analyzed separately, but it is placed here in its entirety to show the full measure of his metacognition during this unit. It should be noted that this thinking was done collaboratively with Bella, a combination that yielded higher results for both of them, with neither dominating. From observation notes:

Answer- Nujood has changed a lot since the beginning. In the beginning I thought she was very shy. Because it seemed to me that she didn’t talk to [sic] much
people in the first few chapters. Then in the middle she got a lot more talkative by
talking to the judge and other in the court. In chapter 6 she really got to taking [sic],
because of the wedding and the divorce case when it was over. Near the end when it was the first day of school for Nujood I felt so excited for her. I was so excited I felt like I was in the story. I could just see how she was happy for her self [sic] when she got into the taxi. But it wasn’t a big deal for Haifa because she used [sic] to going.

*Quote*- “When I opened my eyes this morning, the first thing I felt was my heart beating excitedly. Then I tip toed off to brush my teeth and comb my hair.” *Explanation*- I think she will like the new school that’s she attending, because she going to meet a lot of new and friendlier people.

**Narrative nonfiction summative assessment.** On the Narrative Nonfiction summative assessment, Edward scored in the Proficient range, with a score of Level 5, Stage B, tied for the highest score in the quintain. This was a new experience for him that he enjoyed. Edward’s *Personal Agency Beliefs* were categorized as *Robust*. For the first time, he attempted to answer questions at both Level 6 and 7. An example from Level 5 shows his growth:

[Why might it be a good idea for President Obama to study the lessons of Roosevelt’s presidency? Be sure to use evidence from the essay to help you explain your opinion.] *Answer*- Obama should study the lessons of Roosevelt’s presidency because then he would know what to do if we have another great depression again when he’s president. “When the Great Depression came, Roosevelt set up the ‘TERA,’ to create jobs for people in his state”.

Edward’s overall reflection at the *Robust* level is clear: “I think that this assessment was easy because I know how to do it now from practice with all the responses. I just did it and now I’m very proud of myself.”

**Expository formative assessment.** On the Expository formative assessment, Edward scored in the Basic range, at Level 3, Stage B. He expressed negative *Personal Agency Beliefs* with evidence of a *Fragile* sense of agency throughout the assessment, not attempting to answer any questions after Level 4. His personal reflection in the *Vulnerable* category at the start of the assessment matches his work: “Before I take this
assessment, I am feeling confused because I don’t know how hard its going to be. I think I can do it.” After taking the assessment, Edward’s reflection was Fragile: “I think that this assessment was difficult because it didn’t make sense to me. It was hard to understand the question. I think it is good because I only stoped [sic] at Level 4.”

Consistent with Edward’s fall back to a Basic level of response, the only aspect of thought he demonstrated on this assessment was simple answers, with no explanation or support from the text. The quality of his responses was primarily OK at the concrete levels of comprehension, and poor (Weak and Wrong) when asked to demonstrate understanding at the inferential level. An example of Good work at Level 2 shows the best work on this assessment:

Because of the writing made by the summerians [sic] it allowed them to record what a way of writing on a clay tablet and could bring there ideas everywhere they go and not count on a messagers [sic] memory to remember.

By contrast, a Wrong example from Level 3 and a Weak example from Level 4 show the more prevalent level of responding. At Level 3: [How were kings of the Shang Dynasty buried?] “They would burn your hand and write on it.”. From Level 4: “They found thousands of exercises on the tablets, and found lesson also from the teacher. I think they had a harsh school, with all that work.”

*Expository developmental work.* During the Expository unit work, Edward worked at both Level 3 and Level 4, with Level 3 being the dominant one. Observations place him in the neutral range, vacillating from Tenacious on days when he was not feeling the pressure from other classes and working collaboratively, and Self-Doubting when he either shared that he was tired or that school was too much work and he was asked to work alone.
During the Expository unit, the quality of Edward’s thinking is mixed, with positive attributes in responses, and negative in think-alouds. Throughout the unit work, there were 18 negative Quality ratings, but all in think-alouds and at Level 1. The positive ratings were Good and Strong on Levels 3 and 4. His thinking was distributed across aspects, featuring mostly answering, with explanations about half of the time, and direct support from the text about one third of the time. Edward’s thinking included creating a title to conceptualize his literary response on three occasions. However, his titles were of a concrete rather than an inferential nature. For example, “The Low Point” for a response focused on identifying the stated low point in the depression from a text.

Expository summative assessment. On the Expository summative assessment, Edward scored again in the Basic range, but this time at the top of it with a score of Level 4, Stage A. His Personal Agency Beliefs were again categorized as Self-Doubting, both before and after the assessment. Before the assessment he shared: “Before I take this assessment, I am feeling tired and kinda focused.” After the assessment:

I think that this assessment was difficult because I couldn’t focus at all today I was frustrated with all the work up to my neck In all my other classes. I think I could do better because I didn’t try my best.

For this summative assessment, Edward’s thinking was generally in the OK range, with his responses dominated by simple answers. However, while he did not endeavor to explain or use much in the way of specific support of his thinking from the text, he did demonstrate an understanding of the flaws in his thinking when he reflected on the quality of his answers after reviewing the scoring guide, sharing that “I didn’t have much characteristics of the people.” and “I could have wrote more details and more things that are going on.”
Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

GMRT-4 and SRI. Edward’s performance at the Proficient level on the GMRT-4 is evidence that he closed the gap. He demonstrated a growth of 25 months or 10 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Edward did not demonstrate the same narrowing of the gap with his SRI score. His growth in his SRI score through the end of the year was 50 Lexile points, very close to the average of 51 for his counterparts in the quintain. He did not show any Lexile growth at the semester, placing his progress at the end of the study framework implementation well below the comparison to the other remediation group at the semester of 46.75 Lexiles on the SRI.

PARLA. Edward made gains in all three content areas, as measured by the PARLA proximal assessments. Specifically, on the PARLA-NARR he improved by 3 stages from formative to summative assessment, doubling the average achievement growth of the members of the quintain. On the PARLA-NARR NF, he improved by 6 stages, as compared to the average of just under 4 stages for the rest of the quintain. Finally, on the PARLA-EXPO he underperformed relative to his immediate peers, achieving only 2 stages of growth relative to the group’s average of near 3 stages.

The hierarchical nature of the proximal measures (PARLA-NARR, PARLA-NARR NF, PARLA-EXPO) along with the study design that incorporated a request for
reflection on these measures, provide a more detailed glimpse of Edward’s development than the SRI or the GMRT-4.

During the Narrative unit, Edward’s reflections were consistent with his personal statement to the researcher at the start of the study that he was never going to “get” this “reading thing.” He declined to comment on the pre-assessment for Narrative Nonfiction, but his classroom performance showed tremendous growth, particularly within the context of collaborative work.

Observations and artifacts matched, showing a learner working consistently at Level 4 of comprehension in the framework (basic inference) independently as well as in a leadership capacity within the context of collaborative work. In addition, when working in collaborative dyads and small groups, he performed consistently at Level 5, engaging enthusiastically in thinking at a complex inferential level and supporting this thinking with evidence from the text.

The Narrative Nonfiction summative assessment was the first time, in over 2 years of being Edward’s teacher, that he shared a sense of pride in an academic task accomplished. At this time, his other core teachers began to mention in passing (and in one instance through an email for this purpose) that they were seeing a positive change in his engagement, confidence, and performance in their classes as well, particularly Communication Arts and U.S. History. His mother also corresponded with the researcher through email and indicated that his language about school and himself as a student was changing to positive conversations for the first time since he started middle school.
Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Edward’s self-reporting about his motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASC) did not change, staying in the Modest category throughout (see Table 1. The MST Taxonomy of Personal Agency Beliefs). However, his regular reflections and observations made during discussions showed marked positive change. The most prevalent categories for him throughout the work across the three content areas were: Robust, Tenacious, Modest, Accepting, and Discouraged, in that order, as shown in the pie graph of Figure 12.

![Pie chart showing distribution of most prevalent PABs in Case 3.](image)

**Figure 12.** Distribution of most prevalent PABs in Case 3.

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied: The Robust rankings moved from 31% during Narrative, up to a high of 56% during Narrative Nonfiction, and back to 33% during Expository. The Personal Agency Belief of Accepting was at the greatest level of 47% at the start of the study, dropped during Narrative down to 16%, and then increased to 21% in the final unit. The Modest belief was least susceptible to
change across contents, as shown in the graph in Figure 13. The dominant *Robust* 
*Personal Agency Belief* category featured 22 total occurrences, with 15 being at the literal 
levels of comprehension. When moving into the inferential level of understanding what is 
read, there were seven total occurrences of *Robust*, beginning with the Narrative 
Nonfiction unit. *Tenacious* was next, with seven total items at the literal levels. Two 
instances of *Modest* belief ratings occurred during Narrative Nonfiction. *Accepting* was 
represented with 17 total, 13 during the first unit, and two each for the following two 
units.

*Figure 13. Case 3: Tracking the most prevalent PABs across the study. This figure 
illustrates the variability of PABs throughout the three units of the study.*
Throughout the study, there was one rating at *Accepting* for comprehension at the inferential level during Narrative Nonfiction. Six occurrences of *Discouraged* at the literal levels of comprehension were split evenly between Narrative and Expository unit. Finally, there was one response during Narrative Nonfiction that was at an advanced level of comprehension that was classified as *Discouraged*.

**Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?**

*MARS*I. On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Edward’s self-reported changes in metacognition from the MARSI were greatest in the area of Global Reading Strategies, with a change from a rating of 2.8 at the start of the study to 3.5 at its conclusion. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. The subscale of Problem-Solving Reading Strategies includes slowing down to read carefully, pausing and reflecting, and similar strategies; Edward reported a slight decrease from 3.8 to 3.5 over the study period. Finally, Support Reading Strategies encompass a variety of actions that include note taking, paraphrasing, and discussing the material, among others. Edward’s self-rating improved from 3.1 to 3.8 for these strategies.

Considering the places on the data matrix where assessments meet the MARSI, the growth in self-reporting of use of reading strategies support Edward’s improved reading scores. His reflections about the MARSI are shown in Table 14.
Table 14. 
Reflections on MARS survey by reading strategy for Case 3.

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I’m normally good at predicting, and I’m good at making mental pictures when I’m concentrating on the book or text.”</td>
<td>No comment</td>
<td>“I don’t like or good at taking notes on what I read.”</td>
</tr>
</tbody>
</table>

Think-aloud protocols. The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, the vast majority of the moves in Edward’s think-alouds were at Level 1, as shown in Table 15. Accordingly, this data does not illuminate Edward’s metacognitive growth during the implementation of the PARLI framework. Before each think-aloud protocol, he expressed a desire to skip it. On one occasion he successfully negotiated a reprieve to complete it during the following week when he shared: “I’m just so wiped out, do I have to do it now?”

Table 15. 
Frequencies of Levels of think-aloud protocols for Case 3.

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.76</td>
<td>.03</td>
<td>0</td>
<td>.21</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.43</td>
<td>0</td>
<td>0</td>
<td>.57</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>.82</td>
<td>.09</td>
<td>0</td>
<td>.09</td>
<td>0</td>
</tr>
</tbody>
</table>
PARLA/PARLI. During the Narrative segment of the PARLI framework, the levels of responses were almost identical across Level 1 through Level 4 in Table 16. The Quality of response was predominantly OK or Weak (Table 17.).

Table 16.
Case 3: Frequencies of Levels present in work by unit.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>102</td>
<td>.20</td>
<td>.20</td>
<td>.22</td>
<td>.23</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>115</td>
<td>.21</td>
<td>.23</td>
<td>.18</td>
<td>.26</td>
<td>.08</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>Expository</td>
<td>93</td>
<td>.38</td>
<td>.21</td>
<td>.25</td>
<td>.15</td>
<td>.01</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were almost fairly well distributed across Level 1 through Level 4, with 8% at Level 5, and 2% each at Level 6 and Level 7 as shown below in Table 15. The Quality of response was at the midpoint or above (OK and up) 72% of the time (Table 17.).

Table 17.
Case 3: Frequencies of Quality present in work by unit.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>86</td>
<td>.01</td>
<td>.14</td>
<td>.37</td>
<td>.36</td>
<td>.10</td>
<td>.02</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>91</td>
<td>.04</td>
<td>.15</td>
<td>.53</td>
<td>.21</td>
<td>.04</td>
<td>.03</td>
</tr>
<tr>
<td>Expository</td>
<td>74</td>
<td>.11</td>
<td>.18</td>
<td>.32</td>
<td>.35</td>
<td>.01</td>
<td>.03</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of 84% of responses were at the concrete levels of comprehension (Table 16), while the Quality of
responses were split 29% positive (Strong and Good), 32% in the mid-range, or OK, and 39% in the lower tiers of Quality (Table 17).

Sub-question. 4) Does student performance on the assessments form a pattern of development?

Of the quantitative measures, Edward’s GMRT-4 scores are the only scores that have improved consistently, albeit not greatly. The others fluctuate, but not at a consistent rate. The variable pattern fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process. Edward’s pattern shows improvement following intervention. This was most dramatic when the instructional intervention was over multiple weeks and included reading, responding, and discussion of a full-length text during the Narrative Nonfiction unit.

Summary. Edward performed at the proficient level on the GMRT-4 and on one of the proximal measures (PARLA-Narrative Nonfiction, formative assessment), while being on the upper end of Basic for both the SRI and the Expository PARLA. His formative and summative assessments of agency and motivation remained in the second category of Modest on both, and observational field notes, particularly during the Narrative Nonfiction unit, show many instances of Robust. His metacognitive scores on the MARSI survey changed from the mid-range to the high usage of three of the four categories of reading strategies (Global and Support), bringing all three to a high level of self-reported usage. This shift was also seen in the regular work in greatest measure during the Narrative Nonfiction unit as he utilized a greater breadth of thinking to articulate his understanding of what he read. An evaluation of the evidence of Edward’s development regarding reading comprehension, agency and motivation, and
metacognition across the course of the study makes a case for the PARLI framework being effective with Edward.

**Case study: Dominique.**

**Previous years of schooling.** Dominique (Case 4) was on an Individualized Education Plan (IEP) for speech in elementary school and was exited from that program at the start of sixth grade. She participated in Reading Recovery in third grade, there was no reading intervention in fourth grade and she ended the year at the bottom of the grade level range. In fifth grade she was placed in reading tutoring, but did not progress beyond the fourth grade level. Based on her GMRT-4 reading score at the end of that year, she was placed in a reading intervention that focused on decoding issues for sixth grade.

Dominique has performed at a Basic level on the MAP test in Communication Arts for the past 3 years, with very little change. Dominique’s SRI score was at the middle Basic/Partially Has Met Standard range at the end of last year, but she moved into the start of the Proficient/Has Met Standard Lexile range at the start of this year. Dominique has had consistent increases in Comprehension scores on the GMRT-4 in the last 3 years, but has not yet experienced a full year’s growth; she went from 4.2 at the end of fifth grade, to 4.9 at the end of sixth, and 5.6 in seventh. Dominique started sixth grade with decoding issues, and made small progress there, and was moved up to reading remediation at the Read180 program level for seventh grade largely because her father did not like her being in the decoding class and was vocal in his disagreement about placement, adding to her discomfort about what this placement said about her.

Dominique was resistant to the computer, citing previous negative experience with the program, and did her best to avoid it. She seemed to enjoy reading fiction in
class but resisted reading at home. According to school district records, her family dynamic was highly dysfunctional with a bipolar father. When he was off of his medication, he became very authoritarian and used reading as a punishment, which probably contributed to Dominique’s choices. Dominique embraced participation in the research from the outset and started out very engaged in class work.

**Research Questions.**

Research Question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Dominique. Appendix G features the Event Flow diagram for Dominique that provides a graphic representation for the data.

**Narrative formative assessment.** On the Narrative formative assessment, Dominique scored in the Basic range, between Level 3, Stage C and Level 4, Stage A. She demonstrated Personal Agency Beliefs in the Robust category through Levels 1 and 2, Modest at Levels 3 and 4, and Discouraged at Level 5. Metacognitively, Dominique displayed a mix of quality of thought and articulation of her understanding across concrete and inferential levels of comprehension, but the majority were Weak and Wrong. Her reflections about her thinking process are shallow, showing no insight into her process. An example from Level 3 is illustrative and common throughout the assessment: “The rubric was more detailed.”

**Narrative developmental work.** During think-alouds in the Narrative unit, Dominique rarely commented on the text, and had many miscues and substitutions. In
addition, her reflections at the end of the process were predominantly at Level 2, making a connection between her personally and the key details of the piece. She also made a number of relatively random comments, suggesting a bit of distractibility: “I want a cookie now. They mentioned cookies. I want a cookie.” Her *Personal Agency Beliefs* throughout the work were neutral and variable, moving from *Tenacious* to *Vulnerable* and *Self-Doubting*. During the unit overall Dominique worked at Levels 3 and 4 the most, with Level 1 being the next most frequent. However, Level 4 was almost exclusively achieved with annotation, rather than responses or reflections. An example of this occurred when she highlighted the sentence in the text that describes the neighborhood and who lives there and made a note “art=important” in reference to a mention of artists. She had three times as many *Weak Quality* ratings as those that are positive (*Good*, *Strong*, *OK*). Overall, most of the positive evidence of thinking was in the area of annotating text, while most of the lowest quality of thinking was in the think-aloud protocols, where she mostly read with a great deal of errors.

**Narrative summative assessment.** On the Narrative summative assessment, Dominique scored in the Proficient range with a score of Level 5, Stage A. Her overall *PAB* reflections went from *Vulnerable* to *Robust*. An example from Level 5 shows her level of work, along with her residual decoding issues:

[There seem to be many ways in which Mrs. Jones influences Roger. Name at least three things that Mrs. Jones does that influence Roger. Which one is the most important?] Mrs. Jones influences Roger by taking him home and see how nice she is, giving him 10 dollars to buy what he wants, and making him see what he did wroung [sic]. Making him see what he did wroung [sic] is the important one because he wouldn’t do it any more [sic] and know that he can just ask.

Dominique’s *Vulnerable* reflection prior to the assessment displays the contradictions inherent in her limited metacognitive development: “Before I take this
assessment, I am feeling scared but not because of the assessment and I think I’ll do fine.” Her Robust reflection afterward further reinforces these contradictions: “I think that this assessment was easy because at Level 5 it got hard like I knew the answer but didn’t know how to write it.”

On this assessment, her responses reflect inferential thinking, but her reflections are concrete and do not provide evidence of metacognition. For example: “I just left out one minor detail.”

Narrative nonfiction formative assessment. On the Narrative Nonfiction formative assessment, Dominique scored in the Below Basic range, at Level 2, Stage A. Consistent with her pattern of alternating engagement and disengagement, an example response suggests that perhaps she was not as attentive to the work as she had been previously:

[Specifically, what was the Great Depression? Model--The Great Depression was a severe economic crisis during the 1930s which had millions of people losing their jobs, their houses, and their savings.] Depression is when you fell down and you don’t know if you can do it or not.

Dominique’s thinking was predominantly of poor Quality, with a split between Weak and Wrong. Accordingly, she did not demonstrate a strong sense of agency on this assessment, with mixed PAB ratings throughout concrete level, with equal instances from the top half of the scale (split between Modest and Fragile) and the bottom half of the scale (split between Accepting and Discouraged). Reflections during the assessment reflect this. At Level 2 she said: “I didn’t know what it was and where to find where it was in the essay.” At Level 3, on a question that most in the quintain answered correctly, and whose answer was stated directly in the text: “I don’t know were [sic]to find this at.” At the end of the assessment, she shared her frustration again: “I think that this
assessment was difficult because I couldn’t find out where the answers were and its
different than fiction.”

Narrative nonfiction developmental work. Unlike most of her peers, the work
Dominique did that was not part of a discussion was literal and reflected poor quality of
thinking. She produced more than two times more responses at the literal level than the
inferential. Of the literal, only one instance of Good quality thinking was shown:

The main story moving was that’s no one was listening to her. She was only a
little shorter then [sic] the people wastes [sic]. Then no one could hear her or see
her. Finally a girl saw her and though she was lost. All she wanted is to see a
judge. Then the girl toke [sic] her in to see a judge. I her [sic] and the judge are
connected because with out [sic] him she couldn’t get a divorce.

At the inferential level, only 20% were at the OK level (the best performance).
One example is a question asked as part of a response at Stage C: “Why does Nujood
want to go to the court house and how do you know?”

Think-aloud moves were mostly Level 1, and those were most often Weak
miscues that are not acknowledged or corrected in any way, providing evidence of
persistent problems at the developmental level of phonemic awareness. There was only
one instance of Level 3: “And also, I wonder, is it like, if you drink the salt water, he kind
of gets sick of it, it was salt mixed with oil, that’s disgusting.” She did make some
inferential moves as well. Her Personal Agency Beliefs were split between Robust and
Modest in both artifacts and observations.

Dominique’s work during discussion was more inferential and of better quality.
Of the inferential contributions that represented two-thirds of her discussion points, most
are questions reflecting greater depth of thinking: “Wouldn’t it be dishonor to them
because they were lying half the time about everything?” Her level of engagement was at
its highest during discussion, reflecting a Robust sense of agency, as she had twice as many contributions as the rest of her quintain.

**Narrative nonfiction summative assessment.** On the Narrative Nonfiction summative assessment, Dominique scored in the Proficient range, improving 2 Levels to a score of Level 4, Stage B. While she did not perform consistently at Level 4, she had some success there and at Level 5 that resulted in her score in the middle of Level 4. An example of an incorrect response at Level 4 nonetheless shows an attempt to use evidence from the text to support her thinking:

[What was Roosevelt’s greatest obstacle to a career in politics? Model--Roosevelt’s obstacle was polio because he was in a wheelchair and voters would consider him weak.] That he was a seniter [sic] of New York, he loved politics, he quiet [sic] the line “let me assert my firm belief that the only thing we have to fear is fear itself.” And the most important one was presentdent [sic] because that was his Dream.

A partially correct answer from Level 5, however, stopped short of effective use of details from the text to support the inference made:

[Why did people have so much faith in FDR? Be sure to use evidence from the essay to help you explain your opinion. Model--he was positive, hard working, and he got people back to work.] I think so many people had faith in FDR because they needed some one to lead them. So they wont fell [sic] lost. Also they would have faith in FDR because I think he loved his country and it says that he wanted to end CCC, CWA, WPA, and Great Depression.

Dominique’s Personal Agency Beliefs correlated with this at Tenacious. Her poor performance with the work of the unit overall made this result surprising. Her reflection at the start suggested that focus might be a problem: “Before I take this assessment, I am feeling sleepy but I need to focus on this I think its going to be easyer [sic] because we know all the levels.” However, her final reflection gave credit to the work done in the unit:
I think that this assessment was easy because until level 6 and 7 because I haven’t really done those levels but it was a lot easier then last time. I went up I’m happy that I’m understanding more.

Metacognitively, Dominique’s responses at the concrete level were predominantly of Good quality. At the inferential level, her responses were of higher quality than her reflections about them, showing that she is not yet stable with her metacognitive processing.

*Expository formative assessment.* On the Expository formative assessment, Dominique scored in the Below Basic range, at Level 2, Stage B. She expressed a Tenacious sense of personal agency at the start: “Before I take this assessment, I am feeling Good, But very sleepy.” When she reflected after seeing her results, she was Discouraged: “I think that this assessment was difficult because I got confused on a lot of things and mixed up.” At Levels 1 and 2, she accurately assessed the correctness of her answers, but at Level 3 she lost this accuracy, showing a lack of effective metacognition at a concrete level of thinking and beyond. An example of her best work on this assessment is found at Level 2: [How were kings of the Shang Dynasty buried?] “They buried there kings [sic] with royalty things, everyone in a King’s court, and hundreds of slaves that served him.”

Dominique attempted responding through Level 6, but the attempt was half-hearted, with single sentence responses where paragraphs were called for. The quality of the responses was Weak or Wrong. The following example received .5 out of 2 points possible:

[Use what you read about the class systems of Mesopotamia and the Shang Dynasty class system to explain why making sure all students read and write well is so important in the United States today. Model--The higher classes were taught to read and write. Everyone must know how to read and write now to give
everyone a chance to rise. Those who can read and write have a say in how things work. In our democratic system, the stronger the education of the people, the better the system works for everyone in the society.) it [sic] will be important today because without it we wouldn’t have a way to commicat [sic], or how to write stuff that’s important in time.

*Expository developmental work.* During this unit, Dominique’s think-aloud moves were Level 1 the majority of the time, and were consistent with earlier instances of *Weak* Level 1 moves that were miscues that were not noticed or corrected. However, she did also make some inferential moves that showed some deeper thought: “I’m guessing that it’s like they’re fighting for something they believe in.”

Dominique’s collaborative practice was mostly at literal levels, and attempts at inferential moves were unsuccessful. She did not seem to benefit from collaboration with her primary partner. It is noted that her partner was the one case who did not make marked progress during the study (Rafael). An example of one of their Level 4 attempts shows this weakness with a response that focuses on a piece of text that does not require inference, but is labeled at the inferential level: “Each time a boss would take away a cowboys favorite horse, it would send a sign saying that they wanted them to leave. Because it might mean that the cowboy did something wrong that that boss didn’t like.”

Most of the work was of poor quality at the concrete levels as well, and each of the few inferential attempts was *Weak*, as in the previous example.

Observation field notes show a mix of *Vulnerable* and *Self-Doubting*, based on the partner(s) with whom she collaborated. When working with Alice, Dominique was in the *Vulnerable* range, functioning well some of the time, and waiting for the instructor some of the time. There was one instance of high engagement at the initial exploration of expository text structures. When it came to the work with Rafael as her collaborator, she
down-shifted to *Self-Doubting* and field notes show a constant need for redirection and work stopping unless the instructor was in close physical proximity.

*Expository summative assessment.* On the Expository summative assessment, Dominique scored in the Basic range, going up to a score of Level 3, Stage C. Her *Personal Agency Beliefs* were initially *Robust*: “Before I take this assessment, I am feeling Good [sic], But[sic] very sleepy.” An example of her strongest work on this assessment is at Level 3:

[How did the mountain men contribute to the settlement of the Far West? Model--By acting as guides for wagon trains (exploring the land).] The mountain men knew the land better than anyone else so he would guide the travelers threw the trails.

At Level 5, while responses were incomplete, there was the first attempt to include direct support from the text (Dominique included the support but left off her thinking):

[How were the missionaries and the mountain men similar? Use examples from throughout the essay to support your opinion.] The missionaries would travel on the Oregon trail just like the mountain man and would teach people new stuff. It says “missionaries traveled to Oregon country to teach the Christian religion to Native American.” Then the mountain men “The mountain men knew the Oregon Country better than anyone and needed work. So they began to lead settlers across the rugged country in big groups called wagon trains”.

Dominique’s quality of thinking is split between *Weak* and *Good* quality at the literal levels for both responses and reflections. At the inferential levels, there were two instances of Dominique supporting her thinking with the text and reflecting on her thinking that showed a *Good* quality of thought. Both of these show greater depth of metacognition than she had demonstrated previously. An example of this deeper metacognition is apparent in her reflections, which also included the evidence of her residual phonemic awareness difficulties: “At the time I didn’t really understand what
they was [sic] asking. And it was different by the way they was explaining [sic] stuff from
the knolage [sic].”

At the assessment’s conclusion, Dominique was Discouraged overall: “I think
that this assessment was difficult because I got confused on a lot of things and mixed up.”

Sub-question. 1) Does the PARLI framework contribute to closing both the fiction
and nonfiction reading comprehension gap between struggling readers and non-
struggling, grade level readers?

GMRT-4 and SRI. Dominique’s performance at the Basic level on the GMRT-4
did not provide evidence of narrowing the gap. She demonstrated a growth of 18 months
or a rise of 4 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5
NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for
struggling eighth grade readers receiving remediation in the other four middle schools of
the test district. On the other hand, Dominique’s Proficient performance on the SRI does
represent a narrowing of the gap relative to her non-struggling grade-level peers. Her
increase in her SRI score through the end of the year was 22 Lexile points, and placed her
in the proficient range on this measure.

Interestingly, Dominique expresses doubt in her competence each time the SRI is
taken, reminding all of us, as she did on the initial occasion, that “I’m not good with
reading on the computer,” then in the next breath talking about her new laptop, the desire
for an iPhone®, and her enthusiastic use of social media. There is a possibility that her
lower performance on the SRI is connected to her assertion, perhaps a self-fulfilling
prophecy.
EVALUATING THE PARLI FRAMEWORK

PARLA. Dominique made gains in all three content areas, as measured by the PARLA proximal assessments, providing evidence of closing the gap when she reached grade level expectations in both Narrative and Narrative Nonfiction. When it came to Expository, she did not close the gap, remaining 2 stages below expected grade level. Specifically, on the PARLA-NARR she improved by 1 complete level (4 stages) from pre to post assessment, placing her above the average growth of the members of the quintain. On the PARLA-NARR NF, she improved by 7 stages, as compared to the average of just under 4 stages for the rest of the quintain. On Narrative her final performance was at Level 5, Stage A, and with Narrative Nonfiction her final performance was at Level 4, Stage A, just beginning to understand basic inferences. These results are stronger than most of her written class work. Finally, on the PARLA-EXPO she showed growth relative to her immediate peers, achieving 4 stages of growth relative to the average of near 3 stages, but still ending just below grade level, at Level 3, Stage C.

Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Dominique’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Tenacious to the Vulnerable category (see Table 2. The MST Taxonomy of Personal Agency Beliefs). The most prevalent categories for her throughout the work across the three content areas were: Modest, Robust, and Accepting, in that order, as shown in the pie graph of Figure 14.
Dominique’s residual issues also seem to be playing out in her PAB ratings, with her rating being 2 to 3 levels below what is seen in artifacts and observation. Her initial reflection in scoring herself as Modest was “because my pertenely [sic] best and I’m all ways out going” shows her spelling/decoding/fluency issues along with her reliance on social interaction as an academic tool.

The MAASCM scales do not directly refer to social interactions per se, but Dominique interpreted them as personality descriptors. Interestingly, her pairing of higher PABs with Good and Strong Quality ratings reflects an accurate self-assessment of her performance in practice, despite her lower self ratings in isolation. Her shift in self reporting from Modest to Vulnerable may be explained at the intersection of MAASCM and reading assessments, as this drop fits with her growth of about 6 months over the academic year as measured by the GMRT-4 and SRI.

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied: The Robust rankings moved from 29% during Narrative, up to a high of 46% during Narrative Nonfiction, and
down to 10% during Expository. The *Personal Agency Belief of Accepting* was at 38% at the start of the study, climbed to the greatest level of 55% during Narrative, and dropped down to 10% in the final unit. The *Modest Belief* was the least volatile across contents, varying by no more than 6% from the highest to the lowest, as shown in the graph in Figure 15.

*Figure 15*. Case 4-tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

For Dominique, *Robust* was the second most populated category, with 19 total instances over the time of the study, well distributed across all three contents. *Modest* is the most frequent *PAB* rating for Dominique across the study, with 17 instances at the literal levels of understanding, split across the three genres. In Narrative, she had one response at the advanced levels of comprehension that was categorized as *Modest* in *Personal Agency Beliefs*. When it came to the mid-point in the taxonomy, Dominique had five total *Accepting* ratings at the literal levels of comprehension and 10 for the inferential level of understanding.
Dominique complained about the request to reflect on the assessments, sometimes in ways that did not track, as she did preceding the Narrative summative assessment, when she shares that she is scared, then counters it with an assertion that she will be fine. She does not elaborate about what she is scared about, suggesting that she may be taking what she is feeling and interpreting it as disconnected from being tested in an area of relative weakness. After this assessment, on which she scored in the Proficient range (Level 5, Stage B), she once again provided a reflection that lacks clarity for the reader when she says it was easy but she did not know how to write all of the answers. When it came to Narrative Nonfiction, her reflections about the assessments varied from a straightforward conclusion that the first test was hard to a new reflection of sleepiness prior to each test that made its first appearance here. The class took place immediately following lunch every day, which may have contributed to her drowsiness.

Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?

MARI. On the Metacognitive Awareness of Reading Strategies Inventory (MARI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Dominique’s self-reporting indicated a mix of ratings at all 3 levels for both pre and post measures. She reported virtually no changes in metacognition from the MARI. In the area of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting, and similar strategies. Dominique reported a decrease from 3.8 to 3.6 over the study period. For the subscale of Global Reading Strategies, she remained at 3.0, in the middle of the ratings. Global Reading Strategies include strategies pertaining to setting a
purpose for reading, activating prior knowledge, making predictions, and so forth.

Similarly, Dominique’s lowest self-rating of 2.4 for Support Reading Strategies, which include a variety of actions such as note taking, paraphrasing, and discussing the material, among others, did not change.

Finally, where Dominique’s responses on the MARSI and artifacts meet, the data is mixed. Her MARSI self ratings are spread from high with Problem-Solving Reading Strategies, to medium with Global ones, and low with Support Reading Strategies. She shares: “I know I should use these more and I now [sic] I got an [sic] low” in reference to her use, or lack thereof, of Support Reading Strategies.

Dominique’s artifacts were spread across Level and Quality measures, with her Narrative work demonstrating more high Level and high Quality work, leading to a grade level performance on the post test. When moving to Nonfiction, the utilization of Support Reading Strategies results in greater growth, but she avoided this and worked more at the lower levels, with mixed results. Greater levels of new vocabulary in nonfiction realms may have represented a stumbling point, given her phonemic awareness issues.

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, the vast majority of the moves in Dominique’s think-alouds were at concrete levels as shown in Table 18. Overall, Dominique made 77% of her moves at Level 1.
Table 18.  
*Frequencies of Levels of think-aloud protocols for Case 4.*

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.85</td>
<td>.09</td>
<td>.06</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nonfiction</td>
<td>.68</td>
<td>.14</td>
<td>.02</td>
<td>.14</td>
<td>.02</td>
</tr>
<tr>
<td>Expository</td>
<td>.76</td>
<td>.07</td>
<td>.17</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**PARLA/PARLI.** During the Narrative segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 7, as shown in Table 19. The *Quality* of response was predominantly *OK*, with similar frequencies for *Good* and *Weak* as shown in Table 20.

Table 19.  
*Case 4: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>122</td>
<td>.32</td>
<td>.13</td>
<td>.30</td>
<td>.20</td>
<td>.05</td>
</tr>
<tr>
<td>Nonfiction</td>
<td>107</td>
<td>.34</td>
<td>.37</td>
<td>.08</td>
<td>.21</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>80</td>
<td>.35</td>
<td>.30</td>
<td>.21</td>
<td>.14</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through 7, with the majority being at Level 1 and Level 2, together combining for 61% of the responses. Level 3 and Level 4 combined for 30% of the responses, with the remaining 9% split among the other levels, as shown in Table 19. The *Quality* of response was at the midpoint or below (*OK* and down) 95% of the time, with the remaining 5% at *Good* (Table 20).
Table 20. Case 4: Frequencies of Quality present in work by unit.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>107</td>
<td>.01</td>
<td>.08</td>
<td>.47</td>
<td>.36</td>
<td>.02</td>
<td>.06</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>94</td>
<td>0</td>
<td>.09</td>
<td>.38</td>
<td>.33</td>
<td>.06</td>
<td>.14</td>
</tr>
<tr>
<td>Expository</td>
<td>65</td>
<td>0</td>
<td>.12</td>
<td>.42</td>
<td>.36</td>
<td>.05</td>
<td>.05</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of 72% of responses were at the literal levels of comprehension (Level 1-Level 3), with simple inference (Level 4) at 18%, and complex inference (Level 5) at 9%, as shown in Table 19. While response were attempted up to Level 6, the Quality of responses was rated as Weak or lower 63% of the time, OK 2%, and Good 17% of the time.

Sub-question. 4) Does student performance on the assessments form a pattern of development? Dominique’s scores have been consistent on quantitative measures, but her PARLA scores show variability, in a variable pattern that fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills showing as fluctuating. Dominique’s pattern shows improvement on assessments following instructional intervention with no correlation between the Quality of the work done during instruction and performance on the assessment.

Summary. Dominique performed at the proficient level on the SRI and on two of the proximal measures (PARLA-Narrative and Narrative Nonfiction post-assessments). Her performances on the GMRT-4 and the third PARLA (Expository) were Basic. Dominique’s formative and summative assessments of agency and motivation diminished from a starting place at Modest to an end point near the middle of the scale of Vulnerable. Her metacognitive scores on the MRSI survey did not change. In spite of her limited
engagement in the process on most days, Dominique progressed as a reader in substantive ways, and began to analyze the comprehension tasks from other core courses in order to plan her responses.

While her self rating overall was Vulnerable, this contradicts her confident behavior during instructional time. Demonstrating a Robust sense of agency in daily practice, she consistently took a leadership role in collaborative learning tasks, speaking twice as often during discussions as the next closest member of the quintain. Observational field notes show her pattern of seeking clarification from other peer groups and taking it back to share with her collaborative group in an effort to accomplish a task in a timely manner. However, this was offset by her propensity to get off task and engage in social interaction with peers at inappropriate times that required multiple redirections for the majority of the time during which observational field notes were taken. Not surprisingly, given her developmentally appropriate focus on socializing, she demonstrated her greatest growth in comprehension at the inferential level during discussions, making inferential connections twice as often as literal ones. Despite evidence of her residual phonetic weaknesses, by considering evidence of her development regarding reading comprehension and metacognition across the course of the study, one can make a case for the PARLI framework being moderately effective with Dominique.

Case study: Suzie

Previous years of schooling. Suzie (Case 5) arrived in the district as a non-native speaker at the start of the second grade. She received English Language Learner (ELL) services for her entire elementary experience. In fifth grade, she failed to progress beyond
the middle of the fourth grade level, and was recommended for reading remediation at the level of the Read 180 program in middle school.

Suzie was a student who had been with the researcher since the sixth grade. When she first arrived, she was enrolled in ELL classes and was striving to perfect her English. According to district records, she had often been called upon to handle family matters beyond her years because of her language proficiency relative to her mother. She enjoyed doing well and enjoyed doing better and having a higher rank than as many classmates as possible. It is the researcher’s opinion that she is confident in her intellect and she would politely suggest that others were inferior regarding academic tasks. Suzie exhibits many behaviors of a classic Type A personality. Suzie seemed to always be in a race to read as many books as quickly as possible. She was consistently very vocal about completing another book and announced this in the midst of her classmates, then paused for recognition. She picked similar, realistic fiction most of the time and would quickly abandon a book if she was not immediately comfortable with the story or style of the author. She liked to complete any challenge or task first. She was in Read180 in sixth and seventh grades, and made significant progress, but remained below grade level. She was exited from ELL services at the end of the seventh grade academic year, as she had met the criteria.

According to district records, Suzie has performed at a Basic level on the MAP test in Communication Arts for the past 3 years, with some minor progress. Suzie’s SRI score was in the Proficient/Has Met Standard Lexile range at the end of last year, and she maintained this performance at the Proficient/Has Met Standard Lexile range at the start of this year. Suzie has had inconsistent GMRT-4 Comprehension scores in the last 3
years, but her GMRT-4 Vocabulary performance has steadily improved, resulting in GMRT-4 Total increases over time. Suzie’s GMRT-4 Comprehension score spiked in sixth grade at 6.3 (4.4 in fifth grade and 4.2 in seventh). Her GMRT-4 Total score has improved more than a full year’s growth each year; she went from 4.4 at the end of fifth grade, to 6.2 at the end of sixth, and 7.4 in seventh grade.

**Research questions.**

*Research question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?*

The PARLI framework was effective for Suzie. Appendix G features the Event Flow diagram for Suzie that provides a graphic representation for the data.

*Narrative formative assessment. On the Narrative formative assessment, Suzie scored in the Basic range, at Level 4, Stage A. She demonstrated *Personal Agency Beliefs* in the *Robust* category for Levels 1 through 5, but did not demonstrate breadth of thinking, providing simple answers with no explanations or detailed support from the text as shown in an example from Level 4: [Why does Mr. Tillian bring Rocky to the shop?] “So Rocky can keep him company” (5:Q1:19).

Suzie’s attempts at Levels 5 and 6 were unsuccessful. Metacognitively, her quality of thought and articulation of her understanding at the concrete levels of comprehension were all positive. At Level 1 it was *Strong/OK*, at Level 2 it was *Good/Strong*, and Level 3 it was *Good*. At the inferential levels, Level 4 was *OK*, and Level 5 was *Weak*. 
Narrative developmental work. During think-alouds in the Narrative unit, Suzie made few moves, and the moves she did make were predominantly at Level 1, with miscues that were not acknowledged. She made some moves to make connections at Levels 2 through 4, but they were more incidental than connected to metacognitive processes.

From observational field notes, Suzie exhibited signs of Vulnerable PAB as she resisted the use of the Student Guide, checked in with the instructor every time she had written a sentence or two, and needed to be re-taught in a smaller group. She required several more instances of re-teaching before she was able to perform well and independently.

While working through the Narrative unit Suzie worked exclusively at Levels 3 and 4, with Level 3 being dominant. This Level 4 response illustrates the breadth of thinking:

Title--The girl. Answer--I think that she did because she wanted to controll [sic] it. She also seemed really upset that she did that. Quote--“It must be terrible for you.” Explanation-- I think that this realtes [sic] to me because I can’t stop an asthma attack from happening to me.

Suzie engaged in heavy annotation, using the colors designated for Level 3 and Level 4 to highlight text and symbols to indicate her thinking. In addition, at Level 4 she added notes as she was trying to make sense.

Narrative summative assessment. On the Narrative summative assessment, Suzie scored again in the Basic range with a score of Level 4, Stage A. She also attempted higher levels during this assessment, with limited effectiveness: “Before I take this assessment, I am feeling I am a little nervous but I think I will get a good grade.” Her PAB was predominantly Tenacious throughout, but she did not reflect at the end of the
assessment. An example from Level 4 shows the lack of breadth of thought in still correct responses:

[Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story? Model--He is not really a bad kid, and recognizes that she has given him a chance when he bit off more than he could chew by trying to rob her. He is grateful to her for taking him in, talking to him in a respectful way, and giving him money to get what he wants. It would totally disrespect her to turn around and steal from her. If Roger did this, he would be a bad person.] He doesn’t take Mrs. Jones purse because she helped him.

Suzie’s quality of thought was variable with more of it on the positive side of the continuum. At Levels 1 and 2, *OK* is the predominant quality of thought. At Level 3, *Good* and *Wrong* are equal. At Level 4, thinking was *Good*, and at Level 5 it was *Weak*.

*Narrative nonfiction formative assessment.* On the Narrative Nonfiction formative assessment, Suzie scored in the Basic range, at Level 3, Stage C. She started off with evidence of a *Tenacious* sense of agency: “Before I take this assessment, I am feeling I think I will do fine on it.” This Level 3 example shows the lack of breadth of thought with accurate answers:

[In an essay written when she was 14, why did Eleanor say that it is easier to have no ambition? Model--She said it was easier to not have ambition because you won’t have to face difficulty of disappointment.] She said it was easier to have no ambition cause [sic] most people just face disappointment.

Breadth of thought was absent in the Narrative nonfiction formative assessment. Thinking is predominantly of mixed, but positive quality at Levels 1 and 2, and variable at Level 3 from *Weak* to *OK* to *Good*. With inferential thinking, the quality drops to predominantly *Weak*, with one instance of *OK* at Level 4.

An example of a *Weak* response at Level 4, along with the written reflection showing metacognitive awareness, highlights Suzie’s progress at this time:
[How does Eleanor’s essay on ambition predict the way she would spend the rest of her life? Model--The opinions expressed in her essay on ambition showed that she thought it was not acceptable to not try to do good work and make a difference. She talked about how important that was, then she did a great deal of work to help a lot of people. She is remembered for many things, including many things that she was the first woman to do.] That she would try to overcome her fears and she would help people overcome their fears.

Suzie’s reflection on this response shows her growing awareness of the power of breadth of thought in demonstrating proficiency with reading comprehension: “My answer was ok but I didn’t included [sic] evidence. I think my score would have been better if I would have supported it.” Her confidence increased from the previous unit, and shows signs of stability at Robust through the literal levels. At the inferential levels, her PAB rating drops to Vulnerable, as her performance weakens. At Level 5, her last attempt shows a Discouraged PAB.

Suzie’s reflection at the end of the assessment was Modest: “I think that this assessment was easy because I understand most levels now then [sic] before. I did a lot better than last time.”

Narrative nonfiction developmental work. Think-aloud moves at Level 1 continued to be uncorrected miscues, yet the majority of Suzie’s moves shifted to Level 4 and were connected to the metacognitive processes of making sense of the text. An example is: “I actually like this tribe because it can actually respect women and not like dishonor them.”

During the rest of the unit coursework, there was an even split in quality of thinking between Good and OK at Level 4 as well. Interestingly, her thinking was more stable in the neutral to positive range when engaged in inferential thinking. Level 4 is where most of the work took place. This example shows the type of work she was doing:
Title--The start. Question--Why is his in italics? Answer--I think his is in italics because she doesn’t want to think of him as her husband. I think she just rather get away from him that be with him in the first place. Explanation--In my opinion I think I would do that to and that she didn’t even care about him ever. Quote--I think this because of this quote “I was famished and fairy fell upon the rice and meat that ‘his’ sisters had prepared,”(pg 74). In this quote it shows how it was referring to him as “his”.

From observational field notes, the primary PAB for Suzie was Robust, particularly during group discussions where she was a prolific contributor to the discussion, particularly at the inferential level. She did not hesitate to share her views with her peers, and utilized a breadth of thinking to convey her views.

Narrative nonfiction summative assessment. On the Narrative Nonfiction summative assessment, Suzie scored in the Proficient range with a score of Level 5, Stage A. An example from Level 5 shows the general emerging competence at this level:

[Why did people have so much faith in FDR? Be sure to use evidence from the essay to help you explain your opinion. Model--he was positive, hard working, and he got people back to work.] I think people had faith in him because he was clam [sic] about the situation. He even told people not to fear. The other reason was because he had helped so much that they had faith that FDR could get them through this.

Suzie showed a small gain in breadth of thinking from earlier assessments, but it was well below the level of the working time in this unit, and still dominated by answers only. Her responses were predominantly on the positive end of the continuum of quality thinking, with the only instance of Weak being a single response at Level 5.

Before taking the assessment, Suzie’s reflection shows Modest agency: “Before I take this assessment, I am feeling I think I will do fine on it.” As the nature of the academic task increased on this assessment, moving through levels of literal thinking into simple inference, and on to complex inference, her PAB ratings moved down the continuum. They started with Robust and Tenacious at Level 1, replaced Tenacious with
Modest at Level 2, and replaced Robust with Vulnerable at Level 3. A PAB rating of Accepting was first present at Level 4, and is the only rating at Level 5. Suzie reflected a Tenacious rating at the end of the assessment: “I think that this assessment was easy because I understand most levels now then [sic] before. I did a lot better than last time.”

Expository formative assessment. On the Expository formative assessment, Suzie scored in the Basic range, at Level 3, Stage A. Her score of Level 3, Stage A was due to several partial answers at Level 3. For example:

[The invention of writing allowed the Sumerians in Mesopotamia to keep accurate records of . . . Model--business transactions, religious beliefs, and stories of their people and knowledge of medicine, mathematics, and astronomy.] They kept track of food and there farming.

Suzie expressed a Fragile sense of agency with her opening reflection: “Before I take this assessment, I am feeling I think this will be hard. I don’t really understand it.” This Fragile rating also applied to her reflection at the end of the assessment: “I think that this assessment was difficult because I don’t really read non-fiction so it was hard for me. I did bad because I got a Level 3. I think I needed to concentrated [sic] more. That way my score would have been higher.” These PAB reflections are consistent with her academic history; she had a history of struggle with texts that lacked a consistent story grammar. Expository texts were her primary motivation for choosing to participate in the research study. Further illustrating this, she started out with Robust and Modest as her ratings at Level 1 and dropped from there. At Level 2, her ratings are Fragile, Accepting, and Discouraged and her reflection supports this: “I wasn’t concentrated enough.” The predominance of quality ratings of Weak and Wrong throughout also support her perception.
Expository developmental work. Think-alouds featured few moves, most of them at Level 1, that were miscues that went unrecognized and uncorrected. Suzie’s collaboration with Tanner for the unit coursework demonstrated a breadth of thinking focused on Levels 3 and 4, with 3 dominating. Based on observational field notes, Suzie had a largely Vulnerable stance on her personal agency throughout the Expository unit. She was more confident and engaged when working with her collaborator. An example from Level 3 illustrates the dominant level of work, along with the breadth of thinking that was incorporated:

Title--The Horrible Life of Young Children. Answer--Many kids suffered during the middle ages. But when parents worked many toddlers suffered to. Toddlers were left alone all day until parents and children got home from work. Many toddlers sometimes went with the parents to but the just sat and played with shells imitating there [sic] parents. Explanation--This is a cause and effect relationship because the cause is the parents and children working all day, the effect are toddlers staying alone or playing in shells. In our opinion it is sad how toddlers suffered so much even though they didn’t work, many were kept all alone until there [sic] parents were done working. Quote: “Toddlers wandered about the sheds, playing among the shells and imitating their parents,” (pg43) Question--What did toddlers do while parents and children worked?

An example from Level 4 shows this breadth of thinking as well:

"Title--Kids Dying For a Living. Answer--The reason of the accident rate for children working in mills is twice as high as that for adults because it was already hard enough for adults to operate the machines it would be harder for kids to do it. It would also be twice as high because many of the kids were alone they didn’t have there [sic] parents to protect them. Explanation--Our inference is children weren’t as strong as adults and adults are not as fragile as children. Quote--The quote is: “A twelve-year old doffer boy fell into a spinning machine and the unprotected gearing tore out two of his fingers,” (pg.35). Explanation--Our opinion is children shouldn’t be working in the first place. Question--Why would the accident rate for children working in mills be twice as high as that for adults?

Expository summative assessment. On the Expository summative assessment, Suzie scored again in the Proficient range, dropping to a score of Level 5, Stage A. Her Personal Agency Beliefs were again categorized as Tenacious overall. She began the
assessment with a *Tenacious* reflection: “Before I take this assessment, I am feeling think I will do fine, I am not very scared of what I will get.” She wrote a *Strong* response at Level 4, and a partial-credit one at Level 5. The Level 4 response shows a continuation of the quality of work she did during the latter part of this unit:

[Name at least three qualities that the settlers must have had to survive the Oregon Trail.] Three qualities to survive the Oregon trail would be: strength endurance, and the physical ability. Strength--so they could survive the tough weather. Endurance--because of some people dying on the Oregon Trail. Physical ability--To be able to see and experience so many struggles.

At the literal levels, Suzie remained on the positive side of the continuum of *PAB* categories, but her success with the work did not seem to have generalized to an accurate assessment of her potential success on an assessment related to that work. At the beginning inferential level (Level 4), her *PAB* ratings were *Modest* and *Discouraged*. She did not rate her Level 5 responses, although the partial credit she earned on one response resulted in her Level 5, Stage A score. Suzie included a limited breadth of thought in these responses, but the quality of thought was predominantly *Good* or *Strong*. There were, however, several instances of *Weak* at the inferential levels.

Suzie’s *Tenacious* reflection after the assessment shows confidence and maturity:

I think this assessment was appropriately challenging because some questions were challenging. They made you think hard but I got the story and understood it well. I think I did really good. I had worked a lot on my responses and I think I got it all down.

*Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?*

*GMRT-4 and SRI.* Suzie’s Proficient performances on both the GMRT-4 and the SRI are evidence of her closing the gap. Specifically, she demonstrated a growth of 17
months or a rise of 2 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. She scored well beyond grade level on the comprehension subtest of the GMRT-4 administered at the end of the end of the study in December, equally 11 NCEs of growth, calling the March results into question. Suzie’s increase in her SRI score through the end of the year was 165 Lexile points, and her performances vacillated throughout the year, but remained in the proficient range on this measure.

**PARLA.** Suzie made gains in two of the three content areas, as measured by the PARLA proximal assessments, and reached grade level expectations in both Narrative Nonfiction and Expository, providing evidence of closing the gap in those areas. She did not close the gap with Narrative text, remaining 1 stage below grade level. Specifically, On the PARLA-NARR NF, she improved by 2 stages, as compared to the average of just under 4 stages for the rest of the quintain, and moved into the proficient range with this growth. On the PARLA-EXPO she showed growth relative to her immediate peers, achieving four stages of growth relative to the average of near 3 stages, and ending in the grade level, at Level 5, Stage A. Finally, on the PARLA-NARR she showed no change in performance from the formative to the summative assessment, placing her below the average growth of the members of the quintain.

*Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?*

Suzie’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Robust to the
Tenacious category (see Table 1. *The MST Taxonomy of Personal Agency Beliefs*). The most prevalent categories for her throughout the work across the three content areas were: *Modest, Robust, Tenacious,* and *Accepting,* in that order, as shown in the pie graph of Figure 16.

_Figure 16. Distribution of most prevalent PABs in Case 5._

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of *Personal Agency Beliefs* varied: The *Robust* rankings moved from 29% during Narrative, up to a high of 46% during Narrative Nonfiction, and down to 10% during Expository. The *Personal Agency Belief of Accepting* was at 38% at the start of the study, climbed to the greatest level of 55% during the Narrative unit, and dropped down to 10% in the final unit. The *Modest* belief was the least volatile across contents, varying by no more than 6% from the highest to the lowest, as shown in the graph in Figure 17.
Figure 17. Case 5: Tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

Robust is the leading category of Personal Agency Beliefs for Suzie, with 17 instances at the literal levels of comprehension. These were distributed across contents with six in Narrative and seven in Narrative Nonfiction. There were four Robust reflections at the inferential level. For the second most populated PAB level of Modest, there were three responses at the literal level of comprehension. One was in Narrative, the others in Expository. Moving to inferential thinking, there were eight occurrences of Modest beliefs. One was during Narrative, four during Narrative Nonfiction, and three during Expository. For Suzie, the Tenacious category was very close to the Modest category. At the literal levels of understanding there were four responses, two each in Narrative and Expository. The inferential levels of Tenacious happened only during Narrative, with seven total instances. Finally, Suzie had one category on the lower half of the taxonomy with some traction. The Accepting category features eight responses at the literal levels of comprehension, split evenly between Narrative and Expository. For the
inferential levels, there was a similar split of the six responses, with three in Narrative Nonfiction and three in Expository.

Suzie’s artifacts show increase over the course of the study for PABs in the mid or neutral range (Accepting), and decline for the stronger PABs (Robust, Tenacious and Modest) as the work became more challenging for her, as epitomized by her reflection immediately following the first Expository PARLA about how difficult it was.

Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?

MARSI. On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Suzie’s self-reporting was high across all subscales and virtually unchanged from pre to post measures. In the area of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting and similar strategies; Suzie reported a decrease from 4.6 to 4.5 over the study period. For the subscale of Global Reading Strategies, she remained at 4.3. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. Similarly, Suzie’s lowest self-rating of 3.9 for Support Reading Strategies, which include a variety of actions such as note taking, paraphrasing, and discussing the material, among others, dropped from 4.4 at the start of the study, but both scores were in the high range.

Consideration of Suzie reporting her metacognition relative to her think-aloud artifacts and her reading assessment data supports her assertions at the start of the study.
Her high marks across all three MARSI strategy subscales fit with her strong performances that put her predominantly into the Proficient category in reading comprehension across all measures. Table 21. shows her reflections about this survey.

Table 21.

**Reflections on MARS I survey by reading strategy for Case 5.**

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I think I don’t do some of these things all the time.”</td>
<td>“I do these things alot [sic] while I am reading so I think I did good.”</td>
<td>“I do this all the time so I like my score.”</td>
</tr>
</tbody>
</table>

**Think-aloud protocols.** The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, half of the moves in Suzie’s think-alouds were at concrete levels, as shown in Table 22. Overall, Suzie made 51% of her “moves” at Level 1 and 24% at Level 4.

Table 22.

**Frequencies of Levels of think-aloud protocols for Case 5.**

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.81</td>
<td>.05</td>
<td>.05</td>
<td>.09</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.23</td>
<td>0</td>
<td>.23</td>
<td>.54</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>.74</td>
<td>.09</td>
<td>0</td>
<td>.17</td>
<td>0</td>
</tr>
</tbody>
</table>

**PARLA/PARLI.** During the Narrative segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 6, as shown in Table
23. The *Quality* of response was 51% in the positive range, 14% neutral, and 35%
negative, as shown in Table 24.

**Table 23.**
*Case 5: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>118</td>
<td>.23</td>
<td>.12</td>
<td>.27</td>
<td>.29</td>
<td>.07</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>108</td>
<td>.18</td>
<td>.24</td>
<td>.14</td>
<td>.34</td>
<td>.10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>117</td>
<td>.25</td>
<td>.17</td>
<td>.31</td>
<td>.22</td>
<td>.05</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of
responses were distributed across Level 1 though Level 5, with the majority being at
Level 4, combining with Level 5 to result in 44% of her responses being in the inferential
range, as shown in Table 23. The *Quality* of response was fairly evenly distributed across
the spectrum (Table 24).

**Table 24.**
*Case 5: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>97</td>
<td>.18</td>
<td>.33</td>
<td>.14</td>
<td>.30</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>95</td>
<td>.10</td>
<td>.22</td>
<td>.40</td>
<td>.25</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Expository</td>
<td>94</td>
<td>.06</td>
<td>.17</td>
<td>.33</td>
<td>.36</td>
<td>.02</td>
<td>.06</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of 73% of
responses were at the literal levels of comprehension (Level 1 - Level 3), with simple
inference (Level 4) at 22%, and complex inference (Level 5) at 5% (Table 24). The
Quality of responses skewed to the lower ranges, with ratings of \textit{Weak} or lower 44\% of the time, \textit{OK} 33\%, and \textit{Good} or \textit{Strong} 23\% of the time (Table 24).

Suzie’s dominant level of performance/response was high, but her data was distributed across the matrix with some in the middle range of performance or growth. Her performance on the PARLA Narrative assessment inhabited the sole space on the lower end of growth for the quintain. Observations during the initial unit showed hesitance on her part to engage when she spent most of her time checking in with the teacher/researcher and peers. Her broad range of \textit{Quality}, from \textit{Weak} to \textit{Strong} in artifacts during this unit supports this, showing inconsistency signaling doubt. Suzie has a history of preferring to delay engagement until task parameters are crystal clear, avoiding risk of anything short of a strong performance.

\textit{Sub-question. 4) Does student performance on the assessments form a pattern of development?} Suzie’s scores show little variability overall, marking consistent progress. Her PARLA scores did show fluctuation as she moved to a new content area, with marked improvement after each instructional interventions followed by a return to lower levels when the academic challenge changed during the next unit, followed by another post-instruction rise. This is consistent with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006) that posits uneven growth in the development of complex skills.

\textit{Summary.} Suzie performed at the Proficient level on both standardized reading assessments, as well as two of the three proximal measures (PARLA-Narrative Nonfiction and Expository). Her formative and summative assessments of agency and motivation show change from a starting place of \textit{Robust}, to an ending place of \textit{Tenacious}. Her metacognitive scores on the MARSI survey remained in the high range throughout.
By considering the evidence of Suzie’s development regarding reading comprehension, agency and motivation, and metacognition across the course of the study, one can make a case for the PARLI framework being effective with Suzie.

**Case study: Tanner.**

*Previous years of schooling.* Tanner was on an Individualized Education Plan (IEP) for speech in early elementary and was removed from the IEP at the start of seventh grade. He was the one member of the quintain with clinical diagnoses. According to school records, his diagnoses of Asperger’s Syndrome and Obsessive Compulsive Disorder inform his daily classroom experience in profound ways. He performed above grade level in third grade, at grade level in fourth grade, and dropped below grade level in fifth grade, qualifying him for remedial reading at the level of the Read 180 program in sixth grade. This dropping pattern relative to grade level performance was not surprising given the steady increase of inferential understanding required to achieve grade level as one moves up and Tanner’s preferences for concrete thinking connected to his neurological differences.

Tanner was one of the students who started with the researcher in sixth grade. In his initial days in the classroom he rarely made eye contact. This behavior is typical of individuals who share his Asperger’s diagnosis and being new to middle school was obviously a disconcerting experience for him. Tanner has experienced tremendous improvement in his skills and confidence since sixth grade, and he has worked very hard to do this. According to district records, Tanner has performed at a Basic level on the MAP test in Communication Arts for the past 3 years, with no appreciable difference. Tanner’s SRI score was in the Basic Lexile range at the end of last year. He maintained
this performance at the Basic Lexile range at the start of the year, but with some fluctuating movement within that range tied to test anxiety. Tanner has shown consistent growth in GMRT-4 Comprehension scores in the last three years, from 4.1 in sixth grade to 4.5 at the end of sixth grade, and a significant jump at the end of seventh grade to 6. Tanner struggles with timed measures, as he has both OCD and Aspergers diagnoses which seem to drive him to triple-check everything. While his reading scores and general performance merited this placement, at the same time Tanner’s grades averaged 98% across all core classes throughout both school years.

**Research questions.**

*Research question.* Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Tanner. Appendix G features the Event Flow diagram for Tanner that provides a graphic representation for the data.

*Narrative formative assessment.* On the Narrative formative assessment, Tanner scored in the Basic range, at Level 4, Stage A. An example of his response at Level 4 is clear and to the point: [Why does Mr. Tillian bring Rocky to the shop?] “To help Mr. Tillian keep him company.” Not surprisingly, Tanner exclusively ranks his responses as Robust, as he does not approach the items on which he does not feel competent. Metacognitively, his thinking is *OK/Good/Strong* up through the first Level 4, after which he does not attempt to respond. Tanner does not show breadth of thinking, with only answers provided.
Narrative developmental work. During think-alouds in the Narrative unit, Tanner predominantly made moves at Level 1 to connect to the text with an awareness of trying to actively make meaning, but did have some miscues. One example wherein he questions the text, then notices later when he has clarification, is illustrative:

[Anna Semionovna raised her hand with the rod in it, yanked, and cried out. A little silvery-green fish shimmered in the air.] I wonder what that is. [“My lord, a perch! Ah, ah . . . Quickly! It’s getting free!”--read with appropriate emphasis] So that’s what the silver-green fish in there was.

Tanner’s work throughout the unit was predominantly at Level 3 and 4, with Level 3 being more prevalent. His work at Level 4 with the Concept Diagram and the Lane Comparison organizer demonstrated a developing grasp of this level of understanding. Breadth of thought was present, and OK was the dominant Quality of thought. His annotation, using the colors designated for Level 3 and Level 4 to highlight text, and symbols to indicate his thinking, showed metacognition. This was shown in this Level 3 response at an OK Quality level:

[What does Joanna say about the leaves?] Answer--Joanna says that when every leaf falls, she must go, too. Explanation--This is a cause and effect relationship because she’s counting the leaves and the effect when Joanna says she would die because of the last leaf falling. My opinion is that you can’t die from leaves falling one at a time because that’s impossible! Joanna was just being paranoid [sic] about the leaves falling and her life would be at the end. Quote--“. . . leaves on the ivy vine. When the last one falls, I must go, too”.

Tanner’s sense of agency throughout this unit was Vulnerable to Tenacious, as identified through observational field notes. At the start, Vulnerable was displayed by Tanner checking in with the instructor approximately every 5 minutes. As the unit wore on, he checked in when he had completed a response to be sure he was on track.

Narrative summative assessment. Tanner’s reflection before the assessment was Vulnerable overall: “Before I take this assessment, I am feeling so nervous because it’s a
EVALUATING THE PARLI FRAMEWORK

post test and its [sic] for a grade.” Throughout the assessment, his PAB rankings were variable at Levels 1 and 2, with Tenacious, Robust, and Accepting, before settling into Robust exclusively at Levels 3 and 4. No work was attempted past Level 4. At the conclusion of the assessment his reflection about his performance was Vulnerable overall: “I think that this assessment was difficult because the harder thinking questions kind of stumped me on the whole test and my focus.”

Tanner chose not to answer one of the Level 4 questions, but provided a Strong answer for the other, which was consistent (he focused where he felt competent). This example illuminates the high point of his current comprehension:

[Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story?] Answer--Roger doesn’t take Mrs. Jones’ purse when he had the opportunity later in the story because Mrs. Jones was so kindful [sic] at her house to Roger, that Roger didn’t wanted [sic] to do anything bad to Mrs. Jones like stealing her purse. Also, the woman gave $10 to Roger to buy suede shoes. Explanation--My opinion is that Roger didn’t need to take her purse again because he got $10 from Mrs. Jones to buy him something at the store.

During this assessment his breadth of thinking was largely absent. In addition, his written reflection was sparse, but what was present was evidence of emerging metacognition, as these two samples at Level 1 and Level 2 show: “I was too specific in my answer” and “I answered another question.”

Narrative nonfiction formative assessment. On the Narrative Nonfiction formative assessment, Tanner scored in the Below Basic range, at Level 2, Stage C.

Tanner attempted only one Level 3 question while achieving a Level 2 performance: [What was the first group that Eleanor joined that led her to so many other things?] “The first group was The League of Women Voters.” At Level 1, he ranked
himself with Robust and Tenacious, exclusively Tenacious at Level 2, and Accepting and Robust at Level 3. No attempt to move beyond the literal level was made.

Tanner did not reflect in anticipation of the assessment, but his post-assessment reflection showed a Tenacious PAB rating: “I think this assessment was appropriately challenging because nonfiction isn’t like fiction. This is going to be hard comprehending nonfiction stories.”

Tanner did not demonstrate breadth of thinking on this assessment. In addition, the Quality of his thinking was variable, but predominantly on the positive end of the spectrum. Specifically, thinking was predominantly of OK/Strong quality. Missing responses were present at Levels 1 and 3, and Wrong at Level 3 (in addition to Good and Miss).

Narrative nonfiction developmental work. During think-alouds in the Narrative Nonfiction unit, Tanner continued to build on his moves to connect to the text with only two miscues at Level 1, and the remaining moves focused on connecting to develop greater understanding of what was read, up through Level 4. An example from Level 4 was illustrative of this emerging process: “Oh, I remember Alexander the Great from 7th grade Ancient World Civilizations.”

In all work outside of the think-alouds, Tanner focused predominantly on Levels 3 and 4, with 3 being the most common. He participated in discussions, interjecting his comments and questions in the appropriate place, with confidence, resulting in a Tenacious PAB rating based on observational field notes. One example was this question: “I have a Level 4 question, Is the author, Nujood, saying the Monster is her husband?” Another is when he shared this quote: “Chapter 8, page 124, another Level 4 and it says
Nujood is telling us that her country doesn’t know a lot of the things that we do because she said ‘Nujood, those are computers’ and she’s like, ‘They’re what?’ ‘Computers’.”

Tanner showed breadth of thinking, particularly in his discussion preparation and participation. His work was dominated by thinking of Good quality, at all four Levels in which he participated. At Levels 1, 2, and 4, he also did a bit of work at the OK level of quality. This example response at Level 3 demonstrated these characteristics:

*Answer*--At the beginning of the story, Nujood is ashamed because she doesn’t like talking about her personal things and it’s very upsetting to her. *Explanation*--This is a cause & effect relationship because it’s deeply upsetting to her. My opinion is that she doesn’t want to say anything because they could be inappropriate questions or it might be that Nujood is very emotional. *Question*--Why is Nujood ashamed? *Quote*-- “I’m ashamed of talking about these things. It’s deeply upsetting.”

Narrative nonfiction summative assessment. On the Narrative Nonfiction summative assessment, Tanner scored in the Proficient range at Level 4, Stage B. His *PAB* was *Robust* overall during this assessment, as evidenced by his initial reflection:

“Before I take this assessment, I am feeling like I can do this test because I looked at all my past work and did really good so I studied those things.” *Robust* is the dominant PAB throughout, with one instance of *Accepting* at Level 2 when his response was incorrect. His rating shows a lack of awareness about the error of his thinking.

In the summative assessment Tanner’s breadth of thought was evident. He also showed awareness through some of his reflections. For Level 4, the question he did not answer had a *Self-Doubting* reflection that shows awareness: “I did not answer because I thought it was too hard to answer and was too long in the response. Now that I see the correct answer: I’m like ‘oh so that was the answer to that long question.’” Tanner had this response after he completed a lengthy, and *Strong*, Level 4 response on the preceding
question, so he was making the connection. An example of his breadth of thought can also be found at Level 4:

[Why did Roosevelt tell the American people they had nothing to fear?] *Answer*--I think Roosevelt told the American people they had nothing to fear was that he was mostly saying that he was going to help this crisis [The Great Depression] until the very end. *Explanation*--Roosevelt created a whole bunch of agencies and programs that was going to help The Great Depression. *Quote*--The CCC “used government money to hire young men to work outdoors on public projects like clearing land & building dams.” The CWA “hired men & women to work on other government projects that included building libraries & airports.” The WPA “hired workers that included writers & artists to create guidebooks & public art pieces”.

His *Robust* reflection at the end of the assessment demonstrated the power of the PARLI framework for Tanner:

I think that this assessment was easy because everyone had a lot of time on doing responses and we had a lot of practice on them. I feel really great about myself! I went up 1 level and 2 stages! That is a lot of progress in reading the book!

*Expository formative assessment.* On the Expository formative assessment, Tanner scored in the Basic range, at Level 3, Stage A. This Level 3 response to when the Shang Dynasty began illustrated the level of his work: “The Shang Dynasty began with T’ang. He overthrew the evil emperor of previous dynasty.” During this formative assessment some breadth of thinking was evident, while his quality of thinking was variable. An example from Level 2 showed breadth of thinking:

[How were kings of the Shang Dynasty buried?] *Model*--Kings were buried in deep shafts with a wooden burial chamber, along with their court and slaves and their possession for the next live. Slaves were even buried alive. *Answer*--Kings of the Shang dynasty were buried in impressive burial pits. *Explanation*--It had deep shafts that had a wooden burial chamber. They believed in an after life they buried royalty with things that were valuable for use in their next life.

Tanner’s quality of thinking is spread across *OK, Good, Strong*, and *Missing*, with *Missing* making up half of the entries. His reflection was *Self-Doubting* at the start:
“Before I take this assessment, I am feeling nervous because it’s a whole new unit of reading for me.” That he did not rate his confidence in his responses, together with his Self-Doubting rating, showed his stress on this assessment. He indicated that he stopped after the first Level 3 response because: “I ran out of time doing the assessment.” The assessment was not timed and he could have asked to complete it, but his lack of confidence brought him to the decision to just stop. In line with his obvious stress, his rating upon completing the assessment was Vulnerable:

I think that this assessment was difficult because Expo isn’t the same as fiction and non-fiction. It’s a little too hard!!! I didn’t try hard enough on this assessment and that it was harder than I thought it would be.

Expository developmental work. During this unit work, Level 1 and Level 4 connections dominated the think-alouds. Metacognitively, Tanner’s think-alouds had few moves, but continued his earlier efforts with connections derived from his metacognitive awareness of making sense. In the work overall, Level 3 dominated with complete and complex answers. Level 4 answers were also well-constructed, as seen in the examples. Most of the coursework in this unit was completed collaborating with Suzie, and Tanner appeared to benefit from their collaboration. His work throughout the unit was predominantly demonstrative of Good and OK thinking. Examples of breadth of thinking in Level 3 and Level 4 responses (respectively) illuminate Tanner’s metacognitive growth:

**Title**--Hine’s Great New Job. **Answer**--Hine’s took up photography to help kids out that were being forced into child labor. He took pictures because he wanted to see how child labor was and then he noticed how bad it was. **Explanation**--My opinion is that it is a good job for him because he can see how bad kids have it. **Quote**--“To use his camera as a weapon against the exploitation of children” (pg 19.). **Question**--Why did Hine take up photography?
Title--Kids Dying For a Living. Answer--The reason of the accident rate for children working in mills is twice as high as that for adults because it was already hard enough for adults to operate the machines it would be harder for kids to do it. It would also be twice as high because many of the kids were alone they didn’t have there parents to protects [sic] them. Explanation--Our inference is children weren’t as strong as adults and adults are not as fragile as children. The quote is: “A twelve-year old doffer boy fell into a spinning machine and the unprotected gearing tore out two of his fingers” (pg. 35). Explanation--Our opinion is children shouldn’t be working in the first place. Question--Why would the accident rate for children working in mills be twice as high as that for adults?

From observational field notes a more complete picture of Tanner’s sense of agency develops. He showed a Vulnerable sense of agency while collaborating with his partner. Sometimes, he and his collaborator engaged and discussed their thinking before and during writing responses, at other times they worked in parallel. Given his diagnosis of Aspergers, this is not surprising. In addition, during this unit of study the weather resulted in a 6 snow day closures. This created a back log of work to be completed across all core courses. Tanner was sensitive to this, and notes indicate a number of days when he came in (this class was toward the end of the day) and put his head down. Finally, field notes also indicate his repeated efforts to connect the thinking framework of PARLI to his other coursework across all content areas. One example highlights this: “I think what she wants here is Level 4 thinking. Can you check? If she does, I can do it. I am good at Level 4 now. But if it’s Level 5, that’s harder and I’m not as good at that yet.”

Expository summative assessment. On the Expository summative assessment, Tanner scored at the top of the Basic range, with a score of Level 4, Stage A. Tanner’s sense of agency started out at Self-Doubting for the summative assessment: “Before I take this assessment, I am feeling really nervous because it’s a test and its for a grade, too. I’ll just do my best on it!” Field notes show that he paced the room a few times, came to the instructor and said “I’m really nervous. I had two other tests today. Why do
we have so many tests?” His sense of agency showed a small amount of variability, most likely attributed to his stress level at the start of the assessment, with PAB at Levels 1 and 2 split between Tenacious and Robust. At Levels 3 and 4, they were exclusively Robust. This is not surprising, given his focus on Levels 3 and 4 during the coursework and his stated confidence in his competence at these Levels. His PAB reflections during the assessment showed metacognitive awareness, as in this sample from Level 1: “I didn’t read when it said ‘specifically where is the essay set?’” When it came time to reflect about his performance, his sense of agency changed to Vulnerable: “I think that this assessment was difficult because expository is very challenging and difficult to me. It’s not like fiction and non-fiction at all” shows his metacognitive awareness.

Tanner’s Quality of thought was OK and Good at literal levels. Level 4 was Strong. An example from Level 4 highlights his best work on this assessment:

[How is the time of the Great Migration similar to what is going on in the world today] Answer--The Great Migration is similar to what is going on in the world today is their both economic depressions. Explanation--I can infer that “The Great Migration” was worse than today’s because banks closed, people lost savings, non one [sic] had a lot of money probably because “the people were too desperate,” since they didn’t go to work.

Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

GMRT-4 and SRI. Tanner’s performance at the Basic level on both the GMRT-4 and the SRI are evidence that he did not succeed in closing the gap. He demonstrated a growth of 12 months for an increase of 1 NCE on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the
other four middle schools of the test district. Tanner’s decrease in his SRI score through the end of the year was 84 Lexile points, placing him well below the average performance of 51 for his counterparts in the quintain.

**PARLA.** Tanner made gains in all three content areas, as measured by the PARLA proximal assessments, but remained below grade level expectations in all areas, providing further evidence of his inability to close the gap when assessment data is exclusively used to make the evaluation. Specifically, on the PARLA-NARR he improved by 1 stage from formative to summative assessment, placing him just below the average growth of the members of the quintain. On the PARLA-NARR NF, he showed his greatest growth, improving by 5 stages, as compared to the average of just under 4 stages for the rest of the quintain. On both the Narrative and Narrative Nonfiction areas his final performance was at Level 4, Stage B, becoming stable with understanding basic inferences. This performance was consistent with his written class work, but lower than the thinking he demonstrates in less formal moments.

He shared his sentiment that all written work is “for a grade,” and in his world view, this made it formal and important that he perform well. Tanner generally chose to perform at his guaranteed level of competence rather than risk his reach exceeding his grasp. Again, this is not unusual for students with his neurological diagnoses. Finally, on the PARLA-EXPO, he achieved 3 stages of growth relative to the average of near 3 stages, ending at the beginning stage of demonstrating understanding inferences (Level 4, Stage A).
Figure 18. Distribution of most prevalent PABs in Case 6.

Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Tanner’s self-reporting about his motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Tenacious to the Vulnerable category (see Table 3. The MST Taxonomy of Personal Agency Beliefs). For Tanner, Robust was the most prevalent PAB rating by a sizeable margin, with Modest, and Tenacious being the other categories of some depth, as shown in the graph in Figure 18.

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied. The Robust rankings moved from 32% during Narrative, down to 23% during Narrative Nonfiction, and then to 22% during Expository. The Personal Agency Belief of Modest showed a rise across the course of the study, beginning with 13% during Narrative, growing to 31% in Narrative Nonfiction, and ending with 62% in Expository. The Tenacious Belief began at a relatively high level of 44% in Narrative, changed to 11% in Narrative Nonfiction, and rose dramatically to 89% in the final unit of the study (Expository), as shown in the graph in Figure 19.
Figure 19. Case 6: Tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

Robust is the leading category of Personal Agency Beliefs for Tanner by a tremendous margin. There are 14 instances of Robust PAB in the Narrative section at the literal levels of understanding, with 10 in Narrative Nonfiction and five in Expository. The next category in the taxonomy for which Tanner had a collection of reflections was Modest, with four at the literal levels split between Narrative Nonfiction and Expository and all being labeled as OK. Of the seven Modest occurrences at the inferential levels, one occurrence was in Narrative and five were in Narrative Nonfiction and only one instance in Expository. Tanner’s reflections at the Tenacious category on the taxonomy were predominantly at the literal levels of comprehension. Moving up the inferential level, Robust was the label for five reflections, three in Narrative and one each in Narrative Nonfiction and Expository.

The Narrative summative assessment was one of the many high moments in the process with Tanner, but there were some big challenges, particularly when it came time
to assess for a grade. At least once a week, the pressures of eighth grade in a public school setting were overwhelming for Tanner. One example from observations is typical and illustrates this point:

Tanner is working alone today instead of in dyad or group. He is putting his head down a lot (he had two tests today, both of which he struggled with). He is on almost complete shut down. The other students respect his space and work in dyads and small groups around him. (Field notes, Nov. 3, 2010)

Sub-question. 3) *Do students participating in PARLI demonstrate improved metacognition?*

Tanner embraced working at the inferential level of comprehension and was determined to gain mastery. His efforts were apparent through observations and artifacts. In addition, unlike many of the other members of the quintain, Tanner also engaged in inferential thinking with a frequency averaging about one-third of the time when it came to the think-aloud protocols. He focused on mastering Level 4, and did so consistently, but remained intimidated by Level 5 thinking and work in most cases. That said, he did embrace attempts at Level 5 when working in a dyad or small group; again, having a social learning group as an incentive for him was a big development.

*MARSI.* On the Metacognitive Awareness of Reading Strategies Inventory a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Tanner’s self-reporting indicated high ratings across all three categories of reading strategies for both pre and post measures. His did not report any changes greater than .1 for any of the three subscales. Based on observations of Tanner over 3 years’ time, the researcher agreed with his assessment of his awareness of his thinking processes during reading comprehension.
Tanner gave himself very high ratings, all of them almost perfect, on both the pre and post measures for the MARSI. This in and of itself was not surprising; however, the reflections that accompany his first MARSI suggest that he may have been taking the survey as if he was asked to provide answers about each of the strategy categories (Table 25). These responses suggest some potential confusion. However, he did, in fact, use a significant group of strategies that cut across all three categories with frequency. If he did use all of these strategies at the levels reported, it is most likely that all of his reading comprehension test scores would have reflected greater growth and movement into proficiency across the board.

Table 25.

Reflections on MARSI survey by reading strategy for Case 6.

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“It is rereading the text back.”</td>
<td>“By using predictions”</td>
<td>No comment.</td>
</tr>
</tbody>
</table>

Think-aloud protocols. The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. While, the majority (60%) of the moves in Tanner’s think-alouds were at concrete levels, 29% were at the basic inferential level, as shown in Table 26.
Table 26. *Frequencies of Levels of think-aloud protocols for Case 6.*

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.73</td>
<td>0</td>
<td>.05</td>
<td>.23</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.31</td>
<td>.38</td>
<td>0</td>
<td>.31</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>.61</td>
<td>0</td>
<td>.06</td>
<td>.33</td>
<td>0</td>
</tr>
</tbody>
</table>

*PARLA/PARLI.* One of the most exciting things about Tanner’s participation was his almost immediate transport of the framework to his other courses. Within a week of introduction and practice in the PARLI framework, he began to bring in work from his other core classes, particularly U.S. History and Communication Arts, and ask about the nature of the work he was being asked to do.

On the first occasion of this connection, Tanner first brought a Communication Arts packet and showed it to the researcher, asking, “I just want to check, are they asking for Level 3 thinking here?” The researcher responded that he was absolutely right, which produced a smile. He asked if he might work on it a moment to see if he needed any further help determining the level of thinking being asked of him, and received consent to do so. He went to a table for a few moments (students were working independently) and came back with: “Is this just a little inferring, Level 4, or more like Level 5? I can’t do Level 5.” Again, and not surprisingly, he was correct in his assessment, and this was confirmed while assuring him that he could in fact think at that deeper inferential level.

The researcher reminded him that she had witnessed him doing it already (Field notes, Sept. 2, 2010). The possession of a framework that he could use across the content areas gave him tremendous comfort in approaching his schoolwork with much less hesitation.
Other members of the quintain noticed what Tanner was doing and began similar behavior almost immediately.

During the Narrative segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 4, as shown in Table 27. The Quality of response was predominantly OK or Weak, as shown in Table 28.

Table 27. *Case 6: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>118</td>
<td>.23</td>
<td>.12</td>
<td>.27</td>
<td>.29</td>
<td>.07</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>108</td>
<td>.18</td>
<td>.24</td>
<td>.14</td>
<td>.34</td>
<td>.10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>117</td>
<td>.25</td>
<td>.17</td>
<td>.31</td>
<td>.22</td>
<td>.05</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 5, with the responses at Level 1 through Level 4 evenly distributed, as shown in Table 27. The Quality of response was at the midpoint or better (OK and up) 81% of the time (Table 28).

Table 28. *Case 6: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>97</td>
<td>.18</td>
<td>.33</td>
<td>.14</td>
<td>.30</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>95</td>
<td>.10</td>
<td>.22</td>
<td>.40</td>
<td>.25</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Expository</td>
<td>94</td>
<td>.06</td>
<td>.17</td>
<td>.33</td>
<td>.36</td>
<td>.02</td>
<td>.06</td>
</tr>
</tbody>
</table>
During the Expository segment of the PARLI framework, the levels of 80% of responses were at the literal levels of comprehension (Table 27), while the Quality of responses was rated as at the median score of OK 47% of the time, with the positive rating of Good being the next highest rating, at 25% (Table 28).

Sub-question. 4) Does student performance on the assessments form a pattern of development? Tanner’s scores were persistently at the Basic range on quantitative measures, but his PARLA scores showed some change in a variable pattern that fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process. The progress Tanner made during the course of the study is not readily apparent from the pattern of his test scores.

Summary. While Tanner performed at the Basic level on all reading assessments given, his formative and summative assessments of agency and motivation show no change from Robust, and his metacognitive scores on the MARSI survey changed only slightly. However, his competence in communicating what he understands at an inferential level developed over the course of the study. By considering evidence from observational field notes of his performance, agency, and motivation during the regular coursework in all areas, one can make a case for the PARLI framework being effective with Tanner.

Case study: Sandra.

Previous years of schooling. Sandra (Case 7) arrived in the district in the middle of her third grade year, having been born in the United States with Spanish as the primary language spoken in the home. She received English Language Learner (ELL) services for her entire elementary experience; she progressed at expected grade levels throughout
elementary school. Based on her GMRT-4 reading score at the end of her fifth grade year, Sandra was placed in a reading intervention program that focused on decoding issues for sixth grade.

According to district records, Sandra has performed at a Basic level on the MAP test in Communication Arts for the past three years, with steady progress of about 20 points per year. Sandra’s SRI score fluctuated a bit and was in the mid range of Basic/Has Partially Met Standard Lexile range at the start of this year. Sandra has not shown consistent growth in GMRT-4 Comprehension scores in the last 3 years, from 3.9 in sixth grade to 5.5 at the end of sixth, and dipping down to 5.2 at the end of seventh grade. However, her GMRT-4 Vocabulary scores have seen a steady increase, resulting in an increase of her GMRT-4 Total score overall from 3.3 at the end of fifth grade, to 5.4 at the end of seventh grade, leaving her still significantly below grade level.

**Research Questions.**

*Research Question.* Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Sandra. Appendix G features the Event Flow diagram for Sandra that provides a graphic representation for the data.

*Narrative formative assessment.* On the Narrative formative assessment, Sandra scored at the top of the Basic range, at Level 4, Stage A. All Levels were attempted, with no breadth of thinking present, only simple answers. Her quality of responses was *Good* at the concrete levels and *Weak* at inferential and advanced. Her *Modest* reflections
showed the beginning of effectively using connections between text and self to improve comprehension. An example at Level 4 showed her understanding of the text at an inferential level: [Why does Harry keep walking past the shop after Rocky arrives?]

“Harry keeps walking past the shop because he wants to see what his dad is doing he wants to know what hes [sic] doing with ‘Rocky’.”

An example from Level 5 illustrated that Sandra was at the emerging stage when it came to complex inference:

[How does Mr. Tillian’s behavior affect Harry throughout the story? Be sure to address the beginning, middle, and end of the story. Use at least 2 details and/or examples from the story to support your answer.] At first Mr. Tillian has a young son and really cares about him. Later he buyses Rockey [sic] and embareses [sic] Harry buy talking to the parrot. At the end Mr. Tillian gets hurt and Harry now talks to Rockey [sic] and dosent [sic]care as much.

Sandra was rather stable with her accuracy in rating her confidence in her responses relative to the actual score she earned, which was indicative of metacognitive awareness. One instance of Accepting at Level 2 and Modest at Level 3, otherwise Robust through Level 5. Given the difficulty of the task, it was not surprising that her sense of agency declined to Discouraged at Level 6 and Accepting at Level 7.

_Narrative developmental work._ During think-alouds in the Narrative unit, Sandra did not reflect with any frequency. She made very few moves during the readings, despite repeated trainings and requests to do so. At Level 1, most of her moves were Weak in nature, with half of the responses connected to miscues that went unacknowledged.

Sandra focused on Levels 3 and 4 for the unit coursework, with Level 4 being dominant. Throughout the coursework for this unit, Sandra reflected at Robust and Tenacious PAB when evaluating her work. In addition, field note observations corroborate this. She was confident in her ability to figure out the nature of the academic
task and meet those challenges. Depending on the particular challenge of the day and her frame of mind (*Robust*), Sandra would choose to work independently. On her *Tenacious* days, she would collaborate with a partner and ask the instructor for more guidance.

For this coursework some breadth of thinking was present at both Level 3 and Level 4. The *Quality* of thinking in evidence at this Level was predominantly *Good*. Sandra engaged in heavy annotation of texts, using the colors designated for Level 3 and Level 4 to highlight text, and symbols to indicate her thinking.

Sandra’s work with the Concept Diagrams and Lane Comparison organizers illuminated her depth of metacognition around the nature of reading at different levels of comprehension. She exhibited *Robust* PAB during this work, leading her collaborators through the process.

*Narrative summative assessment.* On the Narrative summative assessment, Sandra scored in the Proficient range, with a score of Level 5, Stage A. At the literal levels, the dominant *Quality* of thinking was *Good*. A Level 3 example illustrated:

> [At the end of the story, what reason does Mrs. Jones give for understanding Roger’s behavior when they first meet?] Mrs. Roger at the beginning was shocked [sic] at what the kid did and was being pretty mean to him, but at the end she notices he’s a kid and we all make mistakes and realizes [sic] she did things in the past to so she was being understanding.

At Level 5, Sandra’s reflection showed metacognitive awareness of the difference between her thinking and what the question demanded of the reader: “I did not put the middle reason why, that’s one of the big ideas.” In her Level 5 responses her thinking was split between *OK* and *Weak* in quality, as a result of her brevity. For example, when a question asked for a paragraph-long response including evidence from the beginning, middle, and end of the story, Sandra wrote the following:
Roger in the beginning is immature because he wants to steal [sic] money to buy some shoes! But at the end I think he feels comfortable with Mrs. Jones like a mom would be, and gains some respect and matures over time knowing what he did wrong.

Sandra’s initial reflection prior to the assessment was Vulnerable: “Before I take this assessment, I am feeling nervous because on the assessment I did not do really well, but I have learned a lot so I feel so-so.” Throughout the assessment, Robust was the dominant PAB rating. At the literal levels, there was one instance each of Vulnerable and Accepting. At the inferential levels, there was one instance each of Modest and Tenacious. Her written reflections showed a Modest PAB as she made a metacognitive statement comparing her process of comprehending and showing what she understood with the sample response that was shared with her after the assessment. She did not demonstrate a breadth of thinking during this assessment, limiting herself largely to answers without any of the supporting thinking. Sandra’s reflection after taking the assessment explained her Robust sense of agency, providing further evidence of her metacognitive development:

I think that this assessment was easy because I read the story carefully [sic] and reread through parts I didn’t understand that helped answering hard questions. Also I knew what the story was about and I felt pretty comfortable. I think this Level is good based on the work that I did.

Narrative nonfiction formative assessment. The Narrative Nonfiction unit was where Sandra showed the greatest spread of competence from start to finish. When taken in light of her other performances, and her lack of an anticipatory reflection on this assessment, it is probable that outside factors contributed to her Below Basic initial performance. Sandra scored at Level 2, Stage C with an overall rating of Vulnerable for her Personal Agency Beliefs. Her thinking on this formative assessment was
predominantly of Weak/Wrong quality, with the absence of breadth of thinking, as she stuck to simple answers and made no effort to support or explain her thinking. The exception was at Level 2, where Good thinking was dominant, making Level 2 the only place where her responses showed competence: “The great depression was basically [sic] a bad depression from everything worldwide, economic crisis, people losing jobs, savings, and homes.” At Level 3, metacognition was absent when at first she did not read one of the questions carefully and responded to something else. Then she did not re-read her response to the second question and her answer was incoherent.

Further highlighting the anomalous nature of this assessment among examples of Sandra’s work, her PAB Rankings were not stable, with Robust being most prevalent, followed by Tenacious at the literal levels. Modest, Vulnerable, Fragile, and Discouraged were also represented in the rankings and the written reflections. Written reflections about responses showed metacognitive awareness about the nature of the cognitive demand of the questions and how the student’s response varies. These reflections predominantly showed a Modest PAB, but one of them at Level 3 was Fragile, reflecting the Discouraged PAB from the ranking on that question: “I did not know at all on what it was talking about I did not understand it.” At the inferential levels, the PAB rankings were Hopeless and Accepting. PAB was Vulnerable overall at the end:

I think that this assessment was difficult because it was not an interesting story and I think I am better at fiction than Nonfiction. I thought it was a difficult story. I feel weird because I did really good on the last assessment and on this I did pretty bad wich [sic] made me feel like when what happened but maybe it was a difficult story.

Narrative nonfiction developmental work. During think-alouds, Sandra demonstrated metacognitive awareness throughout all levels, as her moves articulated her
thought processes aimed at understanding. This began at Level 1 when her moves were
split between miscues that escape notice and moves that showed a basic metacognitive
awareness about vocabulary and how knowing the words is critical to making meaning.
An example of a move at Level 4 illustrates her newly shared metacognition in these
think-alouds: “By the tone of it and the way they’re using it, I think they’re kidnappers
but I’m not sure.” There were almost as many moves at Level 4 as there were at Level 1.

During the Narrative Nonfiction unit coursework, Sandra began with work to
develop competence at Level 3 and 4, and quickly mastered Level 3. She then shifted to
focusing exclusively on Level 4, demonstrating a breadth of thought throughout this
work. Her outcomes (including discussion) were in the neutral to positive range of
thinking across all activities. Specifically, at Level 1, her thinking was split between
*Good* and *Weak* quality ratings. At Level 4, where she spent most of her time working,
*OK* was the dominant quality of thought, with *Good* not very far behind.

Observational field notes show a pattern of *Vulnerable* PAB, as she stayed around
the edges of the discussion about the memoir. When prompted by the requirement to
participate at least three times in each session, she would speak up. When this
requirement was met, she limited her engagement (Field notes Oct, 27, 2010; Nov. 2,
2010; Nov. 4, 2010).

*Narrative nonfiction summative assessment.* On the Narrative Nonfiction
summative assessment, Sandra scored again in the Proficient range, increasing to a score
of Level 5, Stage B. Anticipating the assessment, her reflection was Modest: “Before I
take this assessment, I am feeling confedant [sic] because I feel like in the other one I did
I did not do well so now I’m feeling good that I’ll do good.”
Overall, Sandra’s quality of thinking was Good to Strong and showed breadth of thought, with the exception of the work at Level 4. An example from Level 5 illustrated:

[Why did people have so much faith in FDR? Be sure to use evidence from the essay to help you explain your opinion] Answer--People had so much faith in FDR because he had done a lot of good things and never really messed things up. He had helped people inspired people and had been governor before, Quote--“In 1928 he became governor of New York where he was also reelected and served two terms.” Explanation--That shows that he had been elected more than once because people loved his work. So people had nothing to worry about.

At the literal levels Sandra showed Robust PAB. At Level 4 she stumbled to Fragile, and rebounded to Robust again at Level 5. When reflecting about PAB at Level 4, where she had some difficulty with the responses, she demonstrated metacognitive awareness: “I could have given quotes from the story to support my answer.” At Level 5, she chose to draw a happy face to reflect her pleasure in her successful responses. Her post-assessment reflection appropriately shows a Robust sense of agency:

I think that this assessment was easy because it was a well thought out story and the questions made sense. I feel ok and hope I did good. I think this is a good score and I feel good that I got this. 5B is a good place to be and I accept that.

Expository formative assessment. This assessment marked the highest initial score for both Sandra and the quintain on any PARLA measure, with her starting out at the Proficient level earning a Level 5, Stage B score. She had some limited success all the way up to Level 6:

[Use what you read about the class systems of Mesopotamia and the Shang Dynasty class system to explain why making sure all students read and write well is so important in the United States today. Model--The higher classes were taught to read and write. Everyone must know how to read and write now to give everyone a chance to rise. Those who can read and write have a say in how things work. In our democratic system, the stronger the education of the people, the better the system works for everyone in the society.] What they used back then is important to us today because, they wrote on clay tablets wich [sic] then led on to paper, wich I infer helped lots of people, like expressing feelings, wich led to
great writers today. Also, they wrote in cuniaform [sic] which helped kids with education and led on to us writing English and many other forms to.

Sandra’s PAB was Modest at the start: “Before I take this assessment, I am feeling ok because I did good on the last one and hope to do good again.” Some breadth of thought was present in her responses, but simple answers dominated. As with Narrative, her written reflections showed a Modest PAB as she made a metacognitive statement comparing her process of comprehending and showing what she understands with the sample response that was shared with her after the assessment. Her post-assessment reflection maintained this Modest sense of agency: “I think that this assessment was easy because it had a lot of paragraphs and it just seemed hard. I feel proud of myself. I thought I would do bad but I did pretty good!”

*Expository developmental work.* Sandra’s think-aloud work in Expository continued the trend begun with Narrative Nonfiction. She again demonstrated metacognitive awareness through all levels, as her moves articulated her thought processes aimed at understanding. This began at Level 1 when her moves were split between miscues that escaped notice and moves that showed a basic metacognitive awareness about vocabulary and how knowing the words is critical to making meaning. As with Narrative Nonfiction, her moves were largely split between Level 1 and Level 4.

Level 4 was the main focus of the coursework in this unit for Sandra. Work done here was up to Stage C, writing questions at this level, as well as supporting thinking with quotes from the text. In all of this coursework, she collaborated with Holly, and the breadth of thought shows significant metacognitive development. This example of Level 3 work of *Good Quality* is illustrative:
Title--The records were never found. Question--Why don’t we know much about Confucius’s wife? Answer--We don’t know much about Confucius’s wife. Because in China many other ancient societies take place in home. A lot of women did not as a rule, play a visible role in the public life of ancient China so they were not in any public records but their husbands were. Quote--“During these years, Confucius married. His wife gave birth to a son and a daughter but we know nothing else about her, not even her name. We don’t know how long they were married, if they separated, or if she died. Women did not role a visible in the public life, and while they were influential and behind the scenes very often, they were not always mentioned in historical records.” Explanation--We know this because in the book it said that they didn’t have any record of the wives not even alive.

Similarly, a Level 4 example further demonstrated the Good Quality of thinking shown by this partnership:

Title--Friends Question--What does Yan Hui mean when he says, “While you, Master, are alive, how would I dare to die?” Answer--Yan Hui said this because he really likes Confucius, and with him around its nothing but fun. Quote--“Whenever he went, Confucius seemed to fit in and feel right at home.” Explanation--This to us means the two, Yan Hui and Confucius fit in just right together. Like penutbutter [sic] is to jelly. In our opinion we think Yan Hui cares about Confucius a lot and for him to be alive he wants to be there with him to. Quote--Also, “Confucius worried about him, and when they reunited he was greatly relieved.” Explanation-- And again this meant to us Yan Hui would not be the same without his Master with him.

Expository summative assessment. On the Expository summative assessment, Sandra scored again in the Proficient range, dropping to a score of Level 5 Stage A. Her PAB was Tenacious at the start: “Before I take this assessment, I am feeling confident [sic] I feel I did well last time and maybe I’ll do good again.” Through Level 4 Robust is the PAB rating. At Level 5 it is split between Robust and Accepting, and at Level 6 it is Discouraged.

Good thinking dominated at the literal levels, and up to basic inference. At Levels 5 and 6, Weak and Wrong were predominant. Sandra did exhibit some breadth of thought
on the assessment, but not as much as during the coursework. An example from Level 4 illustrates the *Good Quality* of thinking throughout the first four levels of this assessment:

[Name at least three qualities that the settlers must have had to survive the Oregon Trail.] *Answer*--These settlers had to have speed and strength, *Quote*--“There were many dangerous river crossings along the trail.” *Explanation*--Wich [sic]meant they had to be fast and strong to be able to not drown or trip and get really hurt. You would also have to have skill on how to kill animals, and you have to be prepared for anything, *Quote*--“Rattlesnakes were plentiful on the trail and proved to be quite deadly.” Reading this you know you have to have certain skills.

An example from Level 5 shows an *OK Quality* of answers and thinking at the more advanced levels of inference:

[How is the time of the Great Migration similar to what is going on in the world today?] *Answer*--The Great Migration was an economic depression. And what's [sic]going on today is that the US is having many economic problems such as the “Great Migration.” *Quote*--“In fact, during this time thousands of people in New York, Baltimore and Philadelphia were without work.” Hearing this it made me think about us the US because of all economic problems and many people not having jobs.

Sandra did not engage in much written reflection about PAB, and what was present lacked metacognitive awareness A reflection from Level 5 exemplifies the way in which her reflections escaped metacognitive awareness on this assessment: “Didn’t really give the specific details.” Her final reflection showed a *Tenacious* sense of agency: “I think that this assessment was difficult because it was a hard story. With hard questions. Also I didn’t understand some of the questions that they put. This is ok I did okay so im [sic] pretty proud.”

*Sub-question.* 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?
GMRT-4 and SRI. Sandra’s performance at the Basic level on the GMRT-4 did not provide evidence of narrowing the gap. She demonstrated a growth of 15 months or a rise of 6 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. However, Sandra’s Proficient performance on the SRI does represent a closing of the gap between her and her non-struggling grade-level peers. The increase in her SRI score through the end of the year was 146 Lexile points, and placed her in the proficient range on this measure.

PARLA. Sandra made gains in all three content areas, as measured by the PARLA proximal assessments, and reached grade level expectations in all three as well, providing further evidence that she closed the reading comprehension gap. Specifically, on the PARLA-NARR she improved by one complete level (4 stages) from formative to summative assessment, placing her above the average growth of the members of the quintain. On the PARLA-NARR NF, she improved by 8 stages, as compared to the average of just under 4 stages for the rest of the quintain. Finally, on the PARLA-EXPO she declined by 1 stage, but still ended with a grade level performance. On both Narrative and Expository Sandra’s final performance was at Level 5, Stage A, while with Narrative Nonfiction her final performance was at Level 5, Stage B.

Among Sandra’s artifacts, her richest store was in the annotation for the full-length memoir she read during the Narrative Nonfiction unit. This piece grabbed her imagination, and she annotated rather extensively, where most of her peers made cursory attempts to satisfy the request to annotate. Most of her annotations showed an effort to
connect with the protagonist, Nujood, a young Yemeni woman forced into marriage at the age of 10 (Ali & Minoui, 2010). At the start of the story, her comments on page 23 and 24 were illustrative: “If I was in a situation where I had to agree to do whatever it would be terrible,” followed by “I like running around too and lay on the grass when I’m tired.” Some of Sandra’s later annotations showed her questioning what it was like to be Nujood. On page 38 she wrote: “I wonder how she felt for the very first time she heard waves and all.” On page 93 she wrote: “If I were in her position I would just run away . . . so why didn’t she do something?”

Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Sandra’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) remained constant in the Robust category (see Table 1. The MST Taxonomy of Personal Agency Beliefs). The most prevalent categories for her throughout the work across the three content areas were: Robust, Modest, and Tenacious, in that order, as shown in the pie graph of Figure 20.

![Figure 20. Distribution of most prevalent PABs in Case 7.](image-url)
Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied: The Robust rankings declined throughout, moving from 27% during Narrative, to 21% during Narrative Nonfiction, and down to 19% during Expository. The Personal Agency Belief of Modest was fairly constant at 31% at the start of the study, then dropping slightly to 28% during the second unit and remaining there for the final unit. The Tenacious belief was the most volatile across contents, remaining at 25% for the first two units, and jumping to 50% in Expository, as shown in the graph in Figure 21.

Robust was the leading category of Personal Agency Beliefs for Sandra by a tremendous margin. At the literal levels, there were nine reflections in Narrative, and six each for Narrative Nonfiction and Expository. There were six Narrative instances of Robust, four for the Narrative, and three Expository instances of Robust at the inferential levels. The Modest category begins with a disbursement across the three contents of three in Narrative, three in Narrative Nonfiction, and two in Expository. There was one occurrence of Modest at the advance literacy levels in Narrative. There were two Tenacious ratings attached to different levels of reading comprehension, one each at the literal and inferential levels with Narrative Nonfiction and Narrative, respectively. Sandra’s reflections on the PARLAs during the Narrative Nonfiction and Expository units also illustrated her growth, as demonstrated earlier.
Figure 21. Case 7: Tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

Sub-question. Do students participating in PARLI demonstrate improved metacognition?

MARSI. On the Metacognitive Awareness of Reading Strategies Inventory (MARS1), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Sandra’s self-reporting indicated all high ratings in the pre measure, with one of the subscales falling to the midrange in the post. For the subscale of Global Reading Strategies, Sandra reported a constant 3.5 over the study period. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. The area of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting, and similar strategies; Sandra slightly revised her self assessment from 4.4 to 3.5, remaining at the high rating
for this subscale. Sandra’s lowest self-rating of 3.1 in the post for Support Reading Strategies, which include a variety of actions such as note taking, paraphrasing, and discussing the material, among others, was a decrease from her pre-evaluation of 3.7, and moved her from the high to the middle part of the scale.

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, the majority of the moves in Sandra’s think-alouds were at concrete levels (Table 29). Overall, Sandra made 83% of her moves at Level 1 and Level 2 combined.

Table 29.

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.61</td>
<td>.29</td>
<td>0</td>
<td>.10</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.54</td>
<td>.25</td>
<td>0</td>
<td>.21</td>
</tr>
<tr>
<td>Expository</td>
<td>.59</td>
<td>.16</td>
<td>.02</td>
<td>.20</td>
</tr>
</tbody>
</table>

*PARLA/PARLI.* During the Narrative segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 7, as shown in Table 30. The *Quality* of response was almost identical at *Weak* and *Good* (Table 31).
Table 30.
*Case 7: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th>Case 7</th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>143</td>
<td>.30</td>
<td>.20</td>
<td>.23</td>
<td>.15</td>
<td>.06</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>185</td>
<td>.30</td>
<td>.15</td>
<td>.12</td>
<td>.28</td>
<td>.08</td>
<td>.05</td>
<td>.02</td>
</tr>
<tr>
<td>Expository</td>
<td>165</td>
<td>.28</td>
<td>.13</td>
<td>.15</td>
<td>.32</td>
<td>.06</td>
<td>.06</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 7, with the majority being at Level 1 and Level 4, together combining for 58% of the responses, as shown in Table 31. The *Quality* of response was at the midpoint and on either immediate side 84% of the time.

Table 31.
*Case 7: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th>Case 7</th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>81</td>
<td>.01</td>
<td>.42</td>
<td>.06</td>
<td>.43</td>
<td>.01</td>
<td>.07</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>116</td>
<td>.04</td>
<td>.18</td>
<td>.18</td>
<td>.48</td>
<td>.03</td>
<td>.09</td>
</tr>
<tr>
<td>Expository</td>
<td>93</td>
<td>.08</td>
<td>.30</td>
<td>.12</td>
<td>.44</td>
<td>.03</td>
<td>.03</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the responses were split between literal and inferential levels, with 56% literal and 44% inferential (Table 30). The *Quality* of responses was also split with 50% being negative, and the other neutral to positive (Table 31).
Sub-question. 4) Does student performance on the assessments form a pattern of development?

Sandra’s scores all showed marked improvement from formative to summative. This improvement was demonstrated in a relatively smooth progression.

Summary. Sandra performed at the Proficient level on the SRI and all of the proximal measures (PARLA-Narrative, Narrative Nonfiction, and Expository). She achieved the Basic level on the GMRT-4. Her formative and summative assessments of agency and motivation remained the same at Robust, a rating that is supported by observational field notes across the course of the study. Her metacognitive scores on the MARSI survey changed only slightly, with the category of Support Strategies seeing a slight decline. By considering evidence of Sandra’s reading comprehension development across the course of the study, one can make a case for the PARLI framework as being effective with Sandra.

Case study: Holly

Previous years of schooling. Holly (Case 8) and her mother each independently shared her prolonged academic struggles throughout elementary school with the researcher, but her official records are sparse. She was reading within the grade level range through fifth grade and received reading tutoring to address apparent difficulties. At the end of fifth grade her GMRT-4 score qualified her for remedial reading at the Read 180 program level of intervention.

Holly was one of the students who started with the researcher in sixth grade. She is a polite young woman who lacks academic confidence. Shortly after the school year started, she asked for assistance with core class work, and quickly developed a habit of
staying after school with the researcher two or three days a week for extra assistance. Her gains were usually unstable, and she expressed a great deal of frustration. At the start of the study year, Holly announced to both the researcher and her social studies teacher (both UCLA alumni), that it was her intent to attend the University of California at Los Angeles on an athletic scholarship. Both teachers independently began to process the academic requirements for enrolling in this institution with Holly. It was at this time that she made a noticeable shift in actively engaging in her school work and seeing results, while trusting the gains she made as real. She also stopped referring to herself as *dumb* at about this same time.

According to district records, Holly has performed at a Basic level on the MAP test in Communication Arts for the past 3 years, with very little change (652, 652, and 663). Holly’s SRI score fluctuated a bit, but remained in the mid range of Basic/Has Partially Met Standard Lexile range throughout the study year. Holly showed an improvement in her GMRT-4 Comprehension score last year, with movement from the 4 to 5.6 grade level. She also demonstrated significant growth in GMRT-4 Vocabulary from 5.1 to 6.7. According to statements that Holly made, she did not seem to trust this growth as being real at the start of the year.

*Research questions.*

*Research question.* Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?
The PARLI framework was effective for Holly. Appendix G features the Event Flow diagram for Holly that provides a graphic representation for the data.

**Narrative formative assessment.** On the Narrative formative assessment, Holly scored in the Basic range, at Level 3, Stage B. Holly’s reflections on this assessment do not show an awareness of an error in thinking or contemplating what to do to fix errors. Reflections show *Modest* PAB. Rankings were spread across five of the rankings on the scale. In order of frequency at the literal levels they were: *Modest, Robust, Fragile,* and *Discouraged.* At inferential levels, there was one each of *Robust, Accepting,* and *Discouraged.*

There was no evidence of breadth of thinking on this first assessment. At the literal levels of comprehension, Holly’s thinking was split between *OK* and *Good,* then shifted to an *OK/Wrong* split at the inferential levels. An example from Level 3 highlights her overall score on this formative assessment: [What is different about how Harry is with his father in public than when they are home?] “In public Harry doesn’t [sic] want to be seen with him. At home Harry is happy to see him.”

**Narrative developmental work.** Evidence of metacognition was not plentiful during the unit coursework, but Level 3 work was predominantly *Good* and Level 4 predominantly *Strong.* During think-alouds in the Narrative unit, Holly made very few moves. At Level 1, she corrected her few miscues and just read it aloud, despite coaching on how to conduct the think-alouds and requests to adjust her behavior.

Level 3 annotation was the aspect that dominated the work in this unit and the *Quality* was *Good.* There was some breadth of thinking with a few explanations and one quote. At Level 4 it was again mostly annotation, with *Strong* thinking in evidence.
Holly engaged in heavy annotation, using the colors designated for Level 3 and Level 4 to highlight text, and symbols to indicate her thinking about it. In some instances, she added notes about what she figured out for Level 4.

Holly’s sense of agency during unit work was *Self-Doubting*, moving toward *Vulnerable*. She resisted using the Student Guide and was often confused as a result. She was not engaged with the framework at this time.

*Narrative summative assessment.* On the Narrative summative assessment, Holly scored again in the Basic range with a score of Level 3, Stage B. Her sense of agency was *Fragile* at the outset: “Before I take this assessment, I am feeling. Nervous because I really don’t want to fail and I want to do better than Papa’s Parrot.”

No stability of level of thinking was apparent in Holly’s summative assessment work. Her responses were all over the scales and more dependent on the particular question asked than the level of comprehending needed to answer the question well. Specifically, at the literal levels the *Quality* of her thinking was on the *Strong* side of the scale, while at the inferential levels it was on the *Weak* side. In conjunction with her inconsistent level of thinking, her PAB was also variable throughout assessment with *Modest, Tenacious*, and *Robust* in equal parts, and *Discouraged* and *Hopeless* also present at the literal levels. The inferential levels were almost all in the negative part of the PAB scale with *Hopeless, Self-Doubting*, and *Discouraged* in equal parts. For one instance, she showed a *Robust* ranking at Level 4.

Holly’s summative reflection was *Vulnerable*: “I think that this assessment was difficult because it had some questions that were a little challenging and it required a lot
of thinking. I think it’s ok Im [sic] a little disappointed because I’m behind and need extra
work.”

**Narrative nonfiction formative assessment.** Holly had her weakest result of the
study when she performed in the Below Basic category, scoring at Level 2, Stage C, for
this formative assessment. Holly began with a *Self-Doubting* sense of agency: “Before I
take this assessment, I am feeling I was nervous and felt like crying I’m not good with
tests. I’m always afraid I would do bad and fail.” During the course of the assessment,
she was not consistent in her PAB ratings. At Level 1, *Tenacious* and *Accepting* were the
rankings, but at Level 2 they moved up to *Robust*. Level 3 they declined to *Discouraged*
then *Hopeless*, and Level 4 was *Discouraged* and *Accepting*. Finally, at Level 5, she was
stuck at *Hopeless*.

In this assessment, Holly showed no evidence of breadth of thought. Her thinking
was predominantly of *Wrong Quality*. At Level 1 it was split *OK/Wrong*, while Level 2
was all *Good*. From Level 3 on, her thinking was all classified as *Wrong*. An example
from Level 2 shows the highpoint of her work at this time when she responded to a
question asking her to define/explain what the Great Depression was: “GD was a servere
[sic], worldwide, economic crisis during the 1930s in which millions of people lost their
jobs, their savings, and their homes.”

Ultimately, Holly expressed a Vulnerable sense of agency: “I think this
assessment was appropriately challenging because I really struggled with certain
questions. I’m not happy at all I really would like to be at a Level 4.”

**Narrative nonfiction developmental work.** During think-alouds limited
metacognition was evident, with miscues without corrections making up the majority of
Holly’s moves. There were random statements like “I’m hungry. I want . . . lollipops” sprinkled throughout the think-alouds. In addition, there was evidence of some connecting with obvious intent to bridge understanding of text at Level 4. An example of this came from one of the think-alouds later in the unit: “This is starting to remind me of when my dad was talking about when me and my dad were watching a movie about the Cheriyokees [sic].”

During the bulk of the coursework of this unit Holly worked starting at the lowest Levels, based on her initial assessment. She worked her way up to Level 4. She did her finest work during the discussions, where her weakest Level 4 interaction was at the OK Quality level, and most of them were considered Good or Strong. She worked predominantly at Level 4 and took a leadership role in the discussion. This academic leadership was new behavior. An example of a Strong contribution to the discussion at Level 4 illustrates: “I bet that he is going to take the money then try to keep her.” She also exhibited breadth of thinking. An example of a Level 4 Question: “Why couldn’t her real mom, her first mom, step up and just go to the court to watch her child get a divorce?”

Evidence of PAB in this portion of the unit came from observational field notes. During the unit work, Holly began to increase her engagement. As she moved from Modest to Tenacious, she shifted to asking that the instructor to refrain from helping unless or until she became stuck. She started working earnestly with peers to problem solve and referenced the Student Guide to shape her work.

*Narrative nonfiction summative assessment.* Holly improved markedly from the formative to the summative assessment, moving up to the top of the Basic range with her
score of Level 4, Stage A. Holly showed a *Self-Doubting* sense of agency at the start:

“Before I take this assessment, I am feeling a little anxious and I am freaken [sic] out.”

Her answers at the literal levels of comprehension were on the stronger side (*Good/OK/Strong*), while those at the inferential were in the weaker categories. Level 5 provided an example to illustrate this weaker work:

[Why might it be a good idea for President Obama to study the lessons of Roosevelt’s presidency? Be sure to use evidence from the essay to help you explain your opinion.] It might be a good idea for president Obama to study the lessons of FDR’s presidency because FDR is a good person who wanted to make the nation a better place and Obama hasn’t really been doing that he’s just been helping with the war and making things so faster. President FDR helped other races and made groups to help other races by if they don’t have a home he’s having people build them a home so that they can live president Obama doesn’t do that i don’t think but he should of he does.

Over the course of the assessment, Holly’s *PABs* were unstable, and displayed a disconnect between her rankings and her written reflections. For example, at Level 4 her rankings were *Robust* and *Discouraged*, but the written reflection was *Modest* at an *OK* level of thought: “my answer was different because in the 2 point answer it had a lot more explaining than just a couple sentences.” Similarly, her Level 5 rankings were *Vulnerable*, but written reflection was *Robust*: “I did the great things that FDR did and what Obama should do.” Her final reflection showed a *Vulnerable* sense of agency overall:

I think that this assessment was easy because I had already new [sic] a little about him and I Read the passage once and then everytime [sic] I saw an [sic] question it would be in the passage. A little upset want to be at a higher level.

*Expository formative assessment*. On the Expository formative assessment, Holly scored in the Basic range, at Level 3, Stage B. As she did previously, she began this assessment with a *Self-Doubting* reflection: “Before I take this assessment, I am feeling
scared because it’s a Monday and Im[sic] a little tierd [sic] and social studies Im not so
great at.” Her PABs were variable, with 6 of 10 possible ranks represented at the literal
level. Tenacious and Accepting were identified twice each, and Robust, Vulnerable,
Discouraged, and Hopeless, once each. At the inferential level just Discouraged and
Vulnerable were present.

Throughout this assessment, Holly’s responses showed no breadth of thought
while performing with variability in Quality through all levels of thinking. She displayed
evidence of growing metacognitive awareness, even though her thinking was often faulty.
For literal levels she showed thinking that was Wrong, OK, and Good. At inferential
levels her thinking was classified as Wrong/Weak, Good, and OK. An example of Wrong
at Level 4 illustrates: [We know that the tablets found in Mesopotamia were from schools
because . . . ] “We know this because in the essay it says that we know that Learning to
write was’t [sic] easy.” There was evidence of Holly trying to think things through,
though, demonstrating a shift toward greater metacognition, even if she was pursuing an
inaccurate line:

But I think that when they were learning to write a lot of the letters and words
were hard to spell out so then they wouldn’t be able to pernounce [sic] or say
things write. Maybe letters were hard to write on tablets to.

She closed out the work on this assessment with an Antagonistic reflection:

I think that this assessment was difficult because . . . Im [sic] not veary [sic] good
a social studies and I thinks its really boring for me so I’ve been working on
trying to raise my 4A to a 4b or 4C I’d be happy with either one. Not to happy
because I wanted to get a 4b but I dropped two levels that’s not OK to me and
next time I will definetly [sic]rock the post!

Expository developmental work. During think-alouds Holly made an increase in
connections, although predominantly weak ones, from previous units. Specifically, she
did some connecting of ideas in text to other exposures in history class using metacognitive strategies. Holly’s Level 1 moves during think-aloud featured low level thinking included one omission and two weak connections to the text, but the majority were simple miscues that went unacknowledged. Level 2 and 3 included several weak connections between self and text. She showed evidence of metacognition, but this awareness did not prove to be very effective. During the think-alouds, she shared a Vulnerable reflection connecting to one of the texts: “I think that this may have helped me because in social studies I’m really struggling with this year.”

For the remainder of the work of the unit Level 4 was the main focus. Work done here was up to Stage C, writing questions at this level, as well as supporting thinking with quotes from the text. Holly’s collaborative work done with Sandra featured a breadth of thought that showed significant metacognitive development. This example of Level 3 work of Good quality was illustrative:

Title--The records were never found. Question--Why don’t we know much about Confucius’s wife? Answer--We don’t know much about Confucius’s wife. Because in China many other ancient societies take place in home. A lot of women did not as a rule, play a visible role in the public life of ancient China so they were not in any public records but their husbands were. Quote--“During these years, Confucius married. His wife gave birth to a son and a daughter but we know nothing else about her, not even her name. We don’t know how long they were married, if they separated, or if she died. Women did not role a visible in the public life, and while they were influential and behind the scenes very often, they were not always mentioned in historical records.” Explanation--We know this because in the book it said that they didn’t have any record of the wives not even alive.

Observational field notes placed PABs fluctuating in the positive range between Robust and Modest, with the occasional shift to Tenacious as Holly worked collaboratively with Sandra to master reading and thinking about expository texts. Holly
was determined to develop her proficiency and was engaged and focused on all but one day.

*Expository summative assessment.* On the final proximal measure of the study, Holly once again performed in the Proficient range earning a score of Level 5, Stage A. She began this assessment with a *Fragile* sense of agency: “Before I take this assessment, I am feeling nervous because I don’t want to do bad but I do want to be on a heigher [sic] level.” An example at Level 4 shows her strength:

[How is the time of the Great Migration similar to what is going on in the world today?] *Answer*--The great migration is similar to what is going on in the world today because today our economy is starting to get bad so a lot of people are looseing [sic] their jobs and loosing [sic] their homes and money so a lot of people don’t have the minds to get back out there a look for jobs so they can get back on their feet. *Explanation*--Back then it said that another thing that motivated the Great Migration was that in 1837 the United States was in an economic depression. In fact, during this time thousands of people in New York, Baltimore, and Philadelphia were without work. Banks closed so people lost their savings. Many farmers have borrowed money to purchase seeds and equipment to plant crops, but because most people were broke, the crops did not sell for enough to pay the bills, and they lost everything to.

At Level 5, Holly’s thinking was *OK* and *Wrong*, resulting in an emerging Level 5 score. An example from Level 5 demonstrated further growth in metacognition as she moved into an emergent status with complex inferential comprehension:

[How were the missionaries and the mountain men similar? Use examples from throughout the essay to support your opinion.] *Answer*--The missionaries and the mountain men are similar because of the way they lived. *Explanation*--The Mountain Men are people who learned how to understand land and how to grow crops and live on their own and to learn to hunt with no food or water. Missionaries are people who traveled on their own and new land well and new [sic] how to grow food and make a living in the story it said that the missionaries had to learn and move in groups as for the mountain men they traveled in groups they just do everything on their own.

Holly’s *Robust* reflection following the assessment was appropriate with her performance and illustrates her growth in her sense of her own agency over the course of
EVALUATING THE PARLI FRAMEWORK

the study: “I think that this assessment was easy because I Really tried this time and I was easy because latly [sic] its been easy for me to do.”

**Sub-question. 1)** Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

**GMRT-4 and SRI.** Holly’s performance at the Proficient level on the GMRT-4 is evidence that she closed the gap. She demonstrated a growth of 26 months and an increase of 9 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Holly achieved **Proficiency** with her score on the final measure of the year. Holly’s SRI score improved by 72 Lexile points, placing her above the average performance of 51 for her counterparts in the quintain, but still in the Basic category on this assessment. This showed a failure to close the gap.

**PARLA.** Holly made gains in all three content areas, as measured by the PARLA proximal assessments. She achieved grade level expectations in the final, Expository unit, providing evidence that she closed the gap between herself and her non-struggling grade-level peers. Specifically, on the PARLA-NARR, she improved by 1 stage from formative to summative assessment, placing her just below the average growth of the members of the quintain and ending at Level 3, Stage B. On the PARLA-NARR NF, she improved by 4 stages, right at the average of just under 4 stages for the rest of the quintain and just below grade level, at Level 4, Stage A. Finally, on the PARLA-EXPO, she showed her greatest improvement relative to her immediate peers. She achieved 5 stages of growth
relative to the quintain average of near 3 stages. She scored at the beginning stage of demonstrating understanding of the complex inferential understanding. At this stage the reader understands the relationship between and among multiple inferences across time and texts (Level 5, Stage A).

Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Holly’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Modest to the Vulnerable category (see Table 4. The MST Taxonomy of Personal Agency Beliefs). For Holly, Modest was the most prevalent PAB rating by a sizeable margin, with Robust, and Accepting in comparison to the other categories as being the next most frequent, as shown in the graph in Figure 20. Furthermore, Holly also showed Personal Agency Beliefs in the lower categories of Discouraged and Hopeless with some frequency.

Continuing with her habit of candor, on the initial MAASCM it was surprising when Holly rated herself as Modest, as observations suggested that Vulnerable might be more appropriate, but her reflection supported her rating: “Because I do believe that the environment is OK and I do have trouble thinking that I’m not up to the challenges I face.” The variability of her Personal Agency Belief ratings throughout the study suggest that she was not reflexively choosing options that would reflect favorably upon her, but truly trying to explore where she was in each instance. This was also reflected in the observations of Holly while she was working, within the Contexts of both solitary and collaborative work, in both dyads and small groups. Holly would freely acknowledge when she was not tracking with the work at hand and sought help from peers first, then
adults. Holly’s post rating on the MARSI reflected the gains she made throughout the semester, moving up to the middle of all three scales.

Figure 22. Distribution of most prevalent PABs in Case 8.

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied as shown in Figure 22, with Accepting being the most consistent. The Personal Agency Belief of Modest was at the median level of 28% at the start of the study, rose during Narrative to 39%, and came back down to 21% in the final unit. The Tenacious Belief was stable for the first two units, at 18%, then doubled during the final (Expository) unit to 36%, as shown in the graph in Figure 23. Modest is the leading category of Personal Agency Beliefs for Holly by a wide margin, followed by the Robust category. In the Modest category at the literal levels, responses were predominantly in Narrative, with three Narrative Nonfiction and one Expository as well. When the inferential level of understanding was considered, there was an almost even distribution with two at Narrative, three at Narrative Nonfiction, and three at Expository. Holly also had one instance of advanced levels of comprehension expressed during the Narrative unit. Her next category is Robust, with six Narrative and five Narrative Nonfiction responses at the literal levels. Narrative Nonfiction featured
more inferential thinking, with seven total reflections broken out as two in Narrative, four in Narrative Nonfiction and one in Expository. When exploring Holly’s *Accepting* ratings, only three of the 16 are from evaluations of specific reading comprehension work done, one from each of the content areas. The majority of these ratings come from Holly’s self assessments. By contrast, all 13 of the *Discouraged* ratings come from work products. They are highest during Narrative (six events), drop to a low of three for Narrative Nonfiction, and edge up to four in Expository. Holly’s final category is *Hopeless*, with 10 reflections in total, split between literal and inferential thinking, and spread across the content areas with three in Narrative, five in Narrative Nonfiction, and two in Expository.

![Case 8 graphs](image)

*Figure 23.* Case 8: Tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.
Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?

**Marsi.** On the Metacognitive Awareness of Reading Strategies Inventory (Marsi), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Holly’s self-reporting indicated low ratings across all three categories of reading strategies to begin with and the midrange from the post measures. Holly’s greatest reported changes in metacognition from the Marsi were in the area of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting, and similar strategies; Holly reported an increase from 2.3 to 3.25 over the study period. For the subscale of Global Reading Strategies, she indicated a slightly positive change from a rating of 2.2 at the start of the study to 3.0 at its conclusion. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth. Similarly, Holly’s self-rating improved from 2.4 to 2.9 for Support Reading Strategies, these include a variety of actions such as note taking, paraphrasing, and discussing the material, among others.

On the Marsi, as the only student to rate herself consistently on the Low end of the scale at the start, Holly was most candid about not using strategies (Table 32). She did so while missing the connection that *not* using these strategies with anything might indicate that she should attempt to use some of them to improve her academic competence. She very much wished to become a stronger student, and has aspirations to attend a top-tier university on a sports scholarship, but she still was failing to *connect the dots* on some of the things that would move her along in this pursuit.
Table 32.  
*Reflections on MARSI survey by reading strategy for Case 8.*

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I don’t really do these things I just read usually.”</td>
<td>“I don’t really think about doing them.”</td>
<td>“I don’t think of doing them.”</td>
</tr>
</tbody>
</table>

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. However, the majority of the moves in Holly’s think-alouds were at Level 1, but other levels were also well represented, as shown in Table 33. Overall, Holly made 54% of her moves at Level 1, and 32% at Level 4 or Level 5.

Table 33.  
*Frequencies of Levels of think-aloud protocols for Case 8.*

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.50</td>
<td>0</td>
<td>.25</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.44</td>
<td>.06</td>
<td>0</td>
<td>.38</td>
<td>.12</td>
</tr>
<tr>
<td>Expository</td>
<td>.65</td>
<td>0</td>
<td>.06</td>
<td>.12</td>
<td>.17</td>
</tr>
</tbody>
</table>

*PARLA/PARLI.* During the Narrative segment of the PARLI framework, the levels of responses and the *Quality* ratings of them were distributed across levels, as shown in Tables 34. and 35. respectively.
Table 34.
Case 8: Frequencies of Levels present in work by unit.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>109</td>
<td>.14</td>
<td>.17</td>
<td>.32</td>
<td>.21</td>
<td>.09</td>
<td>.02</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>135</td>
<td>.19</td>
<td>.20</td>
<td>.25</td>
<td>.27</td>
<td>.09</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>107</td>
<td>.25</td>
<td>.10</td>
<td>.20</td>
<td>.39</td>
<td>.06</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 5, with some clustering at Level 3 and Level 4 (Table 34, p. 270), while Quality ratings were also spread across levels, with some clustering at the Good and OK levels, as shown in Table 35.

Table 35.
Case 8: Frequencies of Quality present in work by unit.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>84</td>
<td>.08</td>
<td>.30</td>
<td>.31</td>
<td>.17</td>
<td>.02</td>
<td>.12</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>110</td>
<td>.07</td>
<td>.25</td>
<td>.34</td>
<td>.18</td>
<td>.03</td>
<td>.13</td>
</tr>
<tr>
<td>Expository</td>
<td>116</td>
<td>.03</td>
<td>.20</td>
<td>.41</td>
<td>.24</td>
<td>.03</td>
<td>.09</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of responses were similarly distributed across levels of comprehension, with Level 4 being the level with the most responses (Table 34). When it came to Quality during Expository, Holly’s work was also distributed across categories, with the biggest category being OK, followed by Good and Weak in almost equal measure, as shown in Table 35.
Sub-question. 4) Does student performance on the assessments form a pattern of development?

Holly’s scores have improved over time on all quantitative measures, with her PARLA scores showing variability when moving into a new content area, suggesting the sensitivity of the proximal PARLA instrument to the instructional intervention. This fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process.

Summary. Holly performed at the Proficient level on both the GMRT-4 and the final, and most challenging, proximal measure (PARLA-EXPO). Her score on the final SRI remained in the Basic category. Her formative and summative assessments of agency and motivation show change from a starting place near the top of the scale with Modest, to one near middle of the scale of Vulnerable. What was true with most members of the quintain regarding the moderate reliability of these self-ratings relative to the preponderance of the evidence throughout the study coursework, was particularly true with Holly, and field notes at the end of the study indicate her complaint about having to do the survey again along with an apology to the instructor/researcher for this complaint. Accordingly, more weight was given to her reflections and the sense of agency observed along the way than these survey measures. Her metacognitive scores on the MARSI survey all improved from low to medium usage from pre-study to post-study. Examination of the entirety of the evidence leads one to conclude that the PARLI framework was effective with Holly.
Case study: Bella

Previous years of schooling. Bella’s (Case 9) elementary records indicate that she was reading on grade level in the third grade, but began to lose ground during fourth grade. In fifth grade she participated in reading tutoring. Her GMRT-4 scores at the end of fifth grade qualified her for reading intervention at the middle school level, and further evaluation indicated that she would benefit from direct instruction in decoding to fill gaps in her skills that were slowing her down academically.

Bella started sixth grade with decoding issues. She graduated to the remedial reading intervention level of the Read180 program in seventh grade. She was intellectually curious and tenacious. She was driven to be successful and correlated her successes with hard work. According to school records, she became very anxious during high stakes tests, and her results were generally lower than her demonstrated performance in day-to-day work. Bella was one of the more reflective students. She was an avid reader, despite her challenges with comprehension. Bella has performed at a Basic level on the MAP test in Communication Arts for the past three years, with a slight change in performance, but no significant differences (637, 684, 673).

Research questions.

Research question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Bella. Appendix G features the Event Flow diagram for Bella that provides a graphic representation for the data.
Narrative formative assessment. On the Narrative formative assessment, Bella scored at the top of the Basic range, at Level 4, Stage A. Bella attempted to show understanding at all levels, but was successful only to the level of beginning inference. Her rankings are all Robust at the literal levels, mixed between Robust, Accepting, and Vulnerable at the inferential levels, and Accepting at the advanced level (Level 6). The only aspect of thinking represented in this assessment was answering, showing no breadth of thought at this time. Evidence of metacognition was Good to OK at the concrete level, Good to Weak at the inferential, and Wrong at the advanced levels of comprehension.

An example of a Good response at Level 4 shows Bella’s most effective thinking on the assessment: [Why does Mr. Tillian bring Rocky to the shop?] “So he has company and there is something there to talk to.”

Questions at Level 5 called for a response that was at least one paragraph in length and contained evidence from the beginning, middle and end of the story. A response from this level shows work that was found to be Weak:

[What changes does Harry undergo in the story? Give three examples from the story to explain Harry’s attitudes at each stage of the story and how they change.] Well at the beggin [sic] he doesn’t mind being with his father and his attitude is nice and doesn’t mind, in the middle he starts to pull away from his father and doesn’t need him anymore. Attitude is he doesn’t really care and at the end he is worried and frustrated with bird.

At Level 6, paragraph-length responses were also expected. They required some depth of thought about the world outside of the text and the combining of multiple smaller inferences. This example demonstrated Bella’s attempt at this range of thinking that was not correct:
[What point does the story make about the influences on the relationships between loved ones? Explain your idea using evidence from the story. Use at least 2 details and/or examples from the story to support your answer.] That you are always there for them even when you don’t seem like it and you care about them and do what they need done when they are worried about it.

Narrative developmental work. During the think-alouds in the Narrative unit, Bella made predominantly Level 1 moves, some at Level 2, only two at Level 3, and 10 at Level 4. Level 4 was the second largest category. Level 1 was dominated by Weak Quality to indicate miscue; but, she was one of the few members of the quintain to be obviously aware of many of the miscues and made attempts to correct them. One such example of this metacognition is her repeated struggles with the word inadvertently that ended with this move: “invermeterly . . . in . . . I don’t know, it’s that same word.” Her Level 2 moves were mostly OK quality wherein she made connections to key details. These connections are deliberate, showing metacognitive control: “I would respect him too if he’s been there that long and is that nice.” Bella’s Level 4 moves were almost exclusively evaluated to be a Good quality of thinking: “I guess his dad was there before he was, so his dad probably gave him some tips.”

For the remainder of the coursework for this unit Bella worked exclusively at Levels 3 and 4. Work across this unit showed depth of thought, with a variety of aspects represented. The quality of the responses was predominantly Good for Level 3, with some instances of OK. It was Good/Strong at Level 4, with annotation being the dominant aspect of metacognition. An example from work at Level 3 shows the breadth and quality of the work done during the unit at this depth of understanding:

Answer--The important relationship in the story was Joanna and the leaf because when the tree died she died. Explanation--I thought it was dumb that she thought when the tree died she did I don't think she compar [sic] yourself to a tree or another object. Quote--“And that chance is for her to want to live. In cases like
this, the will to survive is everything. But I'm afraid that Joanna has made up her mind that she's not going to get well.” *Question*--How does Joanna get better in the story?

Similarly, illustration from Level 4 work was also illuminating. During the section of the story when Joanna says that she is tired and ready to “turn loose my hold on everything,” Bella annotates the key parts in which the characters are speaking obliquely in pink to indicate basic inference needed, then annotates with her understanding: “she really wants to die.”

Throughout this work, her *PAB* was variable, being *Robust* on rating scales and *Robust* to *Tenacious* from observation, *Modest* to *Self-Doubting* on verbal reflections. When observed, Bella took the lead readily in launching collaborative work, sought clarification and feedback to keep momentum, but rarely stopped work while waiting for the instructor, and had a generally high level of engagement. She enthusiastically embraced annotation as a learning tool.

*Narrative summative assessment.* Bella’s performance on the summative for the Narrative unit dropped one tier in the Basic range to a score of Level 3, Stage C. She began with a *Modest* sense of agency: “Before I take this assessment, I am feeling I feel ok like I understand what I need to do.” She exhibited a breadth of thought in keeping with her work for this unit at Level 3, but not at Level 4. She earned positive *Quality* ratings at literal levels only (*Good/OK*) and her thinking was classified as *Wrong* at both the inferential and advanced levels. An example of her work at Level 3 shows the high end of her work on this assessment:

[Why does Roger do what he does in the opening lines of the story?] He steals Mrs. Jones pocketbook so that he can by himself a new pair blue suede shoes. Does he need them that badly he already has tennis shoes.[sic] *Quote*--“I wanted a pair of blue suede shoes”.'
An example of ineffective work at the inferential level was also typical: [Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story?] “He knows that she knows his name and he could not carry the purse.”

Bella’s PAB rankings showed metacognitive awareness of her own limitations in comprehending text. She skipped Level 6, commenting that she did not understand the questions, but then attempted Level 7. At the literal levels, her rankings were Robust through Level 3, where one was Discouraged. When she moved to the inferential level Tenacious dominated, but there was also one instance of Discouraged. At the advanced levels she was Discouraged.

Bella’s overall reflection on her performance on this assessment was Vulnerable: “I think this assessment was appropriately challenging because [sic] when the levels get hard I don’t understand the question they are asking me. I understand some but not the whole thing.”

**Narrative nonfiction formative assessment.** Bella’s initial reflection was Vulnerable: “Before I take this assessment, I am feeling a little scared because I don’t know how hard it is going to be.” No breadth of thought was evident in this assessment. Thinking was predominantly of OK and Weak Quality at Levels 1 and 2, and Weak/Wrong at Level 3, moving to only Wrong at inferential.

Level 3 had evidence of metacognition in reflection: “I need to look at the sidebar.” A Wrong example from Level 3 shows Bella’s struggle on this assessment: [In an essay written when she was 14, why did Eleanor say that it is easier to have no ambition?] “Because it would be the same thing every day and do nothing.”
Unlike the Narrative unit, no advanced thinking (Level 6 and 7) was attempted. Ranking was *Tenacious* over all levels of the hierarchy, with some instances of *Robust* at Level 2 only. Written reflection was present only at literal levels and was *Modest/Weak* in nature at Levels 1 and 2, and *Modest/OK* at Level 3. Her final reflection was *Accepting*: “I think this assessment was appropriately challenging because it was way different than fiction. I think I could have done better but I will get better as I go on.”

*Narrative nonfiction developmental work.* During the think-alouds Bella made moves at Levels 1 through 5, with Level 1 dominant, Level 2 the second largest category, and the others equivalent in frequency. Think-aloud *Quality* overall was *OK*. Level 1 *Weak* moves (miscues) are prevalent, with two instances of Bella commenting on her miscue, showing metacognition. Level 2 through 5 are all *OK* and present about equally.

During the remainder of the coursework for the Narrative Nonfiction unit, *OK* quality dominated, but *Good* was well represented and breadth of thinking is present. Bella was a minimum of four times as likely to work at Level 4 as any other level. Level 4 also had solid representation of *Good* thinking, with this being the second largest category after *OK* at Level 4 across all levels. An example of breadth of thinking was found at Level 4:

*Question*--What does the family having two televisions tell us, as readers, about this family? *Answer*--The reason the character is at these people house it that she is staying there while she is trying to get a divorce and the judge doesn’t want her to go back to her house where her parents and her husband can get to her. The reason why they have two TVs is because they have a lot of money and they can afford it and all the other luxury items that they own at there [sic] house. *Explanation*--In my opinion is that I think that they are lucky that they get to have those things while Nujood’s family has nothing. *Quote*-- “There are two televisions in this house-what luxury! (Pg. 47)”.

Observations of class work, including discussions, made metacognition apparent,
as Bella most often shared her thinking aloud with either her collaborator, the instructor, or as a contributor to a larger discussion. She engaged in attempts to systematically build from simpler understandings to make connections to get to the advanced understandings of the highest levels of comprehension.

Bella exhibited variable PAB during the work of this unit. Robust, Modest, and Self-Doubting in written reflection, while Robust and Tenacious based on observation field notes. She was engaged and exhibited leadership with collaborative work. Likewise, she was engaged and animated during book discussion, often challenging the thinking of others in addition to offering her insights.

_Narrative nonfiction summative assessment_. Bella achieved the Proficient level of understanding with a score of Level 4, Stage B. Bella started off with a Robust anticipation of this assessment: “Before I take this assessment, I am feeling like I understand everything that I need to know.” This assessment features few written reflections by Bella. There were two responses each at Level 4 and Level 5, with all responses being at the Modest rank for personal agency. A Level 4 an example highlighted the nature of these reflections: “I need to put a quote and more details.” While her reflections were all of OK quality, regardless of Level, there is a breadth of thinking demonstrated at all levels, from Level 1 through Level 5. OK was the most prevalent Quality category found across all Levels. The PAB rankings were predominantly Robust, with one instance of Vulnerable at Level 4.

In addition, to Bella’s performance at Level 4, she started to show skill at Level 5, as the following example shows:

[Why did people have so much faith in FDR? Be sure to use evidence from the essay to help you explain your opinion.] _Answer_--A lot of people had faith in him
because he was liked by everyone and he was helping everyone and did what was right for the people in America like in the TREA he helped people that didn’t have jobs get them and was in a lot of organizations like CCC CWA WPA. Explanation--I think he was a good man and did a lot of good things and I can see everyone liked him. Quote--“Among them were three programs to help employ workers: the CCC CWA WPA”.

Bella’s culminating reflection was Tenacious, acknowledging her success, but still striving to do better: “I think that this assessment was easy because I understood what was going on. I think I did ok wish I did a little better.”

Expository formative assessment. On the Expository formative assessment, Bella scored in the Proficient range, at Level 4, Stage B. Bella expressed a Vulnerable sense of agency before taking the assessment: “Before I take this assessment, I am feeling that I do not know what it is going to be is it going to be hard or easy.” Breadth of thought remained present (as it had in earlier units of study), but she made no effort to engage in higher level thinking beyond basic inference. This was the first instance in the study when Bella did not attempt to reach beyond her current comfort zone. Good is the most prevalent quality of thought across Levels 1 through 4, while written reflecting was absent. Her PAB rankings were dominated by Robust, with one Tenacious at Level 3 and one Fragile at Level 4. An example from Level 4 illustrates Bella’s best work on this assessment, calling into question the decision to go no further from this point:

[We know that the tablets found in Mesopotamia were from schools because . . . Model--the tablets had writing of experts, thought to be teachers, on one side. On the other side was space for students to practice. We think this was so because one side had work that was full of errors. We think that they were from schools because there were many of them grouped together.] Answer--The tablets had a line that went down the middle and one had the teachers writing and the other side had the kids writing so they could tell that a kid was practicing to write on the tablet and that maybe they were having a hard time since they said it was hard to write on. Explanation--I infer that they could tell was because there was a good side which would be the teachers since they have had a lot of practice and a bad side which would be kids since they are just learning. Quote--“We know that
learning to write was not easy. Among the artifact of the Sumerians are thousands of exercise tablet from schools. On one side, the teacher wrote the lesson. On the other side, the students practiced”.

Bella’s *Modest* reflection at the end connects her results to the work done during the unit: “I think that this assessment was easy because it is basely what we have learned. I think I did pretty good for the first time.”

*Expository developmental work*. Think-alouds featured active work to comprehend the text at all Levels. Specifically, moves at Level 1 were dominated by *Weak* thinking, but Bella was aware of these miscues, and in several cases corrected them (*OK* thinking). In addition, she made active connections several times as well. At Level 2, all of her moves were representative of *OK* thinking. There was a mix between connecting and active comprehension work in which she was metacognitively in control. An example of this connecting was found in one of the later think-alouds: “Oh my gosh! That would be sooo scary!” Likewise, this think-aloud contained this example of a move that involved active comprehension work: “I’m just thinking that from our state capital that would be soooo long, so far, because here to Jefferson is like two hours by bus so that’s gotta be a really long to watch.” Finally, Level 4 moves during the think-alouds were examples of *OK* thinking most of the time. The following is an example of active comprehension at Level 4: “so I’m just trying to imagine how far that is from each other, I mean it wouldn’t be bad if it’s a couple of miles but I mean, like if it’s a lot of miles, that would be really bad.”

Bella collaborated with Edward (Case 3) to complete the unit coursework. They worked exclusively at Levels 3 and 4 only, with Level 3 being where most of the work was focused. Level 3 was all positive quality and predominantly represented with *Strong*
quality of thought, followed by Good and OK. Level 4 was equally divided between Good and Weak, with one instance each of Strong and OK, as well.

The rest of the unit continued with the same breadth of response trend from earlier work. The following Level 3 example shows the best work of Bella and Edward:

*Answer*--The narrator claims that the Big Bull Market happened because of the stocks. The stock market was going up and up so they decided to call it the Big Bull Market. *Explanation*--The cause of the stocks going up and up was because the market was going so well that they wanted to make a name for it because they thought the good times would never end. Because of our economy right now being a disaster it would be great to know that the economy by naming it at the time. *Quote*--“Investors called it the ‘Big Bull Market,’” and it seemed too many that the good times would never end. Stock prices went up and up and up, until those prices were much greater than many companies were worth. That’s when the stock market started to act strangely. Prices dipped suddenly, only to zoom up again on an economic roller coaster ride” (p. 6). *Question*--Why and how did the people call it the “Big Bull Market”?

At Level 4 the quality was often mixed within one response. The following example features a Good answer, a Weak explanation, and a Strong quote:

*Answer*--The reason the character did this was because he couldn’t support his family and help them with what they needed to live. *Explanation*--We would try to support our family to help out by trying to get as much money as we could. *Quote*--“... but for me the low point of the depression will always be the sight of my father that day, crying in the coal. In an era when a father's pride in self-respect depended on his traditional role as the family’s breadwinner” (pg. 4).

Outside of think-alouds, PAB observational field notes show Modest PAB. There were frequent remarks that Expository was more difficult and the need for the instructor to offer more reminders to complete work were indicative of lower levels of engagement.

*Expository summative assessment.* Bella performed again in the Proficient range on the final Expository assessment, achieving a score of Level 5, Stage A. Her initial Robust reflection was not surprising, given her performance on the Expository formative assessment and her work in the unit: “Before I take this assessment, I am feeling I think I
will do fine and do better then last time.” At the literal level her sense of agency was
described by rankings that were mostly Robust, with one at Vulnerable at Level 3. At the
inferential level, rankings were Robust to Tenacious.

Bella made no attempts to engage in advanced thinking on this assessment and
exhibited less breadth of thinking than on the Expository formative assessment, with little
shown prior to Level 5. Her thinking was dominated by OK quality thought at literal
levels, in addition to equal instances of Good and Weak as well. Her thinking was
categorized as Good/OK at Level 4, but all Weak or Wrong at Level 5.

Bella’s Level 5 responses were Weak or Wrong, but in conjunction with the
consistently effective work at Level 4, she earned a Level 5 score overall. The following
example is a Level 5 response that was Weak. In it she used quotes to support her
thinking, but omitted much evidence of actual thinking, leaving a group of quotes as an
answer:

[How were the missionaries and the mountain men similar? Use examples from
throughout the essay to support your opinion. Model--They inspired others to
move West.] They were both always outside, like the mountain men live outside. I
know this because it says this “These adventurous men hiked through the forest,
trapped animals and living off of the land.” Also we know the missionaries were
outside a lot since they helped people travel around “Also during the early 1800s
missionaries traveled to Oregon Country to teach the Christian religion to Native
Americans.” I think that it would really stink to have to live outside all the time
since you never know what the weather is going to be like.

Bella’s honest reflection after taking this assessment was Tenacious as she
acknowledged both the difficulty and the victory: “I think this assessment was
appropriately challenging because parts over it were hard and I couldn’t understand the
question and how they wanted me to answer it. I think I did really well and I think it is a
good point where I should be.”
Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

**GMRT-4 and SRI.** Bella began the study without evidence of a gap earning Proficient scores on both the GMRT-4 and the SRI. This was confounded by evidence from regular coursework in her core classes that was frequently above her independent comprehension level. Bella demonstrated a growth of 17 months and an increase of 2 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Bella achieved Proficiency with her score on the final measure of the year. Bella’s SRI score dropped by 26 Lexile points, placing her below the average performance of 51 for her counterparts in the quintain; however, her scores were all in the middle of the proficiency range and some movement within the range as a function of being a variable human can be expected, so this slight decrease with the proficient range was not cause for concern.

**PARLA.** Bella made gains in two of the three content areas, as measured by the PARLA proximal assessments, and achieved grade level expectations in both Narrative Nonfiction and Expository units, providing evidence that she closed the gap. Specifically, on the PARLA-NARR she declined by one stage from formative to summative assessment, placing her below the average growth of the members of the quintain and ending at Level 3, Stage C. On the PARLA-NARR NF, she improved by 6 stages, exceeding the average of just under 4 stages for the rest of the quintain and at grade level,
at Level 4, Stage B. Finally, on the PARLA-EXPO, she showed her strongest growth relative to grade level, achieving 2 stages of growth at the average of near 3 stages, and ending at the midrange demonstrating understanding of the complex inferential understanding wherein the reader understands the relationship between and among multiple inferences across time and texts (Level 5, Stage B).

Sub-question 2) Do students participating in PARLI report a shift in agency and motivation?

Bella’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Robust to the Modest category (see Table 5. The MST Taxonomy of Personal Agency Beliefs). For Bella, Modest was the most prevalent PAB rating by a sizeable margin, with Robust, and Tenacious being the other categories of some prevalence, as shown in the graph in Figure 24.

![Figure 24 Distribution of most prevalent PABs in Case 9.](image)

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied as shown in Figure 25. with Robust being the most consistent. The Personal Agency Belief of Modest was at
39% at the start of the study, dropped during Narrative Nonfiction to 33%, and decreased further to 22% in the final unit. The Robust category went from 20% during Narrative, to 18% in Narrative Nonfiction, and ending at 23% during Expository. The Tenacious belief was the most variable, changing from 21% at the start, to 47% during Narrative Nonfiction, and decreasing to 16% in Expository, as shown in the graph in Figure 25. Robust is Bella’s top category of Personal Agency Beliefs. At the literal levels of comprehension, Bella had 23 instances of Robust ratings, six in Narrative, eight in Narrative Nonfiction, and nine in Expository. At the inferential level, her five responses were split two to three between Narrative Nonfiction and Expository. With less than half of the associations that Robust has with Bella’s reading comprehension, Modest is the next category. There was one instance at Narrative, five at Narrative Nonfiction, and two for Expository. The Modest category at the inferential level was more populated, with 13 occurrences. There were eight Narrative, one Narrative Nonfiction, and four Expository. While it represented the next category down in prevalence, Tenacious was almost half as dense as Modest. The six literal level instances were all in Narrative Nonfiction. The inferential level was equally represented with Narrative and Narrative Nonfiction each having three occurrences of Modest, and Expository having one.
Figure 25. Case 9: Tracking the most prevalent PABs across the study. This figure illustrates the variability of PABs throughout the three units of the study.

Sub-question. 3) Do students participating in PARLI demonstrate improved metacognition?

*MARSI.* On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Bella’s self-reported ratings indicated high ratings across all three categories of reading strategies at the outset with all of them declining in the post measure. Her greatest reported changes in metacognition from the MARSI were in the area of Support Reading Strategies, these include a variety of actions such as note taking, paraphrasing, and discussing the material, among others, and her score dropped from 4.6 to 3.2. Bella’s self rating for Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting and similar strategies; also declined from 5 to 3.9 over the study period, but remained in the high rating category. For the subscale of Global Reading Strategies, she
indicated a change from a rating of 4.8 at the start of the study to 3.4 at its conclusion, moving her down from a high to a medium rating. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth.

The intersection of MARS, MAASCM, and think-aloud protocols represents a more careful consideration of the measures. Bella’s reflections on two of the MARSI subscales were representative of the strategies she uses directly with her responses on the survey, as shown in Table 36.

Table 36. 
Reflections on MARSI survey by reading strategy for Case 9.

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>No comment</td>
<td>“When I don’t understand something it really helps me to read out loud the text.”</td>
<td>“When I read I take notes so I can remember what I read if I don’t understand it.”</td>
</tr>
</tbody>
</table>

Along with test scores, the shift in frequency of level of comments during the think-aloud protocols away from Level 1 and toward inference at Level 4 and Level 5 was evidence of improved metacognitive awareness.

Think-aloud protocols. The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. Bella’s moves populated the first four levels of reading comprehension across the study time, with just over half (53%) being at Level 1, and 19% at Level 4, as shown in Table 37.
Table 37. Frequencies of Levels of think-aloud protocols for Case 9.

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.65</td>
<td>.13</td>
<td>.04</td>
<td>.18</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.46</td>
<td>.32</td>
<td>.04</td>
<td>.14</td>
<td>.04</td>
</tr>
<tr>
<td>Expository</td>
<td>.48</td>
<td>.25</td>
<td>.02</td>
<td>.25</td>
<td>0</td>
</tr>
</tbody>
</table>

*PARLA/PARLI.* During the Narrative segment of the PARLI framework, the levels of responses and the *Quality* ratings were distributed across Levels, with most of them at the literal levels (Table 38), and *Weak*, as shown in Table 39.

Table 38. *Case 9: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>164</td>
<td>.50</td>
<td>.08</td>
<td>.20</td>
<td>.21</td>
<td>0</td>
<td>.01</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>194</td>
<td>.26</td>
<td>.28</td>
<td>.16</td>
<td>.23</td>
<td>.07</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>147</td>
<td>.29</td>
<td>.20</td>
<td>.16</td>
<td>.28</td>
<td>.07</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 5, with Level 1, Level 2 and Level 4, being roughly equivalent (Table 38), while *Quality* ratings were also spread across levels, and predominantly *OK* to *Weak*, as shown in Table 39.
During the Expository segment of the PARLI framework, the levels of responses were also distributed across levels of comprehension, with clustering at Level 1 and Level 4 (Table 38). When it came to Quality during Expository, Bella’s work was also distributed across categories, with the biggest category being OK, followed by Good and Weak in almost equal measure, as shown in Table 39.

Sub-question. 4) Does student performance on the assessments form a pattern of development?

Bella’s scores improved over time on all quantitative measures, with her PARLA scores showing some variability when moving into a new content area; this fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process.

Summary. Bella performed at the Proficient level on both standardized reading assessments, and all but the first of the proximal measures (PARLA-Narrative). Her formative and summative assessments of agency and motivation show slight change from a starting place of Robust to a concluding one of Modest. Her metacognitive scores on the MARS survey changed in all categories, dropping to the middle category of usage for two of the three. It is likely, given her performance across the study, that these decreases in both surveys were indicative of her having overstated her current level on the initial
survey, as she was one to carefully guard any perceived academic weakness. Given Bella’s initial proficiency on standardized reading measures, consideration of the merits of the PARLI framework for her should focus on her growth in metacognition in complex reading and responding settings that reflect regular classroom academic challenges. Her growth in responding with a breadth of complex thinking during study coursework, and her proficiency with the proximal measures, make the case for the PARLI framework being effective with Bella.

**Case study: Alice.**

*Previous years of schooling.* Throughout her elementary years, Alice (Case 10) demonstrated consistently below grade level performance in reading and received reading tutoring during most of that time. At the end of fifth grade her scores on the GMRT-4 qualified her for reading remediation at the level of the Read180 program.

Alice started sixth grade with the researcher in Read180; when she failed to progress in this program, further evaluation showed an underlying decoding issue, so she was moved to the reading intervention that focuses on decoding for her seventh grade year. The resolution of her decoding issues allowed her to grow as a reader, but she had not yet reached grade level at the end of seventh grade, and was moved up a level in reading intervention to participate in this research study. Alice has performed at a Basic level on the MAP test in Communication Arts for the past three years, with steady improvement that just missed a Proficient rating by two points at the end of seventh grade (648, 650, 678). Alice’s SRI score was at the mid range of the Basic/Partially Has Met Standard at the start of the study. Alice showed a big jump in her GMRT-4 Comprehension score last year, with movement from 4 to 5.8.
Research questions.

Research Question. Is the PARLI framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)?

The PARLI framework was effective for Alice, based in large measure on the fact that Alice performed at the Proficient level on both standardized reading assessments. Appendix G features the Event Flow diagram for Alice that provides a graphic representation for the data.

Narrative formative assessment. On the Narrative formative assessment, Alice scored in the Proficient range, at Level 5, Stage A. Observational field notes reflected her acute focus on the task at hand on this day, along with sharing her goal of a strong performance. At the literal levels of comprehension, Alice’s thought quality was predominantly classified as Good. When it came to inferential comprehension, it was mixed but mostly Good and Strong. While answers showed effective metacognition, she just gave answers with no explanation or text support, and very little reflection about her thinking process.

Throughout this assessment, her PAB rankings were Robust. Her Level 5 response showed her strengths:

[What changes does Harry undergo in the story? Give three examples from the story to explain Harry’s attitudes at each stage of the story and how they change.] 1) When Harry was little he and his friends would always stop by after school to his father’s shop. Harry’s attitude would be happy. 2) When Harry turned twelve he didn’t go [sic] to the shop as often, so his father bought a parrot to talk and Harry felt embarrassed. 3) When Mr. Tillian got very ill he fell and had to stay at the hospital so Harry said he would clean up the mess and feed the bird. When the
bird kept saying, “Where’s Harry?” it got Harry very concerned [sic]. Harry was a little sad and he learned something that his father wanted to see Harry more often.

_Narrative developmental work._ During the think-alouds in the Narrative unit, Alice made predominantly Level 1 moves. Her Level 1 moves were _Weak_. Unlike most of the rest of the quintain, many of her _Weak_ Level 1 responses were random comments. The others were uncorrected miscues. She also made some moves at an inferential depth (Level 4), making connections to her experience. However, she made the connection but did not articulate metacognition about how it influenced her comprehension: “I always used to go fishing when I was younger. Every once in a while I will now.”

For the rest of the coursework Alice worked predominantly at Levels 3 and 4. Her inability to work at Level 5 points to the unlikely scenario that her formative score was accurate. Her Level 3 responses were predominantly of _Good_ quality and represented a variety of aspects of metacognition that included explanations and using the text to support her thinking, as shown in this example:

At the beginning of the story Joanna gets pneumonia and she always stares at the window at a vine and brick houses. Every day she would look out the window watching the leaves fall. She said, “When the last leaf falls then I will die.” I think Joanna is mad for representing her life on a leaf. But, when the last leaf falls, an artist named Behrman painted a leaf on the brick wall of one of the houses and Joanna thinks that's the last leaf. The relationship is yin & yang because Joanna & the leaf go together. _Quote_-- “I thought it would surely fall during the night. I heard the wind. It will fall today, and I will die at the same time,” Joanna said (p. 20). _Question_-- What is the relationship between Joanna and the leaf?

Alice’s Level 4 responses were predominantly _OK_ with a few _Strong_, particularly when it came to annotation, where she used color-coding to indicate the level of her thinking, then added comments as well, like “Joanna wants to die today.” Observational field notes place her sense of agency vacillating between three categories of _Tenacious_, _Vulnerable_ and _Self-Doubting_ without a consistent connection between where she stood
on a given day and either the nature of the work being done, the type of grouping used, or any other instructional strategy or tactic implemented; some days were just better than others. Alice was often behind on her work and made frequent reference to struggles with focusing on the task at hand.

*Narrative summative assessment.* On the Narrative summative assessment, Alice scored again in the Proficient range, but went down by one tier to a score of Level 4, Stage B. Alice began this summative assessment with a *Modest* reflection: “Before I take this assessment, I am feeling I feel pretty good because we have been doing this for about 5 weeks.” This assessment is the only time, throughout the study, that she attempted to think at an advanced level (Level 6). At the literal levels of comprehension her performance was split between *Quality* evaluations of *OK* and *Good*. When Alice moved to the inferential level they dropped off to a split of *Weak* and *Wrong*. No attempt was made to explain answers or support them with the text, as was demonstrated during the work in the unit. An example from Level 4, which called for a paragraph-long response, showed her *Weak* work at the inferential level on this assessment: [Why does Roger not take Mrs. Jones’ purse when he has the opportunity later in the story? “Because she would have probably called the police if he did that.”

At literal levels Alice’s *PAB* rankings were predominantly *Robust*. For inferential work, she indicated slightly more *Modest* levels, but her rankings were all over the scale, including *Robust*, *Modest*, *Accepting* and *Discouraged*. When she compared her work on the assessment with the model, her sense of agency was *Vulnerable*: “I thought it was hard because the questions are more challenging.”
Narrative nonfiction formative assessment. Alice’s reflection before taking the assessment was *Modest*: “Before I take this assessment, I am feeling I feel it will be easier.” Thinking is distributed between *OK* and *Weak* at the concrete levels. This example at Level 3 showed Alice’s stronger work on the assessment:

[In an essay written when she was 14, why did Eleanor say that it is easier to have no ambition? Model-- She said it was easier to not have ambition because you won’t have to face difficulty of disappointment.] It is easier to have no ambition and just keep on the same way every day and never try to do grand and great things. It is better to be ambitious and to do something than to be unambitious [sic] and do nothing.

Throughout the assessment, Alice’s sense of agency ranged from *Tenacious* to *Robust* in reflection of scoring, but *Modest* for her written reflections. When she got to the inferential level, her sense of agency based on the comparison of her actual score with her anticipated score was *Accepting* and she did not engage in written reflection. At this level, generally *Weak* thinking was on display, as shown in this example from Level 4:

[Why was the way Eleanor chose to be First Lady so noteworthy?] Because it was a new organization created to help women develop their political power. These determined women also worked for federal aid to education, international peace, and the end to child labor.

In Alice’s final analysis of her work she reflected a *Modest* sense of agency: “I think that this assessment was easy because the questions were straight from the text.”

Narrative nonfiction developmental work. During think-alouds Level 1 moves were dominant, but some Level 4 moves were also present. Alice’s Level 1 moves were predominantly *Weak* with miscues uncorrected, but unlike during the Narrative unit think-alouds, few were random comments. In addition, some *Good Quality* moves were made. In an example of a Level 4 move Alice was able to connect to the character in the story: “I think that the boy he was happy that he lived in Kandahar because um he made
new friends and it was warm, but it was sad that they didn’t have any stuff.”

For the bulk of the unit coursework, Alice worked equally at Levels 3 and 4. Breadth of thought was shown at both literal and inferential levels. At Level 4, metacognition was evident in the expansion to include explanation and questioning in demonstrating her inferential understanding. The Quality of thought was predominantly OK. An example from Level 3 shows this breadth of metacognition beyond just the answer:

Answer and Explanation--I think kids should ask lots of questions because that’s how you learn in the world. Question--What is one thing that Nujood wants to do, but, doesn’t know how to? Quote--“We children would dash back, to the river, now swollen with water that came up to my neck”.

From field notes of observations, indications were that Alice was operating with a Vulnerable sense of agency and was “checked out” and disengaged from the work. She was careful to comply with requirements for a fixed number of contributions to each discussion, but upon achieving that minimum requirement, she was observed rarely making eye contact or otherwise acknowledging her discussion partners (Field notes Oct 27, 2010; Nov. 2, 2010; Nov. 4, 2010).

Narrative nonfiction summative assessment. On the Narrative Nonfiction summative assessment, Alice scored again in the Basic range, going up one tier with a score of Level 3, Stage B. Alice began this assessment with a Vulnerable reflection: “Before I take this assessment, I am feeling . . . feel a little scared because my sleeping habits were not going well.” The quality of thinking was variable at the concrete level and an exclusively negative (Weak, Wrong) Quality of thought at inferential levels. In addition, there was no breadth of thought that highlighted metacognitive growth as only
answers were given and no additional questioning, explaining, or quotes from the text to support thinking were used.

An example from Level 3 showed the stronger work in this assessment: [Why did FDR created the Federal Emergency Relief Act?] “To help Americans living in the absolute worst circumstances. It set aside $500 million to help these people in the poorest cities and towns across the country.” By contrast, an example from Level 4 demonstrated where her thinking was erroneous, as the inference was absent: [Why did Roosevelt tell the American people they had nothing to fear?] “Because he believed that the only thing they had to fear was fear itself.” Her final reflection matches her performance, remaining Vulnerable: “I think that this assessment was difficult because it was nonfiction & I’m not that good at nonfiction test.”

*Expository formative assessment.* Alice scored in the Basic range with a score of Level 3, Stage C. Her reflection before the assessment was Vulnerable: “Before I take this assessment, I am feeling . . . little scared because I’m not the best at history.”

During the assessment, she engaged in minimal reflecting that illustrated that metacognition was present. Where present it was Weak in Quality. At the literal level confidence ratings were variable with a mix of Robust, Tenacious and Modest at Levels 1 and 2, and a split between Robust and Accepting at Level 3. The language of her reflections was Modest. The lack of stability in Alice’s sense of agency reflected her Basic performance. In addition, the absence of breadth of thought makes her responses weaker overall. An example from Level 3 illustrates this:

[The invention of writing allowed the Sumerians in Mesopotamia to keep accurate records of . . . ] It helped them keep track of laws they wrote, stories of their people & history. It also kept record of business transaction & other community
activities and religious beliefs & knowledge of medicine, mathematics, & astronomy.

Alice’s inferential reflections were almost exclusively ratings and were all Discouraged, further reflecting her lack of skill and confidence with Expository text. At Level 4 she continued to limit her thinking to answering questions, so there was no window into metacognition to provide analysis of how her inferential reasoning was breaking down. When she did engage in verbal reflection, it was very limited as shown in this example from Level 4: “I didn’t answer completely.”

After seeing her work along with a scoring guide, she shared a Vulnerable reflection: “I think that this assessment was difficult because most of the answers are in the passage but some were hard to find. Also, the level 4 + 5 questions were a little confusing.” This reflection at the end of the assessment showed metacognition about her challenges with the test process, rather than just shallow commentary on her products.

Expository developmental work. Alice’s think-alouds for Expository were split between Levels 1 and 4. This was the first time she showed evidence of metacognition at both literal and inferential levels during a think-aloud. At the literal levels about half of her moves were Weak moves of uncorrected miscues. However, metacognition was present more than before with some awareness of miscues and attempts to address them, as well as some basic connections to the text. An example of this can be found in the first Expository think-aloud: “that’s an odd word rashilly [sic]” At inferential levels there was connecting and evidence of metacognition around trying to make the inferences to gain understanding: “We were talking about rights for Americans in history.”

For the bulk of the unit coursework in Expository, Alice worked collaboratively with Shenala. Together, they exclusively worked at Levels 3 and 4 throughout the unit.
Shenala’s frequent absences affected the work production and the benefits of the collaboration, as Alice carried on in her absence, resulting in each of them doing more independent work and less collaborative work than the other teams. Observational field notes show that Alice spent more time out of her seat and interacting with other pairs during this time than other members of the quintain did, and these interactions were most often about the work. This was a significant change for Alice, as she demonstrated engagement more than distractibility. As with her experience in Narrative Nonfiction, the work showed a breadth of metacognition in trying to gain understanding of text and articulate that understanding. The quality of work for this collaborative pair ranged from OK to Weak. An example from Level 3 shows the breadth along with a combination of parts considered to be exemplary of OK and Weak Quality of thought:

*Title*--Confucius’s Ideas and Teachings. *Question*--Why did Confucius spend so much time discussing his ideas with his students? *Answer*--He gave more ideas and details to take informal conversations. He also did not write down his teachings. *Explanation*--I think he didn’t write down his teachings because he taught his students about his teachings into a conversation, so that way he didn’t have to write it down. *Quote*--“Confucius did not write down his teachings. He shared his ideas with his students through the give and take of informal conversations,” (pg.8).

During this unit, students worked in a variety of collaborative groupings, and then were assigned a partner for the largest section of work. From observational field notes, Alice’s sense of agency was Robust throughout, as exemplified by high levels of engagement and a willingness to lead collaboration, even with students with a history of being the top performers.

*Expository summative assessment.* Despite the progress shown in the developmental work during this unit, Alice’s performance on the summative assessment showed no growth, with a repeat of her Basic score of Level 3, Stage A. Her initial
Fragile reflection sheds some light on a potential contributor to this performance:

“Before I take this assessment I am feeling scared because I have a headache.” While the emergence during this assessment of the breadth of thinking shown during coursework was an indicator of metacognitive progress, this was done to varying degrees of quality. Her responses, while structured to show breadth of thought, frequently did not show a careful reading of the questions, with several responses that did not answer the question asked, but appeared to be answering a different question altogether. An example from Level 3 highlights this pattern:

[How did missionaries like the Whitmans attracted other people to settle in Oregon Country? Model--By writing inspirational letters home.] Answer--Marcus was helping the Indians and Narcissa was lonely, missing her family and her husband. Marcus and Narcissa left New York to live and work among the Cayuse Indians in 1836. Quotes--“Marcus Whitman and his wife Narcissa left New York to live and work among the Cayuse Indians in 1836” (Pg 1). When they arrive in Oregon life was hard. Marcus was away helping the Indians and Narcissa was quite lonely, missing both her family.

In addition to examples of nonresponsive work with a breadth of thinking present, Alice also had some simple responses that did not answer the question asked. This example from Level 4 illustrates: [Name at least three qualities that the settlers must have had to survive the Oregon Trail.] “The settlers must have had the train wagon, food, animals to pull the train wagon.”

Alice’s Vulnerable post-assessment reflection does not shed much light on her difficulties: “I think that this assessment was difficult because some of the questions were sort of difficult to find. I stayed the same and I think I did a little better.”

Sub-question. 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?
GMRT-4 and SRI. Alice’s performances of Proficient on both the GMRT-4 and the SRI are evidence of her closing the gap. Specifically, she demonstrated a growth of 21 months and an increase of 7 NCEs on the GMRT-4, as compared to 19.8 months and an increase of 5.5 NCEs as the average growth of the quintain, and 16.95 months, or 4.9 NCEs for struggling eighth grade readers receiving remediation in the other four middle schools of the test district. Alice achieved Proficiency with her score on the final measure of the year. Alice’s SRI score rose 183 Lexiles during the study, and dropped back down 120 Lexile points at year’s end, with a net gain of 63, placing her above the average performance of 51 for her counterparts in the quintain; however, her scores were all in the middle of the proficiency range with some movement within the range.

PARLA. Alice’s performance was highly variable across the three content areas, as measured by the PARLA proximal assessments. Her performance decreased between the formative and summative measures in the first unit, increased in the second, and demonstrated no change in the third unit. When it came to the PARLA measures, Alice did not succeed in closing the gap to achieve grade level performance. Specifically, on the PARLA-NARR, she declined by 2 stages from formative to summative assessment, placing her below the average growth of the members of the quintain and ending at Level 4, Stage B. On the PARLA-NARR NF, Alice improved by 1 stage, below the average of just under 4 stages for the rest of the quintain and ending below grade level, at Level 3, Stage C. Finally, on the PARLA-EXPO she showed no growth, placing her below her quintain’s average of near 3 stages, and ending at Level 3, Stage C, well below grade level.
Sub-question. 2) Do students participating in PARLI report a shift in agency and motivation?

Alice’s self-reporting about her motivation from the Middle School Assessment of Academic Self-Concept and Motivation (MAASCM) changed from Modest to the Tenacious category (see Table 6, The MST Taxonomy of Personal Agency Beliefs). For Alice, Robust was the most prevalent PAB rating by a sizeable margin, with Modest, Tenacious, and Discouraged being the other categories of some prevalence, as shown in the graph in Figure 26.

![Distribution of most prevalent PABs in Case 10.](image)

Across the three content texts--Narrative, Narrative Nonfiction, and Expository--the most prevalent categories of Personal Agency Beliefs varied as shown in Figure 27, with Modest being the most consistent. The Personal Agency Belief of Modest was at 21% at the start of the study, stayed there during Narrative Nonfiction, and increased slightly to 24% in the final unit. The Robust category went from 36% during Narrative, to 16% in Narrative Nonfiction, and ending at 14% during Expository. The Tenacious belief was the one of the two most variable, changing from 6% at the start, to 40% during Narrative Nonfiction, and dropping back down to 35% in Expository, as shown in the
EVALUATING THE PARLI FRAMEWORK

Finally, for the Discouraged category of PAB, Alice’s range of scores was 8% during Narrative, 0 during Narrative Nonfiction and 77% during Expository.

Alice’s Personal Agency Belief ratings were attached to specific work 67% of the time. In the Robust category there were 19 references total at the literal levels, with 10 in Narrative and the rest in Narrative Nonfiction. At the inferential level there were four occurrences in the Narrative unit. For Modest, there were three at the literal level (one in Narrative and two in Expository) and seven at the inferential (five in Narrative and from Narrative Nonfiction). In the Tenacious category, there were three instances in Narrative Nonfiction. Finally, the Discouraged category had three instances at the inferential level, one was in Narrative and the other two in Narrative Nonfiction.

**Figure 27. Case 10: Tracking the most prevalent PABs across the study.** This figure illustrates the variability of PABs throughout the three units of the study.

While artifacts continued to be a mixed grouping of Levels, Quality, and PABs across the Narrative and Narrative Nonfiction units, during the Expository unit, Alice did
not provide any information about her *Personal Agency Beliefs*, suggesting some fatigue with the work and the request to be metacognitive and reflective.

*Sub-question*. 3) *Do students participating in PARLI demonstrate improved metacognition?*

*MARSI*. On the Metacognitive Awareness of Reading Strategies Inventory (MARSI), a sub-score greater than or equal to 3.5 is considered to be a high rating, while from 2.5 to 3.5 is in the mid-range, and 2.4 and below is a low rating. Alice’s self-reporting indicated one high rating and two midrange ratings at the outset, and changed to two high and one low rating at the conclusion of the study. Her greatest reported changes in metacognition from the MARSI were in the areas of Problem-Solving Reading Strategies, which includes slowing down to read carefully, pausing and reflecting and similar strategies. Also in this area are Support Reading Strategies, which include a variety of actions such as note taking, paraphrasing, and discussing the material, among others.

Alice’s scores dropped by .65 and .6 respectively, with the former still keeping her in the high range of the rating scale, and the former dropped her from the midrange to the low range (3.0 to 2.4). For the subscale of Global Reading Strategies, she indicated a change from a rating of 3.3 at the start of the study to 3.6 at its conclusion, moving her up from a medium to a high rating. Global Reading Strategies include strategies pertaining to setting a purpose for reading, activating prior knowledge, making predictions, and so forth.

The intersection of MARSI and MAASCM does not yield illuminating information about Alice’s metacognitive development, but her think-aloud protocols,
while still loaded heavily toward the basic decoding of Level 1, did show some stretch into Level 4 and Level 5 across all three content areas. Alice’s reflections about the MARSI subscales attempt to connect a representation of the strategies she uses directly with her responses on the survey. However, her reflection for Problem-Solving Reading Strategies hints at, more than demonstrates, that she has an understanding of this subscale by her choice of example, as shown in Table 40.

Table 40. Reflections on MARSI survey by reading strategy for Case 10.

<table>
<thead>
<tr>
<th>Global Reading Strategies</th>
<th>Problem-Solving Reading Strategies</th>
<th>Support Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Because I read slow so I know what I’m reading.”</td>
<td>“Because I don’t fully understand prior knowledge.”</td>
<td>No comment.</td>
</tr>
</tbody>
</table>

*Think-aloud protocols.* The think-aloud protocols were part of this study as an effort to gain access to students’ metacognition while reading and comprehending grade-level texts in Narrative, Narrative Nonfiction, and Expository forms. Bella’s moves populated the first four levels of reading comprehension across the study time, with just over half (68%) being at Level 1, and 26% at Level 4 and Level 5 considered together as shown in Table 41.

Table 41. Frequencies of Levels of think-aloud protocols for Case 10.

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>.60</td>
<td>0</td>
<td>.08</td>
<td>.04</td>
<td>.28</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>.61</td>
<td>.03</td>
<td>.03</td>
<td>.30</td>
<td>.03</td>
</tr>
<tr>
<td>Expository</td>
<td>.79</td>
<td>.03</td>
<td>.03</td>
<td>.15</td>
<td>0</td>
</tr>
</tbody>
</table>
During the Narrative segment of the PARLI framework, the levels of responses and their Quality ratings were distributed across levels, as shown in Tables 42 and 43.

**Table 42.** *Case 10: Frequencies of Levels present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>106</td>
<td>.23</td>
<td>.11</td>
<td>.28</td>
<td>.17</td>
<td>.18</td>
<td>.02</td>
<td>0</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>125</td>
<td>.28</td>
<td>.26</td>
<td>.16</td>
<td>.23</td>
<td>.07</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>95</td>
<td>.28</td>
<td>.19</td>
<td>.16</td>
<td>.37</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

During the Narrative Nonfiction segment of the PARLI framework, the levels of responses were distributed across Level 1 through Level 5, with Level 1, Level 2 and Level 4 (Table 42), being roughly equivalent, while Quality ratings were also spread across levels, but still predominantly OK and Weak, as shown in Table 43.

**Table 43.** *Case 10: Frequencies of Quality present in work by unit.*

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>Strong</th>
<th>Good</th>
<th>OK</th>
<th>Weak</th>
<th>Wrong</th>
<th>Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>91</td>
<td>.09</td>
<td>.34</td>
<td>.16</td>
<td>.24</td>
<td>0</td>
<td>.16</td>
</tr>
<tr>
<td>Narrative Nonfiction</td>
<td>105</td>
<td>.01</td>
<td>.20</td>
<td>.36</td>
<td>.33</td>
<td>.01</td>
<td>.09</td>
</tr>
<tr>
<td>Expository</td>
<td>139</td>
<td>.01</td>
<td>.12</td>
<td>.45</td>
<td>.32</td>
<td>.01</td>
<td>.09</td>
</tr>
</tbody>
</table>

During the Expository segment of the PARLI framework, the levels of responses were also distributed across levels of comprehension, with clustering at Level 1 and Level 4 (Table 42). When it came to Quality during Expository, Bella’s work was also
distributed across categories, with the biggest category being OK, followed by Weak, as shown in Table 43.

Sub-question. 4) Does student performance on the assessments form a pattern of development?

Alice’s scores have improved over time on all quantitative measures, with her PARLA scores showing only slight variability when moving into a new content area; this fits with Fischer’s Dynamic Skill Theory (Fischer & Bidell, 2006), with learning of complex skills being a fluctuating process.

Summary. Alice was only Proficient on one of the proximal measures (PARLA-Narrative formative assessment). Her formative and summative assessments of agency and motivation went from a starting place of Modest, to and ending place of Tenacious. Her metacognitive scores on the MARSI survey showed movement in both directions. An evaluation of her overall performances and observational field notes allow one to make a case for the PARLI framework being effective with Alice.

Summary of the Chapter

The data shows that for 9 of the 10 cases, the PARLI framework was effective. The means by which the cases in the quintain demonstrated their growth are variable. The most illuminative data for further development of the PARLI framework come from observations and reflections. While the changes in the cases studied based on standard reading assessments are not statistically significant, the small sample size and case study methodology were selected to evaluate the merits of proceeding with the development of the PARLI framework and provide guidance for this development. The variety of data representations across the cases suggest the importance of continuing development of the
framework as a flexible thinking framework that supports differentiated instruction at deep levels of academic reading comprehension.
Chapter 6: Discussion

The final chapter of this dissertation provides a brief overview of the study, including a statement of the problem and the major methods involved. The majority of the chapter is devoted to a summary and discussion of the study hypothesis and the five research sub-questions. In addition, a discussion is provided on the implications for action and recommendations for further research.

Summary of the Study: Problem and Purpose

This study conducted an initial evaluation of the Pragmatic Analytical Reading Level Instruction (PARLI) framework with eighth grade struggling readers at a Midwestern middle school. This framework provides a systematic and targeted means for flexible instruction to remediate the reading comprehension deficits of a diverse population of struggling adolescent readers based on a transdisciplinary approach that integrates theories from the fields of reading research, cognitive development, motivation and engagement, and the emerging field of educational neuroscience. The purpose of this study was to conduct a formative assessment of the PARLI framework.

Hypothesis and Research Questions

It was hypothesized that participation in the PARLI curriculum framework for one academic semester would result in growth of reading comprehension among struggling readers. The overall research question was Is the Pragmatic Analytical Reading Level Instruction (PARLI) framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)? The answer was sought through responses to the research sub-questions in this study:
1. Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

2. Do students participating in PARLI report a shift in agency and motivation?

3. Do students participating in PARLI demonstrate improved metacognition?

4. Does student performance on the assessments form a pattern of development?

5. How can measurement tools, including observations, used with struggling readers result in better understanding of these students and how to optimize their learning opportunities?

Review of the Methodology

This study evaluated the PARLI framework in a formative capacity. As such, a mixed methods approach, in a naturalistic setting, emphasized a pragmatic paradigm that utilized the quantitative and qualitative research traditions applied in a practical manner to elucidate an understanding of the complex interaction between an instructional framework and the development of academic literacy. Case study methodology was used as the research framework to carry out the interpretive task of understanding the complex factors that contribute to the development of reading comprehension among struggling adolescent readers.

Limitations

A number of limitations were identified in relation to the use of think-alouds, videotaping of observations, teacher-made Common Assessments, and time available for teaching and learning. The think-aloud methodology was selected to provide explicit metacognitive data relative to how the readers were developing complex comprehension
skills. Despite multiple training sessions, students did not demonstrate comfort and proficiency with this tool, remaining at the level of predominantly reading aloud and revealing decoding and fluency struggles, as shown in the reported frequency data. To address this, an additional observer could have been recruited to sit with the student and provide prompting during this process.

Initial videotaped observations yielded little usable data and were undecipherable to the panel of literacy experts (three secondary reading specialists), as the physical realities of the classroom context did not allow for placement of the camera in a location that could yield both context and enough detail to be illuminating regarding any particular case, or the interaction between and among them. The recording equipment picked up multiple speakers at once, such that the voices were layered upon each other and it was not possible to distinguish who was speaking and what was being said. A videographer would have been advantageous to address this limitation, had one been available. In the absence of this resource, the videotaping did not continue as planned.

Additionally, one of the planned measures to allow for comparisons between the performance of members of the quintain and their non-struggling peers was the teacher-made Common Assessments. Unfortunately, these assessments were not examined for efficacy prior to administration, and yielded district-wide average performances clustering at the 60% and below levels, without normal distribution, creating an entirely inadequate vehicle for evaluating the gap between student groups. The lack of a disperse distribution of performance, and the skew to the low end of performance call the reliability and validity of these measures into question. It may also have been that the assessment was too difficult for the intended participants. The researcher did not
participate in the development of these measures or the option for the district to select a more reliable and valid assessment to use across all students, leaving the GMRT-4 as the only tool to measure relative performance and growth across groups. The SRI and PARLA measures were used to evaluate the gap based on expected grade level performance rather than a direct comparison between groups.

Another limitation of the study might be in the sample. This had the potential to be problematic because it limits the generalizability of the results to the larger population of struggling readers; however, this is a limitation of case study methodology as a whole. Further research could endeavor to conduct a broader mixed methods evaluation of the effectiveness of the framework or qualitative research on a larger population of struggling readers.

A final limitation of the study was the duration of both class sessions individually and the study as a whole. The development of high levels of student reading comprehension across academic contexts is complex and time-consuming. Introducing a novel framework that emphasizes thinking structures over the predominant mode of using questions to be answered at the end of a reading (often in multiple-choice format) is a significant adjustment for learners. Building this complex and variable means of response to reading takes time. Frequently, students would just begin to grasp a new concept or skill and it would be time to conclude class. Students often needed re-teaching and revisiting, as well as time, to recover their place during the next class session. At least twice a week one or more of the students complained about not wanting to leave and/or remarked on their surprise that class was over. Additionally, the pacing and total time for the Narrative and Narrative Nonfiction units seemed appropriate, but students
complained throughout the brief two weeks of the Expository unit that we were moving too quickly. Furthermore, the adjustment of the pacing of the PARLI instructional framework to respond, at least in part, to learners’ needs, resulted in the elimination of the planned Descriptive (science-based) unit from the study. Overall, it was clear that the time for learning did not meet the needs of the student participants.

The study could be enriched if it included a larger sample size of students across the spectrum of reading proficiency. The researcher would recommend that data be collected across such a broad sample, with qualitative data being provided from a smaller, representative sample from three groups: struggling readers, on grade level readers, and excelling readers performing consistently above grade level.

A battery of reading assessments may not be particularly useful as a primary means of guiding effective literacy plans for struggling adolescent readers. The presence of growth in confidence and competence demonstrated in the classroom data and self-reports of the quintain provided detailed and actionable information. Further, a consistent pattern of progress was absent across the variety of reading assessments utilized resulting in a lack of direction that would have been provided if such patterns of progress had been present. Taken together, these realities lend support to a diminished reliance on test scores when striving to develop the reading lives of struggling adolescent readers.

**Major Findings**

**Cases.**

*Shenala.* From Shenala’s (Case 1) results the intersections of artifacts and think-aloud protocols with the reading comprehension GMRT-4 and SRI were most illuminative. Her strong performances on these standardized assessments matched the
depth of thinking that she regularly demonstrated. She moved into the Advanced category on the GMRT-4 and the Proficient category with the SRI, after being very expressive in her think-alouds. Shenala was one of the most engaged in sharing her thinking while she read during the think-alouds. She repeatedly demonstrated an unusual combination of fluency when reading aloud with remarkably slow processing when reading silently. It was also unusual that her rate of three to four times slower than her slowest peer did not impede her ability to maintain what was read in her memory.

Shenala was able to respond at very deep levels of comprehension with a Strong quality of response, when time allowed. It seemed that the request to think while reading, and share her reflections, suited her thought process well, provided that time was available. In addition, the intersections among the two reading comprehension measures, observations, artifacts, and think-alouds with the MARSI also provided some useful insight. She reported a growth in the use of reading strategies that matched her results. However, Shenala’s growth across reading comprehension measures was not consistent. Her high degree of comprehension was not evident in her PARLA assessments for either the Narrative Nonfiction or the Expository assessments, but was apparent in class work within the PARLI framework, when she was present and completed assignments.

Shenala was not assertive within the group, but was generally willing to go along with the consensus when working collaboratively. She expressed some frustration with not having enough time, but did not correlate that with her absenteeism. As with the other students, she was particularly concerned with not having the same amount of time to develop the Expository practice as was provided during Narrative Nonfiction. However, her work production dropped significantly during this time in conjunction with her being
absent three and four days each of the two weeks of this unit. Her strongly positive performance on the GMRT-4 was justifiable cause for much celebration, and the focus of her ending reflection: “My parents are happier because I am doing better this semester. They are also happy that I am more confident.”

Shenala made the greatest gains among her peers on the GMRT-4, but remained largely disengaged from the process, in no small part due to her absences. She was, ironically, absent when students wrote their ending reflections and did not complete it. Through conversation she did share that she felt that she had gained skills in reflecting on what she read and responding in a sophisticated manner, which gave her confidence for high school. She did acknowledge that her rate of absence was likely to cause her considerable problems at the high school level, and asked the researcher to speak with her father about it so that he could engage her high school counselor in offering her some extra support to be academically successful. This showed an assertiveness that came and went with her throughout the study time.

**Rafael.** Overall, Rafael’s (Case 2) performance was unreliable. At times he showed glimpses of competence, but did not sustain this level. His most consistent level of performance throughout the study placed him two grades below his expected level. Rafael was not assertive within the group, except during periods of work limited to discussion, and was generally happy to give the appearance of productivity when working collaboratively. He was well practiced in hiding his lack of competence with his strong verbal skills, his robust people skills, and the ability to appear busy. Overall, he did not display the motivation and engagement of his peers, yielding the poor results that the research literature would predict.
Rafael’s results from the intersections of MARSI and MAASCM scores with all other data were most illuminative. His lack of coherent thought connecting the MARSI survey with his reflections show a pattern of being disconnected. Further, his reflections share his limited use of strategies. This fits with his poor performance. When his sense of agency is considered, his shift from *Tenacious* to *Vulnerable* also matches his poor performances.

Throughout the study, he expressed some frustration with not having enough time, but did not correlate that with his lack of production during class. It is worthwhile to note that he was involved in tutoring for several hours two days after school in math, and successfully improved his math performance at school during this time of stagnation in reading. In addition, he is a competitive tennis player at a high level and was very involved in practice and tournaments. It is likely that these competing demands sapped his energy and he had little left to give to building his reading comprehension during this time.

Rafael’s final reflection is indicative of his largely neutral, leaning toward positive, stance that he took throughout, regardless of his actual performance: “I think my progress has been some [sic] because I worked really hard on the Narrative Nonfiction stuff and I think I did make good progress.”

Qualitative data suggest that one-on-one intensive work is called for, as multiple instructional interventions in small groups have not yielded sufficient results, either as measured by class work or assessments.

*Edward.* From Edward’s (Case 3) results the intersections of observations with artifacts and each of the reading comprehension measures (GMRT-4, SRI, and PARLAs)
were most illuminative. They demonstrate his consistent growth in metacognition and resulting reading comprehension that generated improved test scores. In addition, the intersections among the three reading comprehension measures and these measures with the MARSI also provided some useful insight. They show a positive shift in his sense of self agency through self-reporting, observation, and analysis of artifacts.

Edward was among the cadre of students who initiated the request with the researcher/teacher to extend the Narrative Nonfiction unit at the expense of the Expository unit. He took his time and was thorough in the work in the unit, building his sense of agency as the unit progressed. This unit allowed him the time he needed to flourish. During the Expository unit, he expressed some frustration with not having the same amount of time to develop this practice as was provided during Narrative Nonfiction. However, his work production also dropped a bit. He explained this as a function of the increased workload in his other classes, as the school was trying to compensate for the loss of more than a week of instruction prior to the semester summative exams in core classes as a result of school closure due to inclement weather.

Edward did not make the greatest gains among his peers in formal assessments, but showed evidence of profound change throughout the course of the study. He became proud and confident, wanting to share his work with peers and his family. His body language changed from a young man who mostly looked down and spoke softly, to a young man with a habit of looking others in the eye, explaining his position, arguing respectfully, advocating for himself, and laughing a great deal, even during academic tasks. Toward the end of the study, Edward said that the framework made the thinking required more clear to him and he knew how to approach it. Observations trace a rather
steep change from constant checking in with the teacher, to independent ownership of his thinking, with periodic checking with peers, and less frequent checking with the teacher. When asked at the study’s conclusion how he felt about himself as a reader, his reply was true to his history of understatement, while still being illustrative of who he became: “I feel really good in my progress that I made. I have been reading more at home when I have nothing else to do.” Given the novelty of his academic success during the study year, and the support that helped him to be successful, further literacy support would be beneficial in the form of tutorials to aide his successful work completion in core classes.

**Dominique.** Exploration of the data matrix show that Dominique’s (Case 4) performance tended toward moderate growth relative to her peers. Dominique improved on the PARLAs, albeit only slightly, but this improvement was not evident when she was observed, with her best work happening during the first unit (Narrative). The intersections among the three reading comprehension measures show the weakest growth in SRI.

As it relates to Dominique’s performance, the intersection of the reading measures on the matrix shows better growth on the proximal measures than the standardized measures; this is generally an anticipated pattern of performance, with standardized performances lagging behind those on proximal measures. However, it may be that this also points to a consideration of incorporating the opportunity for more low level work (at the decoding and fluency levels) within the PARLI framework. Dominique is neither the first nor the last student to prematurely be moved on from the decoding level of intervention for social reasons.
Regarding Dominique’s performance, the intersection of think-aloud protocols with observations and artifacts may be the most revealing in telling her story about the underlying reasons for her mediocre performance overall. Her think-alouds were almost exclusively at the surface level (Level 1) and revealed residual decoding issues, most of which she made no attempt to address. This deficit was also expressed in her weak spelling across artifacts. She was taken out of the school’s reading intervention class dedicated to resolving the base-level decoding and phonemic awareness issues for largely social reasons, as discussed in the Previous Years of Schooling for this case.

When taken along with her classroom performance in her core classes being consistently in the B range, as Dominique utilizes her array of adaptive strategies, the decision was made to move her up a level in reading intervention. Her adaptive strategies are visible during observation; a primary strategy is the utilization of peers to check in with and confirm her thinking. Dominique rarely checks in with the teacher and rarely demonstrates any concern for correct spelling and syntax in her writing, but her speech is age appropriate and clear.

Dominique’s mode of navigation through the learning environment is largely fueled by an expectation of positive results. After the conclusion of the study, she did move toward more high level work with peers, and her core teachers commented on the progress they witnessed. Her concluding self reflection shares this positive approach: “I feel that my reading has improved. I have been reading more then I use too and has been liking more books.” Dominique would probably not benefit from further reading intervention, as she is satisfied with her current skill level. Given her residual deficits, access to extra time and tutoring would benefit her academic progress.
Suzie. One of the more interesting intersections on Suzie’s (Case 5) matrix is an exploration of her self-assessment PAB ratings relative to those from observations and artifacts. Like Dominique, her self-rating seems to be more a measure of her personality, with her comment: “Tenacious is more predictable.” She is, in fact, a tenacious young woman and was likely referencing this rather than carefully responding to the scale.

Observations revealed sensitivity by Context, with uncertainty at the beginning of each unit, proceeding to a big drop in member checking during the middle, with a rise at the end attached to concerns about performance and grades across all Aspects and Products. As a perfectionist who receives counseling services to help her maintain balance, her strong self-perception was plagued by self-doubt that makes both her PABs and her general work production somewhat fragile and variable. However, when it was time to perform, she was generally prepared and did well, as evidenced in most of the reading measures; PARLA Narrative was the one exception. This is particularly interesting in light of her voracious enthusiasm for fiction, as she sped through novel after novel, sometimes consuming more than one per week during the course of the study.

Suzie’s final reflection reviews her middle school journey as a reader well: “I think I have improved alot since the last time in 6 grade that I was in reading class. I am very happy where I am right now.” Suzie’s final performances suggest that no further academic support is needed for her continued school success.

Tanner. Tanner (Case 6) represented the greatest incongruence between test scores and the work completed with attendant gains of all in the quintain. He is diagnosed with both Aspergers and Obsessive Compulsive Disorder and exhibits both for the casual observer to notice. Given these traits, it is not surprising that he rated himself in the
highest $PAB$ category on the MAASCM. In reality, while he is academically competent and a very intellectually capable young man, his fear of failure, defined as anything less than perfect, was profound.

Tanner’s Aspergerian nature presented a host of challenges daily that come with the gifts that neurotypical students do not face. He was pulled by a strong force to stay in the zone of what was concrete and could be tightly governed, if not by him, then at least by a trusted adult. His agreement to participate in this study, which featured a framework with the development of inferential thinking as its core, was a bit surprising. He and the researcher spent considerable time in the spring of the previous year discussing what participation would entail and the potential benefits. In Tanner’s unique way, he would discuss for a few moments then leave the researcher’s close proximity and go work independently. For the first few days, he came back more than once during a class. Then he came for a brief dialogue daily. After about 2 weeks he informed the researcher that he trusted her and thought this would help him prepare for high school, which was very intimidating to him at this time, and that he would assent to participation.

The intersection of MARSI and the reading assessments was another place where some understanding of the complexities of being a young person with the clinical diagnoses Tanner has, and working to develop deeper comprehension skills, was available. Through observation there was a record of Tanner’s enthusiastic engagement with rubrics. He used them as a focal point of interaction with his peers and strived to master their use in ensuring that he was doing his best thinking and sharing it on paper. The rubrics were a vehicle for him to experience significant growth in $Personal\ Agency\ Belief$ in practice. Tanner’s adeptness with the Student Guide and the rubrics helped
strengthen his position in the community; students were vying to work with him and he basked in their praise. For the first time ever, he earned a nickname from a peer, T-Man, which made him grin from ear to ear and brought tears of joy to the adults who love him.

He worked in a constantly shifting configuration of dyads and groups, always with the Student Guide and rubrics in hand. He consistently used a variety of reading strategies that agreed with his MARSI survey results. He spent class time seeking input from peers instead of the teacher, representing tremendous growth. Tanner engaged with his peers, leading discussions of material read and how to respond. He truly became a full member of this community and on many days, for either the entire period or a significant portion of it, he was indistinguishable from his peers to the casual observer. Tanner repeatedly requested a celebratory party at the end of the study, providing further evidence of his shift in his sense of self agency as a full member of the classroom community.

The importance of building rapport and a safe place for students to let themselves be known and take academic risks is truly underscored by Tanner. The PARLA framework is inherently flexible and provides a vehicle for the establishment of differentiated instruction that respects individual learners’ needs, while providing a focus on achieving high academic literacy goals. Tanner’s closing comment shows the young man he has become and how generously he shares credit for this growth: “This year helped me a lot! My confidence level went up because I knew I could do it and had an awesome teacher that helped me through all these years!” Providing Tanner with continued easy access to extra academic and social support will ensure his continued
growth in competence and self agency that will fuel his success academically and in life outside of school.

**Sandra.** Sandra (Case 7) was the only student in the quintain to reach the level of Proficiency across all three of the content areas on the PARLAs. Despite this strong showing, she proved to be fairly enigmatic. Examining her matrix, she was also proficient on the SRI, but remained in the middle of the Basic range on the GMRT-4. Historically, most students demonstrate gain on the GMRT-4 last, probably in part because the GMRT-4 is a timed assessment; whereas with the other tests, students are allowed as much time as they need.

The intersection of artifacts and PARLA results suggests that Sandra’s many attempts to work at the highest levels in the hierarchy, which most of her peers did not attempt, fed her progress on the PARLAs. Her performance on the summative test for the Narrative Nonfiction unit was the strongest showing of the quintain. Interestingly, throughout the unit she was generally behind schedule with work completion. She embraced independent use of the Student Guide, and generally grasped the concepts at each new level quickly. Sandra’s reflections were all fairly guarded and surface. Hers is a story about generally rushing through class work to have time for play. She aspires to be an actress and spent part of each class entertaining her peers, such that her humor was used during instruction as an example of inference. Sandra made solid gains throughout the study. She rather quietly focused on stretching her written responses to the best of her capacity, as this is where she perceived that both her deficit and the more powerful academic skills reside. Her final reflection shows this as well: “I think I improved in many ways like improving on my writing skills (responces)[sic], because It has helped
me with Comm [sic] Arts ALOT! and I think it shows science [sic] last year.” Sandra’s final performances point to no further need for academic remediation.

**Holly.** The intersection of MARSI and MAASCM with artifacts and PARLA assessments illuminates Holly’s (Case 8) development into a competent reader over the course of the study. Her self-reported growth in the use of reading strategies matched her progress with outcomes. In her habit of candid reflection, her self-reported sense of agency fit her growth as a reader without being overly confident relative to her actual competence.

Another place of interest on her matrix is the intersection of the three reading assessments. The fact that she reached proficiency on the PARLA Expository assessment, arguably the most challenging of the six, and the GMRT-4, usually the last measure to show evidence of growth with striving readers, while her SRI score remained firmly in the center of the Basic range suggests that the SRI may not have provided an accurate measure. That said, Holly improved more than 1 year’s progress on this measure, and her lack of previous success with the Read180 program lends further support to a mismatch between her most desirable learning modes and this software program. The SRI is the proximal measure for the Read180 program, and students who participated in Read180 during sixth and seventh grades like Holly did, took the measure three to four times each year. It could also be that her past mediocre experience with the program had her anticipating similar results; however, she also has had extensive negative experiences with the GMRT-4, so this is not likely. Her artifacts also show her exploration of the higher levels in the framework hierarchy, with a generally accurate PAB that aligns with the Level and Quality of her work across all three content areas. In addition, Holly was
one of the few members of the quintain who consistently attempted to gain awareness of her metacognition at higher levels during the think-aloud protocols, with consistent efforts and Level 4, and frequencies in the mid to high teens for both Narrative Nonfiction and Expository think-alouds. Taken in the gestalt, her positive test results seem to be a true measure of her growth.

Holly’s final reflection is both accurate and understated: “I feel that from the Beginning [sic] of school to now I have More confidence that I will and can hit grade level or above. Im [sic] Happy with every thing [sic] that’s happened.” Holly’s success with challenging texts, and her exporting these strategies to her core classes suggest that no further formal intervention is needed. However, the novelty of her competence will become habit if she has access to differentiated instruction with literacy assignments in her core classes. This instruction would entail targets with tremendous specificity as to the level of thought expected. When the depth of metacognition needed is clear to Holly, she now has the means to achieve the academic tasks required of her.

**Bella.** Bella (Case 9) was Proficient on most of the reading comprehension measures and all of those that were taken after the Narrative unit; PARLA Narrative is the one exception. Ironically, Bella was gaining mastery of fiction last year and opted to participate in this study to bolster her nonfiction skills for high school; a look at the data shows that she has done just that.

In addition, exploration of artifacts with observations adds depth to this understanding, showing an accuracy of positive PAB ratings with Strong Quality ratings, as well as an improvement in moving toward higher levels of work during the first two units, with these higher level responses remaining constant during the final unit.
While artifacts remained constant from Narrative Nonfiction through Expository, PARLA scores did not. The greatest growth took place during Narrative Nonfiction, but Bella built on this with her highest achievement overall in the Expository unit. This is particularly noteworthy given the abbreviation of this final, challenging unit.

Bella’s reflections reveal a young woman who found the PARLA assessments to be appropriately challenging if she could *not* master them. This is a young woman who sets high standards for herself and takes pride in true accomplishment, not the completion of learning opportunities of little challenge. Her comments also demonstrate her recognition that this framework encompasses all levels of reading comprehension, not just those that pertain to middle school, and that there are multiple levels at which one can understand a text. Her final reflection provides a glimpse of a young woman ready for the challenges of high school: “I think I am doing better and feel more confident about myself. I also am enjoying reading more.” Given her tremendous and largely stable success, no further reading intervention is recommended for Bella.

*Alice.* Alice (Case 10) did not present a predictable pattern with her reading comprehension test scores. Her fluctuating SRI scores throughout the course of the study along with her mediocre performance on the proximal assessments did not predict the Proficient performance on the GMRT-4. Exploring the intersection between artifacts and the GMRT-4 does not point to this performance either. Alice’s work was largely spread over most of the levels with inconsistency in *quality* in evidence, and a mix of *PABs* that made her belief in her ability to be successful, and thus the fuel for her engagement, hard to pin down.
Observing Alice, one could miss her struggles to focus if not paying close attention. She has mastered the ability to look engaged, regardless of whether or not anything is happening. Her attention and focus were inconsistent. Her behavior included being mature, directed and focused one day, followed by her dissolving into constant giggles and being in perpetual motion the next day, and sitting quietly seeming to work but accomplishing nothing the third. She is an engaging and articulate young woman who interacts well with adults.

Alice was able to work in any dyad or small group, but had issues with fitting in beyond this classroom community that fueled some periods of unrest for her when new groups were forming throughout the study. Some of her worry about fitting in may be driven by her very diminutive stature. With the desire to be like the crowd that governs much of the middle school social decisions, she seems a bit self conscious about how small she is, commenting about it with frequency in casual conversation, and bristling visibly when others mention it. It is worthwhile to note that at this time Alice was also developing a relationship with her first boyfriend. This provided a fair amount of distraction during class with her classmates, as Alice shared the drama with the class whenever possible, and was a bit distracted by it herself the rest of the time. Her performance on the Narrative PARLA is the only one that earned her a Proficient rating, while her other performances were largely stagnant.

Given the mixed results, it is disappointing that Alice’s reflections were not particularly instructive in trying to understand what transpired for her over the course of this study. While her results were certainly variable, from her perspective the experience had the desired effect in preparing her for high school:
I feel more confident about going to high school. I don’t think I would want to go to high school because at the beginning of the year I wasn’t confident at all. I think I’ve improved a lot and I thank Mrs. -- for everything!

Alice’s difficulties with work completion and inconsistent focus and performance suggest that further academic intervention focused on the executive functions required for school success would be most beneficial.

Cross-case analysis. The previous sections of this chapter have shared discussion of the case studies of the 10 students who participated in this study. This section provides information resulting from a cross-case analysis. Taken on the whole, the quintain illustrates the complex and dynamic interactions that the five research categories that inform the PARLI framework predict. The development of reading comprehension is complex and involves “working parts” that must interact with each other in a highly coordinated manner to achieve results, as illustrated in the PARLI heuristic. The information in this section is organized to answer the research questions for the quintain as a whole. In accordance with previously detailed limitations of this study, caution is advised when making generalizations to a larger population from this analysis. The overarching question was Is the Pragmatic Analytical Reading Level Instruction (PARLI) framework effective with struggling middle school readers, based upon both the outcomes it produces (reading comprehension, agency and motivation) and the nature of the processes contained within it (metacognition)? When one analyzes both quantitative and qualitative data, the PARLI framework is effective. Specifically, 7 of the 10 members of the quintain (Cases 1, 4, 5, 7, 8, 9, and 10) no longer qualify for remedial services, based on the school district’s criteria. It is interesting to note that the greatest gains within the proximal measures (PARLAs) were made during the middle unit of Narrative Nonfiction,
when students were largely able to build on their experience with Narrative, and move it successfully into the nonfiction arena. The optimum circumstances of this unit provided a context that facilitated the interactions and outcomes predicted by the research literature in the five categories that inform the PARLI framework as represented in the PARLI heuristic. Students were motivated, engaged, and ready to build on their recent cognitive development from the Fiction unit. The instructional strategies, both reading specific and general, were structured to yield results predicted by the literature under optimum circumstances. The Expository unit featured half as much instructional time, a likely contributor to the less impressive growth therein, as most complained that they were not yet ready for a summative assessment.

Sub-question: 1) Does the PARLI framework contribute to closing both the fiction and nonfiction reading comprehension gap between struggling readers and non-struggling, grade level readers?

The members of the quintain can be categorized in four groups, based on conventions of the field of literacy: a) English Language Learners (ELLs); b) those with residual decoding issues, or RD; c) those with a specific reading comprehension deficit without a decoding problem, what Cutting et al., (2009) labeled as S-RCD, and; d) an individual with a clinical diagnosis. While the members of the quintain did not close the gap in reading comprehension relative to their non-struggling peers to a level of statistical significance, they did show gains in reading comprehension across all measures, including closing the gap in a number of specific instances.

For 9 of the 10 cases in the quintain, improvement in reading comprehension skills was appreciable when considering both their reading assessment scores and their
regular coursework. All of those in the RD group were successful on at least two of the three reading assessments (the district criteria for grade level performance). Both of the ELLs were also successful, with one closing the gap on all measures and the other on two out of three measures.

For the S-RCD group, two were successful on two of the three measures and one was not successful in closing the gap on any of them; nor was he successful with coursework. Finally, the case with clinical diagnoses of Asperger’s and Obsessive-Compulsive Disorder was not successful on any of the reading assessments. He did, as highlighted in the results, demonstrate grade-level work with coursework by the end of the study. While only two of the cases (Bella and Suzie) closed the gap on all reading comprehension measures, the majority of the 10 cases (6 of 10) closed the gap on two of the three reading assessments that were included in the study. Given the natural fluctuations in performance among adolescents from event to event, two out of three successes is evidence that they indeed made progress to closing the gap. When taken in conjunction with the additional evidence from coursework artifacts that also show successful work at an inferential level appropriate to their grade level, support is present for the assertion that they did, indeed, close the achievement gap.

Sub-question: 2) Do students participating in PARLI report a shift in agency and motivation?

Students exhibited and reflected the strongest sense of personal agency during the middle, Narrative Nonfiction, unit. This unit was, based on the research literature, optimally executed. Observation and artifacts also reflect greater levels of competence here than in the abbreviated Expository unit. It is likely that equal time spent in
Expository work (making it equivalent to Narrative and Narrative Nonfiction) would have contributed to greater growth overall on the GMRT-4. When it comes to a sense of agency overall, the majority of the quintain (8 of the 10) demonstrated a sense of agency in the Robust or Modest range more than the other categories on the scale. The third and fourth most prevalent PABs across the quintain were Tenacious and Accepting. Not surprisingly, students varied over time in their sense of agency, both reported and observed. This fits with the variable nature of the academic tasks across the three units of the study. Across the study, the analysis of Personal Agency Beliefs based on observations more closely reflected development in reading comprehension skills.

Sub-question: 3) Do students participating in PARLI demonstrate improved metacognition?

Cross-case analysis did not prove to be particularly informative for this question. Overall, demonstrable changes in metacognition, beyond reading assessment scores, were particular to the individual cases, with the exception of the S-RCD subgroup. The S-RCD subgroup had some consistent changes that hint at the possible strength of the framework with this subgroup in particular. The quality of their thinking showed appreciable growth in the Good category with a coordinating drop in the Weak category during the most-productive unit (Narrative Nonfiction).

This group also moved from concrete level thinking to inferential level thinking during the Narrative Nonfiction unit, but was unable to sustain this greater degree of inferential thinking during the Expository unit. The shorter duration of the Expository unit is a likely contributing factor. The RD group did not show such consistency, although 3 of the 4 shifted to more inferential thinking in the Narrative Nonfiction unit.
and held onto these gains through the Expository unit, without any similar consistency in quality of thinking. While the overall success of the quintain points to improved metacognition for all but one of the cases, consideration of these changes based on belonging to one of the four sub-groups did not yield actionable insights.

Sub-question: 4) Does student performance on the assessments form a pattern of development?

The PARLA assessments, each representing only one genre rather than a blend of genres, showed the strongest consistency when compared with Fischer’s (2008) developmental theory. The assessments were parallel in construction regarding depth of thought needed at each of the levels of the hierarchy. The text genres moved from the most straightforward (fiction), to slightly more complex narrative nonfiction, to the most challenging (expository) genre. This hierarchical structure within (Levels 1-7) and across PARLAs (narrative through expository) makes these measures sensitive to incremental growth. As predicted by Dynamic Skill Theory (Fischer, 2008), the quintain improved overall from formative to summative in the Narrative unit, then declined again as the academic challenge rose with the formative for the Narrative Nonfiction unit. Following the similar predictable pattern, progress made in reading comprehension skill during the Narrative Nonfiction unit was sensitive to the increased demand of the Expository unit formative assessment. This was apparent in the pattern of declining performance between the summative PARLA-NNF and the formative PARLA-EXPO. In contrast, the standardized reading measures do not emphasize a taxonomic structure. They are separate instances of a similar assessment that does not greatly vary the academic demands. This could lead to an anticipated steady, if slow, rise in performance as skills develop, rather
than the fluctuating pattern predicted by Dynamic Skill Theory (Fischer, 2008) and shown in the PARLAs.

Sub-question: 5) How can measurement tools, including observations, used with struggling readers result in better understanding of these students and how to optimize their learning opportunities?

There is no single solution to the complex problems of adolescent literacy. Although the planned analysis was a focus on evaluating the quintain, when contemplating the merits of further development of the PARLI framework, illumination of the specific experiences of each case provides potentially more valuable, subtle, and complex insight about the ways in which the framework contributed to the academic literacy and proficiency of the cases in the quintain. By recognition of the traits shared among individuals who struggle to attain literacy in general and those participating in this study specifically, this particular insight may prove helpful to those charged with developing struggling readers’ academic literacy. The Expository unit was shorter than the previous two (Narrative and Narrative Nonfiction), in response to student progress and engagement with Narrative Nonfiction. As mentioned previously, during the Narrative Nonfiction unit students were highly engaged and making significant progress. They requested that the unit be extended; this was, to say the least, an unusual request from students in middle school. Honoring a commitment to responsive teaching, the request was granted. One of the outcomes of this request being granted was an abbreviation of the Expository unit.

In evaluation of the PARLI framework, the evidence that students participating in the study moved into the Proficiency category on the GMRT-4 at a higher rate than their
peers receiving other forms of reading remediation (based on district-wide archival data), along with the difference in time spent per unit and resulting performances, point to the merit of proceeding with the development of the framework.

**Conclusions**

The results showed that a) based on the preponderance of the data, the PARLI framework contributed to reducing the comprehension gap between struggling readers and non-struggling, grade level readers; b) that the students participating in the PARLI did report a shift in agency and motivation; c) that most of the students participating in the PARLI did demonstrate improved metacognition; d) it was generally the case that student performance on the PARLA assessments formed a pattern, but on the standardized measures (GMRT-4, SRI) it did not; and that e) understanding struggling readers in depth is a complex business. It necessitates an individual approach to exploring how a host of measurement tools can be considered in the development of a story of an individual’s literacy life and the most advantageous way to approach guiding this continued development to write a story of high levels of academic literacy marked by power and choice.

One of the most powerful realizations for the researcher occurred as some photographs were being taken at the end of the school year. This community of 10 learners arrived in my class as a group of young people who were very guarded and spent most of their time looking down. Their experiences with failure in school informed all that they did. The majority of them began their secondary reading journey in my sixth grade class. As we were laughing and having a good time celebrating our successes, I was easily distracted and did not get the full impact of the pictures that were taken until a
bit later in the day. Each and every one of these students was looking at the camera, standing up, and smiling or laughing. This was important for me as a teacher, as it was confirmation that these individuals had been transformed in a meaningful way by their experiences of interacting with text in complex and challenging ways.

**Implications for Action**

The PARLI framework merits further development according to the findings of the current research study. As educators in the 21st Century, we have a primary obligation to structure learning experiences that provide opportunities for our young people to develop the highly complex literacy across disciplines that is required to become successful, independent contributors to society. When considering the challenges of advancing comprehension among struggling readers in particular, the PARLI instructional framework provides specific scaffolding through direct instruction. This finding supports the work of the Rand Research Group (RRSC, 2002), Snow (2002), Dymock and Nicholson (2007), and Vacca (1998) that specifically considered reading instruction as well as research from the field of instructional strategies (Bennett & Rolheiser-Bennett, 2001; Bulgren, Deshler, & Schumaker, 2003; Marzano et al., 2001; and Tomlinson, 1991, 2001). This structure engages learners to focus on their metacognitive processes and gain the heightened awareness of the techniques and strategies for approaching complex reading comprehension challenges. The transparency of the complexity and hierarchical nature of the reading process (Hillocks) within the PARLI framework supports this metacognition and gives learners the opportunity to see how to use effective comprehension techniques. This finding supports earlier work in the role of metacognition in reading comprehension (Abromitis, 1994; Cutting et al., 2009;
Where historically we have provided readers with individual strategies rather than a complex, coordinated framework across content areas (Alexander, 2003; Alvermann, 2002; Beers, 2003; Biancarosa & Snow, 2006; Greenleaf & Hinchman, 2009; Harvey & Daniels, 2009; Torgesen et al., 2007; Vacca & Vacca, 2005) and tested comprehension, PARLI supports the actual teaching of comprehension. This finding supports the work from the field of adolescent literacy including that of Alvermann (2002); Beers (2003); Daniels and Steineke (2004); Harvey and Daniels (2009); Harvey and Goudvis (2007); Keene and Zimmerman (1997); Robb (2008); and Schoenbach et al. (1999).

These struggling readers were scaffolded in their focus on owning their own metacognitive processes and really learning when and how to apply particular strategies to succeed in gaining understanding of complex, secondary-level material read. Pivotaly, they recognized a difference between this instructional framework and their previous experience, resulting in a willingness to engage with the process despite previous failures. For these struggling readers, the framework provided them with the knowledge that understanding what is read happens at more than one level, some more simple and on the surface, some deeper. As a result of this, the research on the PARLI supports that of Ford (1992), Bandura (2006), and Guthrie et al. (2007) connecting engagement and motivation to complex learning, as well as that of Bennett and Rolheiser-Bennett (2001) and Bulgren, Deshler, and Schumaker (2003) that highlights the power of transparency in instructional practices. This knowledge along with the direct instruction to practice these
skills, and the explicit instruction regarding the transference from one reading context to the next, built in these previously struggling readers the ability to regulate and control what they knew and were able to do, as well as defining the nature of the task at hand to recognize when support might be needed. This reality was most clearly articulated by Tanner (Case 6).

Consideration of the results garnered through the PARLI framework in this study, relative to the research literature base that informed its development is most clearly carried out by a return to the heuristic (Figure 1) from Chapter 2. The heuristic for the PARLI framework features motivation and engagement as the lead gear. The PARLI framework provided a structure that supported these struggling readers in a way that was consistent with both the American Psychological Association’s emphasis on learner-centered instruction (APA, 1993) and autonomy-enhancing instruction (Guthrie, 2007). The resulting positive shifts in agency demonstrated by the overwhelming majority of the cases is in concurrence with both Bandura’s (2006) Social Learning Theory, and Ford’s (1992) Motivation Systems Theory (MST).

While this study did not incorporate fMRI data that could reveal correlations between the hierarchical nature of the framework and what is currently known about the cortical network cycle illustrated in Chapter 2 (Figure 6), the evidence does support this research (Fischer, 2008). Accordingly, the PARLAs do provide data of student development in a pattern that is consistent with Fischer’s (2008) developmental theory as highlighted in the cross-case analysis early in this chapter. At the same time, the relative brevity of the study did not allow for a clear mapping of the progression of skill development, from Representations through Principles, that takes place as one develops
into a sophisticated, literate learner and generator of knowledge. Consistent with Fischer’s (2008) illustration of the cyclical spurts of development in Figure 4 of Chapter 2, improved performance in reading comprehension was greatest as they were guided with high support up the progression of reading skill. It is a reasonable hypothesis that the proximal measures (the PARLAs) provided a greater degree of scaffolding than the standardized tests, accounting for these PARLA results being consistent with Fischer’s Dynamic Skill Theory, while students’ scores on standardized measures reflected the flatter development of independent functional levels, as portrayed in Figure 4 of Chapter 2. Given a longer study period, perhaps even over more than one year, the potential to gain support for the efficacy of the PARLI from its coherence with Dynamic Skill Theory increases.

This research study did not support previous research on the power of think-aloud protocols in strengthening and displaying growth in metacognitive awareness. While Coté and Goldman (2004) found correlations between strategies that students reported using and reading comprehension scores, this study did not, as highlighted in the Limitations section earlier in this chapter.

Finally, pure neuroscience directly applied to specific teaching and learning objectives in a complex learning environment is in the embryonic stages (Fischer, et al., 2007). While this neuroscience connection may well evolve into an effective resource for anticipation, preventions, and focused remediation before struggles with reading serve to de-motivate and fuel disengagement of this significant portion of the population, the science is not yet there. The results of this study are consistent with the findings in the growing literature base (Bunge et al., 1999; Cooke et al., 2001; Cutting et al., 2001, 2001;
Dehaene, 2009; Della Chiesa et al., 2007; Giedd et al., 1999; Schmalhofer & Perfetti, 2007; and Shaywitz, 2004), but the absence of specific neurological data gathering in the study design precludes any conclusions about these patterns.

Authors like Biancarosa (Biancarosa & Snow, 2004), Greenleaf (Greenleaf & Schoenbach, 2001; Greenleaf & Hinchman, 2009), Scammacca et al. (2007), Schoenbach et al. (1999, 2003), Snow (2002) Snow et al. (2003) and others are all calling for schools to focus on the development of complex, deep literacy across the content disciplines. This call is joined by groups like the Council of Chief State School Officers and the National Governors Association (2010) in their drive toward curriculum like that articulated in their 2010 report. Flexible, research-based frameworks that support this complex work will be needed to provide a variety of paths to the same end, since one size most certainly will not fit all. The PARLI is one such framework. Deep thinking takes time. This framework provides the potential structure to support collaboration across content areas that holds the greatest promise for leveraging progress in abstract, complex thinking. Students, particularly at the middle school level, do not see connections; they must be walked across the bridges teachers build with explicit guidance regarding these connections (Alvermann, 2002; Beers, 2003; Daniels & Steineke, 2004; Harvey & Daniels, 2009; Harvey & Goudvis, 2007; Keene & Zimmerman, 1997; Langer, 2001; Robb, 2008; Schoenbach, Greenleaf, Cziko, & Hurwitz, 1999; Vacca, 1998).

The key improvement suggested for the PARLI framework by the results of this research is the duplication of the structure and time of the Narrative Nonfiction unit. To allow this greater time, implementation should be explored across a full academic year, such that each of the four text genres (narrative, narrative nonfiction, expository, and
descriptive) receives equal time for practice. In addition, if following several weeks of instruction and practice time with each unit in a reading or literacy class, teachers coordinated the infiltration of this framework into the appropriate core class by genre, it is likely that the gains in reading comprehension and students’ sense of personal agency that fuels further engagement and growth across the academic context, would be significant. By developing a step-by-step means for this expansion, the PARLI framework would be greatly improved. Once the PARLI framework is expanded, literacy specialists and content area teachers at the middle and high school levels can use this improved framework and the findings about the challenges that learners faced with expository text and the use of metacognitive strategies to work collaboratively to build instructional opportunities. Effective opportunities would allow learners to benefit from explicit instruction and practice in reading comprehension with these complex texts. As shown in the comparison between the Narrative Nonfiction and the Expository units of this study, the application of metacognitive strategies needs to be demonstrated and practiced explicitly over time to yield reliable reading comprehension growth.

Additionally, colleges of education can utilize the hierarchical framework to organize the development of effective instructional pedagogy for literacy development. This framework can help practicing and novice teachers teach students how to read and think critically. As a hierarchical structure, it supports building expertise with effective learner-centered literacy instruction. In addition, the hierarchical levels provide a clear structure for effectively differentiating instruction.
**Recommendations for Further Research**

The complexity of reading comprehension at high levels of academic work dictates that its development among young people will never be straightforward; there is no magical program that will work for all. As such, future research with the PARLI framework should be expanded to include students across the spectrum of reading proficiency, from special education populations, to typically developing readers, to advanced readers not currently being effectively challenged to continue to develop in this area of strength. Expanding the research into the PARLI framework should also include teachers of diverse backgrounds and teaching experiences. As technology becomes more readily available, incorporation of fMRI (or the next generation of this tool) to gain insight into what the brain is doing could be incorporated to further strengthen the framework.
References


EVALUATING THE PARLI FRAMEWORK


developmental assessments: Do they measure the same thing? *Cognitive Development, 18*(1), 61. doi:10.1016/S0885-2014(02)00162-4


doi:10.1177/105345129503100203


Boston, MA: Pearson/Allyn & Bacon.


document from the Center on Instruction. Portsmouth, NH: RMC Research Corporation, Center on Instruction.


Appendix A- Sample PARLA Assessment

**PARLA**

**NARRATIVE NONFICTION — ELEANOR ROOSEVELT**

A biography is the story of someone’s life, from birth to death. Biographies are generally full book length. This shortened biography of Eleanor Roosevelt emphasizes her role as a humanitarian, reformer and diplomat. If you are interested in learning more, your teacher can direct you to several full length biographies about her.

Brought up in a wealthy family and married to the president of the United States, Eleanor Roosevelt could have led a comfortable life filled with parties and entertainment. Instead she chose to work hard trying to help some of the poorest, most neglected people in American and working for peace and human rights around the world.

**Childhood**

Eleanor was the oldest of three children. She was plain, shy, and awkward recalling that, “I was always afraid of something...Anything I accomplished had to be done across a barrier of fear.” As a young lady being introduced to upper class society, she began volunteering and started to learn a great deal about people beyond her wealthy friends and family.

**The Early Years**

In 1905, Eleanor married her distant cousin, Franklin Delano Roosevelt. As Franklin’s political career continued to develop, so did Eleanor’s commitment to helping poor people. When Franklin was New York state senator and then Governor of the state of New York, Eleanor worked for better public housing, sanitation and health care programs. She helped pass laws to improve educational, living and working conditions for the women and children of their state. The life of a society lady did not appeal to Mrs. Roosevelt and she searched for worthwhile things to do. She began “drifting far afield from the old influences” and “thinking things out” for herself. She joined the
PARLA
NARRATIVE NONFICTION — ELEANOR ROOSEVELT
League of Women Voters. It was a new organization created to help women develop their political power. These determined women also worked for federal aid to education, international peace, and the end to child labor. With the help of her new friends in the League of Women Voters, Eleanor overcame her fear of public speaking and became an effective and popular speaker.

The Great Depression
Simply put, The Great Depression was a severe, worldwide, economic crisis during the 1930s in which millions of people lost their jobs, their savings, and their homes.

The First Lady
When FDR became president in 1933, Mrs. Roosevelt was determined to continue her work. She traveled around the country talking to all kinds of Americans who were still not being helped by relief programs. She encouraged people to write to her and hundreds of thousands did. She knew what was happening in the lives of regular people and supported them in every way she could. The First Lady wrote a daily column, gave speeches, wrote books and magazine articles, and talked on the radio (which was like TV is for us today). She did everything she could think of to share the problems she saw and encourage people to come up with creative solutions. She worked 16-20 hours each day.

Mrs. Eleanor Roosevelt had many firsts as a First Lady, or president’s wife. She was the first president’s wife to give all-female press conferences and travel by plane. Her job as Director of the Office of Civilian Defense meant that she was the first First Lady to hold and lose a government job. Mrs. Roosevelt appeared before a congressional committee, which no First Lady had ever done. As if that was not enough, she also wrote a syndicated, or national, newspaper column and gave a speech at a national convention. Eleanor became such a skilled speaker that she was also receive payment for lectures and to be a radio commentator, yet two more
PARLA
NARRATIVE NONFICTION — ELEANOR ROOSEVELT

firsts. Finally, Eleanor Roosevelt was the first wife of a president to publicly disagree with her husband's policies. She did all of this while raising five children too.

Death of a President and Work with the United Nations

Throughout her husband's presidency, from the Great Depression through World War II, Eleanor Roosevelt fought for human rights in the U.S. and around the world. After FDR's death in 1945, Harry Truman became president and the United Nations was formed. Truman appointed Eleanor to be one of the five U.S. delegates to the newly formed United Nations (UN). She did such a fine job, she was appointed to a more permanent position as the U.S. representative to the UN's governing body, the General Assembly. Next she became chairperson of the Human Rights commission of the UN. Mrs. Roosevelt resigned her official United Nations job after six successful years. She continued to work for justice for all as a private citizen until 1962 when she became ill and died.

"Ambition"

When Eleanor was fourteen, she wrote the following in an essay about ambition:

Of course it is easier to have no ambition and just keep on the same way every day and never try to do grand and great things, for it is only those who have ambition and who try to do who meet with difficulties and they alone face the disappointments that come when one does not succeed in what one has meant to do...

Is it best never to be known and to leave the world a blank as if one had never
PARLA
NARRATIVE NONFICTION — ELEANOR ROOSEVELT
References


### Level 1: Basic Stated Information

1. Who was Eleanor Roosevelt?

2. Who was Franklin Delano Roosevelt?

3. What happens in the essay?

### Confidence Ratings:

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I have no idea</td>
</tr>
<tr>
<td>2</td>
<td>I might be right</td>
</tr>
<tr>
<td>3</td>
<td>I'm almost sure</td>
</tr>
<tr>
<td>4</td>
<td>I know I'm right</td>
</tr>
</tbody>
</table>

What the confidence ratings mean:

- 1: I have no idea
- 2: I might be right
- 3: I'm almost sure
- 4: I know I'm right
Please answer the following questions using a single complete sentence in the space provided. Then, rate your confidence in your answer on the scale at the right (the meanings are listed at the bottom of the page).

**Level 2: Key Details**

4. Specifically, what was the Great Depression?

5. How did Eleanor react when Franklin got started in politics?

6. What was the first group that Eleanor joined that led her to so many other things?

7. What important job did Eleanor do after her husband died?

**Confidence Ratings:**

1. I have no idea
2. I might be right
3. I'm almost sure
4. I know I'm right
Level 3: Stated Relationships
8. When did Eleanor first start working with people not of her social class?

9. In an essay written when she was 14, why did Eleanor say that it is easier to have no ambition?

10. Who were the two groups of people that President Roosevelt helped because his wife influenced him to do so?

Confidence Ratings:

What the confidence ratings mean:

1. I have no idea
2. I might be right
3. I'm almost sure
4. I know I'm right
For the next set of questions, please answer with complete sentences, including an inference beyond exactly what the author says. Your answer should include your opinion and specific support from the text. Be sure that your answer is complete. Then, rate your confidence in your answer on the scale at the right (the meanings are listed at the bottom of the page).

**Level 4: Simple Implied Relationships**

11. How does Eleanor's essay on ambition predict the way she would spend the rest of her life?

**Confidence Ratings:**

11. Why was the way Eleanor chose to be First Lady so noteworthy?

**Confidence Ratings:**

What the confidence ratings mean:

1. I have no idea
2. I might be right
3. I'm almost sure
4. I know I'm right
For the next set of questions, please answer with complete sentences, including an inference beyond exactly what the author says. Your answer should include your opinion and specific support from the text. Be sure that your answer is complete. Then, rate your confidence in your answer on the scale at the right (the meanings are listed at the bottom of the page).

**Level 5: Complex Implied Relationships**

13. Although Eleanor was a shy, self-conscious girl, she was one of the great women of history. How do you think she was able to do this? Use evidence from the essay to support your opinion.

14. In what way might the lessons of Eleanor Roosevelt’s time spent as First Lady be helpful to First Lady Michelle Obama? Use evidence from the essay to support your opinion.

**What the confidence ratings mean:**

1. I have no idea  
2. I might be right  
3. I’m almost sure  
4. I know I’m right
For the next two questions, please re-read through the piece and think about the author’s opinions about these big ideas and how you know, then choose only one to answer. Your answer should include your opinion and specific support from the text. Be sure that your answer is complete. Then, rate your confidence in your answer on the scale at the right (the meanings are listed at the bottom of the page).

**Level 6: Author’s Generalization**

15. What does the author suggest about the generalization that perseverance is related to greatness? Support your opinion with evidence from the text.

16. What does the author convey about the responsibility of those who have wealth? Support your opinion with evidence from the text.

What the confidence ratings mean:

1. I have no idea
2. I might be right
3. I’m almost sure
4. I know I’m right
EVALUATING THE PARLI FRAMEWORK

PARLA
NARRATIVE NONFICTION — ELEANOR ROOSEVELT

For only one of the following questions, pretend you are a historian. Re-read through the piece and write a short essay response using evidence from the text to support your ideas. Then rate your confidence in your answer on the scale at the right (the meanings are listed at the bottom of the page).

Level 7: Structural Generalization
17. How does the presentation of information in this piece emphasize the role of leaders? Support your opinion with evidence from throughout the text.

18. How could the author ensure that the text is largely free of bias and truly a credible account of this significant figure of U.S. History? Support your opinion with evidence from throughout the text.

What the confidence ratings mean:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have no idea</td>
<td>I might be right</td>
<td>I'm almost sure</td>
<td>I know I'm right</td>
</tr>
</tbody>
</table>

Confidence Ratings:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix B MARSI Survey Measure

Metacognitive Awareness of Reading Strategies Inventory

Directions: Listed below are statements about what people do when they read academic or school-related materials such as textbooks or library books. Five numbers follow each statement (1, 2, 3, 4, or 5), and each number means the following:

1. means "I never or almost never do this."
2. means "I do this only occasionally."
3. means "I sometimes do this" (about 50% of the time).
4. means "I usually do this."
5. means "I always or almost always do this."

After reading each statement, circle the number (1, 2, 3, 4, or 5) that applies to you using the scale provided. Please note that there are no right or wrong answers to the statements in this inventory.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I have a purpose in mind when I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I take notes while reading to help me understand what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I think about what I know to help me understand what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I preview the text to see what it's about before reading it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. When text becomes difficult, I read aloud to help me understand what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I summarize what I read to reflect on important information in the text.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. I think about whether the content of the text fits my reading purpose.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I read slowly but carefully to be sure I understand what I'm reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I discuss what I read with others to check my understanding.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. I skim the text first by noting characteristics like length and organization.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. I try to get back on track when I lose concentration.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. I underline or circle information in the text to help me remember it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. I adjust my reading speed according to what I'm reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. I use reference materials such as dictionaries to help me understand what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. When text becomes difficult, I pay closer attention to what I'm reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. I use tables, figures, and pictures in text to increase my understanding.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. I stop from time to time and think about what I'm reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. I use context clues to help me better understand what I'm reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. I paraphrase (restate ideas in my own words) to better understand what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. I try to picture or visualize information to help remember what I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. I use typographical aids like boldface and italics to identify key information.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. I critically analyze and evaluate the information presented in the text.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. I go back and forth in the text to find relationships among ideas in it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. I check my understanding when I come across conflicting information.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. I try to guess what the material is about when I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. When text becomes difficult, I reread to increase my understanding.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. I ask myself questions I like to have answered in the text.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. I check to see if my guesses about the text are right or wrong.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. I try to guess the meaning of unknown words or phrases.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

developed by Kousider Mokhtari and Carla A. Reichard, Copyright © 2002 Kousider Mokhtari, used with permission.
# EVALUATING THE PARLI FRAMEWORK

## Metacognitive Awareness of Reading Strategies Inventory

Student name: ____________________________  Age: ________  Date: ____________

1. Write your response to each statement (i.e. 1, 2, 3, 4, or 5) in each of the blanks.
2. Add up the scores under each column. Place the result on the line under each column.
3. Divide the subscale score by the number of statements in each column to get the average for each subscale.
4. Calculate the average for the whole inventory by adding up the subscale scores and dividing by 30.
5. Compare your results to those shown below.
6. Discuss your results with your teacher or tutor.

<table>
<thead>
<tr>
<th>Global Reading Strategies (GLOB subscale)</th>
<th>Problem-Solving Strategies (PROB subscale)</th>
<th>Support Reading Strategies (SUP subscale)</th>
<th>Overall Reading Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>8.</td>
<td>2.</td>
<td>GLOB</td>
</tr>
<tr>
<td>3.</td>
<td>11.</td>
<td>5.</td>
<td>PROB</td>
</tr>
<tr>
<td>4.</td>
<td>13.</td>
<td>6.</td>
<td>SUP</td>
</tr>
<tr>
<td>7.</td>
<td>16.</td>
<td>9.</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>18.</td>
<td>12.</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>21.</td>
<td>15.</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>27.</td>
<td>20.</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td></td>
<td>28.</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **GLOB score**: 
- **PROB score**: 
- **SUP score**: 
- **GLOB mean**: 
- **PROB mean**: 
- **SUP mean**: 

**Key to averages**: 3.5 or higher = high  
2.5-3.4 = medium  
2.4 or lower = low

**Interpreting your scores**: The overall average indicates how often you use reading strategies when reading academic materials. The average for each subscale of the inventory shows which group of strategies (i.e., global, problem-solving, and support strategies) you use most when reading. With this information, you can tell if you score very high or very low on any of these strategy groups. Note, however, that the best possible use of these strategies depends on your reading ability in English, the type of material you read, and your purpose for reading it. A low score on any of the subscales or parts of the inventory indicates that there may be some strategies in these parts that you might want to learn about and consider using when reading.

Developed by Kouider Mokhtari and Carla A. Reichard, Copyright © 2002 Kouider Mokhtari, used with permission.
EVALUATING THE PARLI FRAMEWORK

Appendix C - MAASC Survey Measure

Self-Concept and Motivation
Middle School Assessment of Academic Self-Concept and Motivation by Janney A. Gordon, Copyright © 2003, and the assessment of Personal Goals (APG), Copyright © 2007, Mark A. Ford.

Adapted with permission from the assessment of Academic Self-Concept and Motivation by Janney A. Gordon, Copyright © 2003, and the assessment of Personal Goals (APG), Copyright © 2007, Mark A. Ford.

Your Emotions = How do I feel about it?
Category A
The Importance of Value = Is it important to me?
Category B
Your Control = Can I control it?
Category C
Environmental Responsiveness or Opportunity = Does school help me do it?
Category D
Your Ability or Capability = Can I do it?
Category E

and those in the yellow box for Category:
For each category, you will choose the statement that matches your beliefs and experience best. Use the statements in the blue box for categories A through D.

The Middle School Assessment of Academic Self-Concept and Motivation is divided into four sections of goal statements and five categories which address the environment in mind as you complete it.

School or academic experience, your experiences you may have at home, in your neighborhood, or at work. Please keep your school activities, experiences, and

Katherine Berdjick

3/15/2010

The Middle School Assessment of Academic Self-Concept and Motivation is a questionnaire about your beliefs. This questionnaire refers specifically to your
<table>
<thead>
<tr>
<th>General Cognitive Goals</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Understanding the materials/information taught in my class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Coursework assigned by my teacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Performance well on tests which cover class material</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Being good grades in my courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Reading cognitive goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How do I feel about these goals?</th>
<th>E</th>
<th>D</th>
<th>C</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very happy, excited, positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slightly positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somewhat positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somewhat negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very negative, disappointed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use these ratings for columns A-D.
The image contains a table and a chart, which appear to be part of a larger document. The table is titled "Middle School Assessment of Academic Self-Concept and Motivation by Kimberly A. Gordon House, Copyright © 2003, and the assessment is adapted with permission from the assessment of Academic Self-Concept and Motivation by Kimberly A. Gordon House, Copyright © 2003.

The chart includes two columns: "Personal Goals" and "Social Goals." Each goal has a rating scale from 1 to 5, and there are descriptors for each rating level. The chart also includes a section for columns A-D, which seems to be related to the assessment.

The text on the right side of the image is not clearly visible due to the angle and quality of the image.
Appendix D Permissions

FRANCIS HOWELL SCHOOL DISTRICT
4546 Central School Road • St. Charles, MO 63304-7113
Phone: 636-851-4000 • Fax: 636-851-4093 • www.fhsdschools.org

April 12, 2010

Katie Bradarich
48 Piney Ridge Court
Saint Charles, MO 63304

Dear Katie,

You have permission to conduct a research project at Barnwell Middle School on Middle School Reading Comprehension. Your study will provide the District with valuable information in research to determine the methods to differentiate instruction to improve the reading comprehension of the middle school age reader. You have permission to use Barnwell Middle School archival data from August 12, 2010 to April 29, 2011 as well as to conduct surveys of students in your classes during their time enrolled in your class. During our conversation, you indicated that you will be following the Francis Howell School District curriculum and you will minimize the risk to the students and staff from their participation in your surveys/interviews and will keep all personally identifiable information confidential.

I wish you the best on your study which promises to provide valuable research information to the field of education in general and to the Francis Howell School District in particular.

Sincerely,

[Signature]

Dr. Renée Schuster, Superintendent
Francis Howell Schools
Informed Consent for Child Participation in Research Activities
Evaluating the Pragmatic Analytical Reading Level Instructional (PARLI) Framework

Participant ____________________________________________

Principal Investigator Katherine Bradarich PI’s Phone Number 636-851-4155

1. Your child is invited to participate in a research study conducted by Katherine Bradarich, and Gwendolyn Turner. The purpose of this research is to see how effective a method of teaching reading to students who struggle is with 8th graders. Your child will be participating in all regular reading classroom activities. The researcher is asking for permission to study the regular work that your son or daughter does as part of reading class, and the progress he or she makes in building reading comprehension. As part of the regular reading classroom activities, students will be video and audio recorded to evaluate reading fluency and performance. These recordings will be viewed by the teacher and the students, and available for parents to view on request. According to APA guidelines, these recordings will be kept in a locked, secure location for five years before being destroyed.

2. Approximately ten subjects will be involved in this research. The amount of time involved in your child’s participation will be the regularly scheduled reading class each day, for the fall semester, 2010.

3. Consenting for your child to take part in this study means that the researcher will be closely evaluating the regular work your child does in reading and how he or she changes and progresses with reading comprehension over the course of the study. If you don’t want your child’s information to be used in this study, you don’t have to let me use it. Allowing me to study your child’s work for this study is up to you, and no one will be upset if you don’t want me to do it or if you change your mind later and want me to stop examining your child’s work for research purposes. Your child’s grade will not be affected whether his or her information is included in the study or not.

4. We will do everything we can to protect your child’s privacy. As part of this effort, your child’s identity will not be revealed in any publication or presentation that may result from this study. In rare instances, a researcher’s study must undergo an audit or program evaluation by an oversight
agency (such as the Office for Human Research Protection). That agency would be required to maintain the confidentiality of your child’s data.

5. If you have any questions or concerns regarding this study, or if any problems arise, you may call the Investigator, Katherine Bradarich at 636-851-4155 or the Faculty Advisor, Gwendolyn Turner at 314-516-5096. You are also welcome to email me at katie.bradarich@fhsdschools.org. You may also ask questions or state concerns regarding your child’s rights as a research participant to the Office of Research Administration, at 516-5897.

I have read this consent form and have been given the opportunity to ask questions. I will also be given a copy of this consent form for my records. I consent to my child’s participation in the research described above.

________________________________________
Parent’s/Guardian’s Signature         Date

________________________________________
Parent’s/Guardian’s Printed Name

________________________________________
Child’s Printed Name

________________________________________
Signature of Investigator or Designee     Date

________________________________________
Investigator/Designee Printed Name
Division of Teaching and Learning

To be read to students.

Assent to Participate in Research Activities (Minors)
Evaluating the Pragmatic Analytical Reading Level Instructional (PARLI) Framework

1. My name is Mrs. Bradarich.

2. I am asking you to take part in a research study because we are trying to learn more about teaching middle school students to better understand what they read.

3. If you agree to be in this study, you will be working on becoming a stronger reader in several different kinds of reading, as you have in reading class before. You will be doing all of the regular things we do as part of a reading class. I will be carefully studying the work you do and how you are growing as a reader.

4. You might find being in this study helps you become a stronger reader in one or more subject areas in school. You might also find that this experience teaches you something about how you learn best that you can use in other situations.

5. If you don't want your information to be used in this study, you don't have to let me use it. Remember, allowing me to study your work for this study is up to you, and no one will be upset if you don't want me to do it or if you change your mind later and want me to stop examining your work. Your grade will not be affected whether your information is included in the study or not.

6. You can ask any questions that you have about the study. If you have a question later that you didn't think of now, you can call me at 636-851-4155, or send me an email at katie.bradarich@fhsdschools.org.

7. Signing your name at the bottom means that you agree to let me use your reading work in this study. You will be given a copy of this form after you have signed it.

Participant’s Signature  Date  Participant’s Printed Name

Participant’s Age  Grade in School

Submit an original and 1 copy of this application, with attachments (number all pages), to the Office of Research Administration, 341 Woods Hall.
Appendix E Initial Codes

Codes preparing for deeper analysis Jan 16, 2011

Code Families

HU: DISSENTATION MERGED HU
File: C:\Users\Katie\Documents\Scientific Software\ATLAS.ti\TextBank\DISSENTATION MERGED HU.hpr6
Edited by: Super
Date/Time: 01/16/2011 12:08:00 PM

Code Family: Annotation and Response Codes
Created: 10/27/2010 06:19:57 AM (Super)
Codes (32):
1. [ANNO L1- naked]
2. [ANNO L3-naked]
3. [ANNO L3-symbol]
4. [ANNO L4- naked]
5. [ANNO L4-symbol]
6. [ANNO L4 - note]
7. [ANNO L5-naked]
8. [ANNO misc-naked]
9. [ANNO misc-note]
10. [IV_1 think she waited so long bec...
11. [QUES- given_1]
12. [QUES- weak]
13. [QUES-no fit]
14. [QUES - comp]
15. [QUOTES-complete_1]
16. [QUOTES-weak]
17. [QUOTES - weak]
18. [RES- L1 No EXP-]
19. [RES- L3 explain_1]
20. [RES- L3 miss]
21. [RES- wrong]
22. [RES-L3 relationship_1]
23. [RES-L3_1]
24. [RES-L4 explain_1]
25. [RES - L1_1]
26. [RES L1- explain_1]
27. [RES L2-NOT_1]
28. [RES L4_1]
29. [RUB - match_1]
30. [TITLE- fit_1]
31. [TITLE- ok_1]
32. [TITLE- weak]

Quotation(s): 435

Code Family: Conceptual Thinking of Levels
Created: 11/06/2010 07:40:37 PM (Super)
Codes (10):
1. [Conc Comp L3 vs L4 -]
2. [Conc Comp L3 vs L4 +]
3. [Conc Comp L3 vs L4 ok]
4. [Conc EX -] [Conc EX +]
5. [Conc EX ok]
6. [Conc L3 +]
7. [Conc L3 ok]
8. [Conc L4 +]
9. [Conc L4 ok]

Quotation(s): 49
### Codes preparing for deeper analysis Jan 16, 2011

**Code Family: IN-VIVO CODES**

*Created: 01/16/2011 12:06:12 PM (Super)*

**Codes (59):**

1. [IV_ Scared because it's a Monday a..]
2. [IV_ tired and kinda focused]
3. [IV_ tired and sleep]
4. [IV_Also I didn't understand some ..]
5. [IV_and it's different than fiction..]
6. [IV_answers did not make sense and..]
7. [IV_bad because I don't really lik..]
8. [IV_because it was nonfiction]
9. [IV_because it's a test and its fo..]
10. [IV_because some o' the questions ..]
11. [IV_but I was getting on the right..]
12. [IV_But very sleepy..]
13. [IV_Expo isn't the same as fiction..]
14. [IV_hard comprehending nonfiction ..]
15. [IV_I am not very good with non-fi..]
16. [IV_I answered another question.]
17. [IV_I couldn't really find out wer..]
18. [IV_I couldn't understand the ques..]
19. [IV_I did not answer because: I di..]
20. [IV_I did really good on the last ..]
21. [IV_I didn't do it I didn't unders..]
22. [IV_I didn't had enough time!]
23. [IV_I didn't understand the story ..]
24. [IV_I don't really read non-fiction..]
25. [IV_I got confused on a lot of thi..]
26. [IV_I just couldn't think about it..]
27. [IV_I think I am better at fiction..]
28. [IV_I think I did good for non-fic..]
29. [IV_I think I did ok wish I did a ..]
30. [IV_I think its right because nonf..]
31. [IV_I think she waited so long bec..]
32. [IV_I thought it was too hard to a..]
33. [IV_I was frustrated with all the ..]
34. [IV_I was too specific in my answ..]
35. [IV_I'm inferring that the judges ..]
36. [IV_I'm not happy at all I really ..]
37. [IV_I'm not into all the history s..]
38. [IV_IDK Big words.]
39. [IV_It said it right there in the ..]
40. [IV_It was nonfiction & I'm not th..]
41. [IV_It was nonfiction is hard for ..]
42. [IV_It was way different than fict..]
43. [IV_It's not like fiction and non-..]
44. [IV_Little scared because I'm not ..]
45. [IV_Native Peoples of the Southea..]
46. [IV_negotiations negaions, negata..]
47. [IV_Nervous because it is about An..]
48. [IV_nonfiction isn't like fiction..]
49. [IV_Now that I see the correct ans..]
50. [IV_presedent]
51. [IV_presdentent]
52. [IV_seniter]
53. [IV_So they won't fell lost]
54. [IV_the class ended and I didn't h..]
55. [IV_the questions were straight fr..]
56. [IV_They all look at each other in..]
57. [IV_they are discussing Nujoood's w..]
58. [IV_What was Roosevelt's great..]
59. [IV_Who were the two groups of..]

**Quotation(s):** 61
Codes preparing for deeper analysis Jan 16, 2011

**Code Family: DISCUSSION CODES**

Created: 12/13/2010 08:49:08 PM (Super)

**Codes (42):**

1. [ARIEL]
2. [ASHLEY]
3. [DALTON]
4. [DISC- acknowledge]
5. [DISC- agree/continue]
6. [DISC- cite pg]
7. [DISC- connect]
8. [DISC- contradict]
9. [DISC- correct]
10. [DISC- cycle back]
11. [DISC- details]
12. [DISC- IDK]
13. [DISC- interrupt]
14. [DISC- Lead]
15. [DISC- no reply]
16. [DISC- opinion/judgment]
17. [DISC- procedure]
18. [DISC- question]
19. [DISC- read]
20. [DISC- REF ANNO]
21. [DISC- respond]

22. [DISC- shift]
23. [DISC- social]
24. [DISC- specific Level]
25. [DISC- TEACH ask]
26. [DISC- TEACH guide]
27. [DISC- unknown words]
28. [DISC-asking TEACHER]
29. [DISC-inferring-S]
30. [DISC-pronunciation]
31. [DISC-Teach Dir]
32. [DISC-Teach Respond]
33. [EMILY]
34. [IV_They all look at each other in..]
35. [IV_they are discussing Nujood's w.]
36. [KAYLA]
37. [MATT D]
38. [MATT H]
39. [NANCY]
40. [RUBI]
41. [SAMMY]
42. [Unknown: Where are we?]
Codes preparing for deeper analysis Jan 16, 2011

**Code Family: PARLA codes**

Created: 11/06/2010 07:33:46 PM (Super)

**Codes (24):**

1. [P-Inference weak]
2. [P-Not Responsive]
3. [P-Awk]
4. [P-Confidence Mismatch -]
5. [P-Confidence Mismatch +]
6. [P-Confidence NOT rated]
7. [P-essence]
8. [P-Evidence]
9. [P-Incomplete]
10. [P-Incorrect]
11. [P-Level 1 match]
12. [P-Level 2 match]
13. [P-Level 3 match]
14. [P-Level 4 match]
15. [P-Level 5 match]
16. [P-Not Attempted]
17. [P-Not Detailed]
18. [P-Not in text]
19. [Time= 12:45-WPM=< 81 = < 2nd grade]
20. [WPM = 109-130 = 2nd grade]
21. [WPM = 88 = 2nd grade]
22. [WPM => 2X lowest]
23. [WPM = 148-161 (5)]
24. [WPM = 82-108= 2nd grade]

**Quotation(s):** 941
EVALUATING THE PARLI FRAMEWORK

Codes preparing for deeper analysis Jan 16, 2011

<table>
<thead>
<tr>
<th>Code Family: Reflection codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Created: 11/06/2010 07:35:57 PM (Super)</td>
</tr>
</tbody>
</table>

**Codes [S1]:**

1. [ASSESS- REF- challenge]
2. [ASSESS- REF- Confident]
3. [ASSESS- REF- Diff]
4. [ASSESS- REF- Easy]
5. [ASSESS- REF- hierarchy]
6. [ASSESS- REF- Improve]
7. [ASSESS- REF- nervous]
8. [ASSESS- REF- neutral]
9. [ASSESS- REF- no comment]
10. [ASSESS- REF- not confident]
11. [ASSESS- REF- Overall -]
12. [ASSESS- REF- Overall +]
13. [ASSESS- REF- Overall NO COMM]
14. [ASSESS- REF- specific reason]
15. [ASSESS- REF- understanding]
16. [ASSESS- REF- just right]
17. [ASSESS- REF- Level]
18. [ASSESS- REF - disappoint]
19. [ASSESS- REF - optimism]
20. [IV- the class ended and I didn't h...]
21. [IV- because it was nonfiction]
22. [IV- I did not answer because I di...]
23. [IV- I think its right because nonf...]
24. [IV- Now that I see the correct ans...]
25. [REF- attribute of self]
26. [REF- confidence higher than merit]
27. [REF- confidence lower than merit]
28. [REF- did not know answer]
29. [REF- did not respond to question but no why]
30. [REF- Did not understand question]
31. [REF- example longer]
32. [REF- example more detail]
33. [REF- example said it differently]
34. [REF- incorrect answer]
35. [REF- Instrument & self match]
36. [REF- Instrument higher than self]
37. [REF- kind of different]
38. [REF- misunderstanding instrument]
39. [REF- read question wrong]
40. [REF- reflection some specifics]
41. [REF- simple reflection on behavior]
42. [REF- specific]
43. [REF- statement]
44. [REF- wrong answer]
45. [REF- confidence higher than merit]
46. [REF- misread question]
47. [REF- no confidence marked]
48. [REF- no reflection]
49. [REF- Not different]

Quotation[s]: 856
Codes preparing for deeper analysis Jan 16, 2011

**Code Family: RESPONSE CODES**

Created: 12/13/2010 08:48:16 PM (Super)

**Codes (51):**

1. [ARIEL & MATT D]
2. [ASHLEY & SAMMY]
3. [DALTON & NANCI]
4. [EMILY & MATT H]
5. [QUES- comp]
6. [QUES- given]
7. [QUES- no fit]
8. [QUES-miss]
9. [QUES-ok]
10. [QUOTES- complete]
11. [QUOTES- Inc]
12. [QUOTES- comp]
13. [QUOTES- L4 Not Int]
14. [QUOTES- no cite]
15. [QUOTES-weak]
16. [QUOTES- miss]
17. [QUOTES- not fit]
18. [RES- L1 No EXP]
19. [RES- L2 No EXP]
20. [RES- L3-opinion]
21. [RES- L3 explain]
22. [RES- L3 strong opinion]
23. [RES- L4 INC Infer]
24. [RES- L4 opinion]
25. [RES- L5- opinion]
26. [RES- L5 explain]

Quotation(s): 448
EVALUATING THE PARLI FRAMEWORK

Codes preparing for deeper analysis Jan 16, 2011

Code Family: Think-Aloud Codes
Created: 11/06/2010 07:31:10 PM (Super)

Codings (23):
1. A-Conscious Inference
2. A-Judgment/Opinion
3. A-Questioning
4. A-Reflecting
5. A-Summarizing
6. D-connect
7. D-Conscious Inference
8. D-Paraphrase/interpret
9. D-Pause/Reflect
10. D-Predict
11. D-questioning
12. D-Identifying important info
13. D-Integrating parts of text
14. D-Judgment/Opinion
15. D-Miscue
16. D-Repeat/Restate
17. Decoding trouble not addressed
18. [MA- Monitoring decisions]
19. [MD-Activate Process]
20. [MD-Activate processing]
21. [MD-Meaningful process]
22. [MD-Recognize Prob]
23. [MD-Text Characteristics]

Quotation(s): 892
## Appendix F Definition of Codes and Coding Rules

<table>
<thead>
<tr>
<th>Definition</th>
<th>Coding Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level One is a dimensional code for LEVEL reflection the first Level of comprehension in Hillocks' hierarchy. This is a Literal Level of comprehension.</td>
<td>Level One is Basic Stated Information which deals with important, usually both prominent and repeated obvious information that the reader is able to demonstrate he/she understands/notices when reading the text. Also included in Level One are the open codes from the Think-Alouds that refer to the most basic of reading/decoding strategies and actions. It includes pronunciation, miscues, and substitutions.</td>
</tr>
<tr>
<td>Level Two is a dimensional code for LEVEL reflection the second Level of comprehension in Hillocks' hierarchy. This is a Literal Level of comprehension.</td>
<td>Level Two is about Key Detail, which are those details that are important to the twists and turns of the plot that the reader is able to demonstrate he/she understands/notices when reading the text.</td>
</tr>
<tr>
<td>Level Three is a dimensional code for LEVEL reflection the third Level of comprehension in Hillocks' hierarchy, and the last Literal Level of comprehension.</td>
<td>Level Three is about Stated Relationships which are those wherein the author does specifically state the relationship between at least two pieces of information in the text, and the relationship is usually only stated once in the text. Understanding at this Level means that the reader is able to demonstrate he/she understands/notices things that the author explicitly states go together, as well as the nature of that relationship (yin/yang, cause/effect, co-occurrence) when reading the text.</td>
</tr>
<tr>
<td>Level Four is a dimensional code for LEVEL reflection the fourth Level of comprehension in Hillocks' hierarchy, and the first Inferential Level of comprehension</td>
<td>Level Four is about Simple Implied Relationships. This Level involves parts of the text where the author hints at a relationship between two pieces of information, but the reader must take into account denotative and connotative clues, and relate them to prior knowledge or experience, to successfully make the inference. Understanding at this Level means that the reader is able to demonstrate he/she can make simple and reasonable inferences about things that the author does not explicitly address when reading the text.</td>
</tr>
<tr>
<td>Definition</td>
<td>Coding Rules</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
</tr>
<tr>
<td>text. Several open codes from the T-A protocols were subsumed within this code during axial coding and include pausing and reflecting, making conscious inferences at a simple level, and making simple connections to the world outside the text.</td>
<td>Level_Five is about Complex Implied Relationships which are relationships between and among a large number of details over a major portion of the text. This contrasts with Level 4 because it is possible that these details may simultaneously imply more than one pattern or set of relationships, and it usually involves a pattern that develops throughout the text. Understanding at this Level means that the reader is able to demonstrate he/she can develop an understanding of a pattern, supported by evidence and reasonable inferences about growth, change, and development over the course of the text. The author does not state these conclusions, but implies them in a variety of ways that the reasonable and competent reader at this Level will infer. At the stage of axial coding, earlier codes of integrating parts of a text and integrating ideas and concepts across texts, aka more complex inferential thinking, from T-A protocols were folded into this dimensional code of Product.</td>
</tr>
<tr>
<td>Level_Five is a dimensional code for LEVEL reflection the fifth Level of comprehension in Hillocks' hierarchy. It is the second Inferential Level of comprehension, and involves significant abstract and complex thinking.</td>
<td></td>
</tr>
</tbody>
</table>

| Level_Five is a dimensional code for LEVEL reflection the fifth Level of comprehension in Hillocks' hierarchy. It is the second Inferential Level of comprehension, and involves significant abstract and complex thinking. | |

| Level_Five is about Complex Implied Relationships which are relationships between and among a large number of details over a major portion of the text. This contrasts with Level 4 because it is possible that these details may simultaneously imply more than one pattern or set of relationships, and it usually involves a pattern that develops throughout the text. Understanding at this Level means that the reader is able to demonstrate he/she can develop an understanding of a pattern, supported by evidence and reasonable inferences about growth, change, and development over the course of the text. The author does not state these conclusions, but implies them in a variety of ways that the reasonable and competent reader at this Level will infer. At the stage of axial coding, earlier codes of integrating parts of a text and integrating ideas and concepts across texts, aka more complex inferential thinking, from T-A protocols were folded into this dimensional code of Product. | |
### Definition

The Quality dimension of *Product* addresses the nature of a particular product along a continuum and further delineates the characteristics of responses at different Levels where both dimensions are applicable, and also describes the *Aspects* such as reflections or words per minute.

### Coding Rules

| Strong: The dimension of “strong” is for *Products* or *Aspects* that leave no question about the student's ability to think at this Level, communicate this characteristic clearly, demonstrate support for their thinking from the text, express confident and positive opinions and beliefs; in a word, mastery. | QUAL_strong is for a response that leaves no question about the student's ability to think at this Level and demonstrate support for that thinking from the text. Mastery
Subsumes Open Codes of strong, positive, opinion, confident etc |
| Good: The dimension “good” shows positive characteristics, but just a slightly lesser level than does “strong”; in a word, proficiency. | QUAL_good is for a response that shows specific and detailed thinking or support for thinking, but mastery is not unequivocal. Proficient.
Subsumes Open Codes of note, specific, details |
<table>
<thead>
<tr>
<th>Definition</th>
<th>Coding Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK: Novice-level, less sophisticated thinking and products are coded with “ok” to indicate that evidence of the students’ mastery is not unequivocal, but is certainly suggested.</td>
<td>QUAL_ok is for a response that has the essence of what is needed, but is not as detailed and specific as needed to be sure that the student truly understands at this Level and can support that understanding. Novice. Subsumes Open Codes of: essence, ok, neutral, detailed, incomplete TAD Repeat/Restate</td>
</tr>
<tr>
<td>Miss: “Miss” reflects Quality that is defined as the absence of an expected <em>Aspect</em>.</td>
<td>QUAL_miss is a response or reflection, on assessment or class work, that is missing an Aspect of a complete and well done response or reflection. Missing an entire Aspect. Includes Open Codes of miss or no comment</td>
</tr>
<tr>
<td>Weak: The Quality dimension “weak” is one that begins to hint at quality, or is of a very basic nature.</td>
<td>QUAL_weak is when the student makes an unsupported statement or annotates with no further reflection or comment. Thinking is not apparent; student may or may not be competent at this Level, but one cannot tell from this response. Basic Subsumes Open Codes of: naked/no support Decoding trouble not addressed</td>
</tr>
<tr>
<td>Wrong: The Quality dimension of <em>Product</em> of “wrong” labels a present, but incorrect response.</td>
<td>QUAL_wrong is for incorrect, ridiculous, nonsequitor, not from the text responses that do not make sense. Below Basic Subsumes Open Codes of: nonsense, not in text, incorrect, wrong, random</td>
</tr>
<tr>
<td><strong>Definition</strong></td>
<td><strong>Coding Rules</strong></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Anno: Is the Aspect subcategory that encompasses annotation of text, be it a written or verbal aside or comment</td>
<td>ASP_anno is for the action/interaction of annotating a text.</td>
</tr>
<tr>
<td>Answer: Is the Aspect subcategory that encompasses the action/interaction of the student providing an answer/response that just covers the basic response, no explanation or &quot;because&quot; to support it.</td>
<td>ASP_answer is for the action/interaction of the student providing a brief and correct answer/response that just covers the basic response.</td>
</tr>
<tr>
<td>Explain: Is the Aspect subcategory that encompasses the action/interaction part of the response in which the student explains his or her thinking about the response and/or the importance of the segment of text that he or she is responding to.</td>
<td>ASP_explain is for the action/interaction part of the response in which the student explains. It may also be used to code the entire response if it does not make sense to separate out the answer and the explanation.</td>
</tr>
<tr>
<td>Ques: Is the Aspect subcategory that encompasses the action/interaction of creating a question that will solicit a response at the desired Level. The ability to do this effectively is a key indicator of mastery at that Level.</td>
<td>ASP_ques is for the infrequent action/interaction of the student creating his own question as part of a complete response to the text. This code is not used when the student repeats a provided question as part of the response.</td>
</tr>
<tr>
<td>Quote: Is the Aspect subcategory that encompasses the action/interaction of supporting the thinking at the desired Level with a quote drawn from the text. The ability to do this effectively is a key indicator of developing proficiency/competence at that Level.</td>
<td>ASP_quote is for the action/interaction that is clearly from the actual text. This code is used whether or not the quote supports the thinking, and is a coding of the presence of said quote. Quality codes address the appropriateness/strength of it.</td>
</tr>
<tr>
<td>Definition</td>
<td>Coding Rules</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Ref: Is the Aspect subcategory that encompasses the student reflecting about his performance.</td>
<td>This is a more specialized code and applies only to particular Contexts and Products.</td>
</tr>
<tr>
<td></td>
<td>ASP_Ref is for the reflection of the student about his or her performance. They may reflect about their anticipated performance overall, their performance on a particular question, or their performance after they have looked at their scores.</td>
</tr>
<tr>
<td></td>
<td>If it is a reflection about a particular response, the Level of the response is coded also.</td>
</tr>
<tr>
<td></td>
<td>ALL Reflection codes also have PAB codes.</td>
</tr>
<tr>
<td>Title: Is the Aspect subcategory that encompasses the action/interaction of creating a title for the response that captures the essence of the response. The ability to do this effectively is an indicator of competence at that Level.</td>
<td>ASP_title is for the generation of an original title to label the response.</td>
</tr>
<tr>
<td></td>
<td>This code is used whether or not the title supports the thinking, and is a coding of the presence of said title.</td>
</tr>
<tr>
<td></td>
<td>This code is not used if the “title” consists of a label of the Level and Stage of the response, the title of the piece, or both of these.</td>
</tr>
<tr>
<td></td>
<td>Quality codes address the appropriateness/strength of it.</td>
</tr>
<tr>
<td>WPM: Is the Aspect subcategory that encompasses rate of reading.</td>
<td>This is a more specialized and applies only to particular Contexts and Products.</td>
</tr>
<tr>
<td></td>
<td>ASP_WPM= words per minute that the student reads the piece. This is only calculated in an assessment and solo written setting, so no CON coding is needed.</td>
</tr>
<tr>
<td>Context refers to the conditions in which the actions/interactions (Aspect and PAB) and consequences (Product-Quality and Level) happen.</td>
<td>Data is most often coded with more than one Context.</td>
</tr>
<tr>
<td>Definition</td>
<td>Coding Rules</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CON_assess is for the context of</td>
<td>For the actions/interactions and consequences that are specifically generated</td>
</tr>
<tr>
<td>assessment- the student is</td>
<td>in an assessment context, Con_assess is used. This may be paired with Con</td>
</tr>
<tr>
<td>responding during an assessment.</td>
<td>For those that are not generated during assessment, there is no code.</td>
</tr>
<tr>
<td></td>
<td>For the actions/interactions and consequences that are specifically</td>
</tr>
<tr>
<td></td>
<td>generated in a context of more than one student working together, Con_Collab</td>
</tr>
<tr>
<td></td>
<td>is used.</td>
</tr>
<tr>
<td>CON_collab is for the context of</td>
<td>For the actions/interactions and consequences that are specifically</td>
</tr>
<tr>
<td>students working in collaboration</td>
<td>generated in a context of an individual student working independently, Con_Solo</td>
</tr>
<tr>
<td>with others to develop a response</td>
<td>is used.</td>
</tr>
<tr>
<td>to text.</td>
<td>For the actions/interactions and consequences that are specifically</td>
</tr>
<tr>
<td></td>
<td>generated in a context of the student speaking aloud, Con_Talk is used.</td>
</tr>
<tr>
<td>CON_solo is for the context of</td>
<td>For the actions/interactions and consequences that are specifically</td>
</tr>
<tr>
<td>students working in isolation/</td>
<td>generated when the student is writing, Con_Write is used.</td>
</tr>
<tr>
<td>independently to respond to text.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Definition</td>
<td>Coding Rules</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PAB is the code for Personal Agency Beliefs, used as a measure of the confidence the Case has in his or her ability to be successful in a particular context.</td>
<td>PAB coding varies by Context. PAB Subcategories for assessments (Assess) are a combination of the points achieved and the confidence rating of the student regarding his performance. PAB Subcategories for the remaining Context codes are based on the language and observed behavior of the students.</td>
</tr>
</tbody>
</table>
| PAB_1-Robust is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a positive context and a strong belief in one's capability. There is an expectation that goals will be achieved, even in the face of obstacles, difficulties and failures. Negative outcomes are temporary. | For Con_Assess:  
For Levels 1-3: 1 point & C=3/4  
For Levels 4 & up: 2 points & C=3/4  
For all other Contexts: Language and actions that convey messages such as,  
"I can do anything if I put my mind to it," and  
"I'm not going to let this little thing get me down/stop me/discourage me" get coded as Robust. |
| PAB_2-Tenacious is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a neutral or variable context and a strong belief in one's capability. Some degree of environmental unresponsiveness is viewed as predictable and unsurprising and effortful persistence is present. | For Con_Assess:  
For Levels 1-3: 1/2 point & C=3/4  
For Levels 4 & up: 1 points & C=3/4  
For all other Contexts: Language and actions that convey messages such as,  
"This isn't easy, but I can do it if I try," and  
"I won't let this tough situation get to me," get coded as Tenacious. |
<table>
<thead>
<tr>
<th>Definition</th>
<th>Coding Rules</th>
</tr>
</thead>
</table>
| **PAB_3-Modest** is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a positive context and a moderate or variable belief in one's capability. This is generally a green light pattern. The self is more fallible than the context and the context is a source of strength rather than an obstacle. | **For Con_Assess:**  
For Levels 1-3: 1 point & C=2  
For Levels 4 & up: 2 points & C=2  
For all other **Contexts:** Language and actions that convey messages such as,  
"If I get a little help, I can do it," and "If I set reasonable goals, I'll be ok" get coded as Modest. |
| **PAB_4-Vulnerable** is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a neutral or variable context and a moderate or variable belief in one's capability. There is uncertainty or vacillation between favorable and unfavorable goal expectations. May get some anxiety, worry, and be cautious about goals, but not negative enough to inhibit growth toward important goals. | **For Con_Assess:**  
For Levels 1-3: 1/2 point & C=2  
For Levels 4 & up: 1 point & C= 2  
For all other **Contexts:** Language and actions that convey messages such as,  
"Sometimes I think I can do it, others I fear the worst," and "I try to avoid letting anyone know I can't do it" get coded as Vulnerable. |
| **PAB_5-Accept** is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a negative context and a strong belief in one's capability. There is a significant degree of distrust of the environment and a blaming of the context for the task performance. Accepting lacks a pattern of anger about it. | **For Con_Assess:**  
For Levels 1-3: 0 point & C=3/4  
For Levels 4 & up: <1 points & C= 3/4  
Needs to have comment that fits in reflection.  
For all other **Contexts:** Language and actions that convey messages such as,  
"I know I can't stop this bad result, but I can figure out when it will come," and "All I can do is accept it and try not to think about it" get coded as Accept. |
<table>
<thead>
<tr>
<th>Definition</th>
<th>Coding Rules</th>
</tr>
</thead>
</table>
| **PAB_6**-Antagonistic is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a negative context and a strong belief in one's capability. There is a significant degree of distrust of the environment and a blaming of the context for the task performance. Antagonistic features a pattern of anger about it. | For **Con_Assess**:
For Levels 1-3: 0 point & C=3/4
For Levels 4 & up: <1 points & C= 3/4
Needs to have comment that fits in reflection.

For all other **Contexts**: Language and actions that convey messages such as,
"I'm not quitting until I make this darn thing work!" and "I'll show those who don't think I can do it!" get coded as Antagonistic. |
| **PAB_7**-Fragile is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a positive context and a weak belief in one's capability. The general belief in the context's adequacy combined with a significant degree of self-deprecactions and self-devaluation. People generally blame themselves rather than the context for problems and failures. | For **Con_Assess**:
For Levels 1-3: 1 point & C=1
For Levels 4 & up: 2 points & C=1

For all other **Contexts**: Language and actions that convey messages such as,
"I've just been lucky, I don't really get it," and "I need help; I can't do this on my own" get coded as Fragile. |
| **PAB_8**-Self-Doubting is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a neutral context and a weak belief in one's capability. This is similar to Fragile. Motivation is more seriously impaired because the context is less likely to be seen as a reliable source of help and support. | For **Con_Assess**:
For Levels 1-3: 0 point & C=1
For Levels 4 & up: <1 points & C= 1

For all other **Contexts**: Language and actions that convey messages such as,
"I know I will blow it," and "Only a miracle will help me now" get coded as Self-Doubting. |
<table>
<thead>
<tr>
<th>Definition</th>
<th>Coding Rules</th>
</tr>
</thead>
</table>
| PAB_9-Discouraged is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a negative context and a moderate or variable belief in one's capability. Any remaining hope for good outcomes is the self rather than the context. Less likely to focus on personal deficiencies than on the impossibility of making progress in the current context. | For Con_Assess:  
For Levels 1-3: 0 point & C=2  
For Levels 4 & up: <1 points & C=2  
For all other Contexts: Language and actions that convey messages such as,  
"I can't seem to make this work. I should just do something else," and "This is awful- what am I going to do now?" get coded as Discouraged. |
| PAB_10-Hopeless is for the action/interaction of the student's Personal Agency Belief about the task at hand that reflects a negative context and a weak belief in one's capability. Neither the self nor the context are seen as having any potential to improve current or anticipated negative events. Bad outcomes are inevitable; good outcomes are impossible. | For Con_Assess:  
For Levels 1-3: 0 point & C=1  
For Levels 4 & up: <1 points & C=1  
For all other Contexts: Language and actions that convey messages such as,  
"It's a lovely idea, but it will never happen," and "It's no use- I give up." get coded as Hopeless. |
EVALUATING THE PARLI FRAMEWORK

Diagram: Flowchart showing the process and development of reading comprehension and reading fluency improvement.
EVALUATING THE PARLI FRAMEWORK

Figure DC: Event Flow for Tenant (case g) shows development of reading comprehension, metacognition, and personal agency beliefs over time of study.
EVALUATING THE PARLI FRAMEWORK

Figure 2: Flowchart for standards (Case 2) showing development of scoring framework.

The diagram illustrates the development process of a scoring framework for evaluating the PARLI (Personal Assistant for Reconciliation of Language, Information, and Knowledge) framework. The flowchart depicts the steps involved in developing a scoring framework, including the identification of criteria, the application of weights, and the assessment of documents. The process is iterative, with feedback and revisions being integrated at each stage.

The diagram is labeled as follows:
- **Overall PARLI Framework**: The starting point for the development of the scoring framework.
- **Criteria Development**: The initial step involves defining the criteria that will be used for evaluation.
- **Weight Assignment**: Each criterion is assigned a weight based on its importance.
- **Document Assessment**: Each document is assessed against the criteria and weighted scores are calculated.
- **Feedback and Revision**: Based on the assessment, feedback is collected and the framework is revised.
- **Final Scoring Framework**: The final version of the scoring framework is ready for implementation.

The diagram is a visual representation of the iterative process of developing a scoring framework for evaluating the PARLI framework.
Figure DI: Event Flow for Ages (C) 10 (shows development of reading)}