

**EFFECTS OF THE *JACOB'S LADDER READING COMPREHENSION*  
*PROGRAM* ON READING COMPREHENSION AND CRITICAL THINKING  
SKILLS OF THIRD, FOURTH, AND FIFTH GRADE STUDENTS IN RURAL,  
TITLE I SCHOOLS**

---

**A Dissertation**

**Presented to**

**The Faculty of the School of Education**

**The College of William and Mary**

---

**In Partial Fulfillment  
Of the Requirements for the Degree  
Doctor of Philosophy**

---

**by  
Tamra Stambaugh  
March 2007**

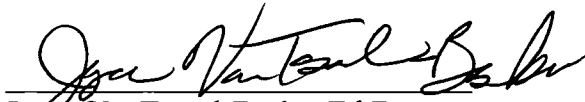
EFFECTS OF THE *JACOB'S LADDER READING COMPREHENSION PROGRAM* ON  
READING COMPREHENSION AND CRITICAL THINKING SKILLS OF THIRD,  
FOURTH, AND FIFTEH GRADE STUDENTS IN RURAL, TITLE I SCHOOLS

By

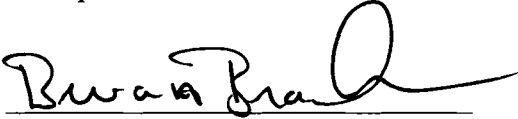
Tamra Stambaugh

---

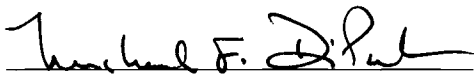
Approved March 2007 by



Joyce VanTassel-Baska, Ed.D.  
Chairperson of Doctoral Committee



Bruce A. Bracken, Ph.D.



Michael F. DiPaola, Ph.D.

## Table of Contents

### Chapter I

Introduction.....	2
Statement of Problem.....	6
Conceptual Framework.....	8
Rationale for Study.....	11
Purpose and Research Questions.....	13
Methodology Synopsis.....	13
Definition of Terms.....	14

### Chapter II

Introduction.....	19
Scaffolding.....	20
Reading Comprehension.....	25
Thinking Skills.....	35
Disadvantaged Populations.....	43
Reform Models in Education Innovation.....	49
Differentiation.....	61

### Chapter III

Methodology.....	74
Study Procedures.....	90
Data Analysis.....	97
Limitations and Delimitations.....	99

## Chapter IV

Reading Comprehension Findings.....	103
Critical Thinking Findings.....	106
Findings by Subgroups.....	109
Focus Group Findings.....	113
Summary of Findings.....	123

## Chapter V

Discussion.....	127
Implications.....	138
Future Research.....	140
Appendix A: Performance-Based Assessment.....	145
Appendix B: Treatment Fidelity Form.....	159
Appendix C: Ladder Templates.....	161
Appendix D: Ohio Content Standards.....	166
Appendix E: Focus Group Themes.....	170
References.....	175

## **Acknowledgments**

Many persons have been influential throughout the dissertation process. The two school districts that participated in the study were cooperative and professional. I could not imagine a better group of teachers and administrators to join me in this endeavor.

My committee should also be recognized: Dr. DiPaola for his moral support; Dr. Bracken for his critique and assistance in making this dissertation stronger; and of course, Dr. VanTassel-Baska, my advisor and mentor, for her wisdom, feedback, and encouragement. In addition, Dr. Annie Feng and Dr. Carol Tieso were invaluable for their statistical knowledge and guidance.

I would like to thank my Ohio friends (Kay, Mary, Kelly, Diana, Pauline, and Sharon) who made sure I had a home, food, diversions, and great conversation while I was in Ohio collecting data. My colleagues (Heather, Susannah, Bess, Catherine, Elissa, and Val) also provided support in numerous ways.

Finally, my parents, my parents-in-law, my sister, and most of all, my husband, Tim, have contributed unconditional emotional support, prayer, encouragement, and tangible contributions during this process. I am blessed to have them in my life. They are the greatest!

## List of Tables

Table 1: Comparison of Effective Reading Curriculum.....	56
Table 2: Summary of Research Findings by Literature Strand.....	67
Table 3: Jacob’s Ladder Features and Research Citations.....	71
Table 4: Overview of Research Questions and Data Collection.....	75
Table 5: Comparison of School Districts and Buildings in Sample.....	76
Table 6: Number of Participants by Factor and Condition.....	78
Table 7: Alignment of Bloom’s Taxonomy and Jacob’s Ladder.....	88
Table 8: Timeline of Tasks for Study Implementation.....	92
Table 9: Statistical Analysis by Research Question.....	98
Table 10: Means and Standard Deviations for ITBS.....	104
Table 11: Means and Standard Deviations for TCT.....	107
Table 12: Paired-Samples t Test results by Grade for PBA.....	108
Table 13: Number and Percent of Participants by Factor and Condition.....	109
Table 14: Treatment Fidelity Form and Frequency of Observed Behaviors.....	112
Table 15: Number and Percentage of Readings by Genre & Teacher.....	114

**THE EFFECTS OF THE *JACOB'S LADDER READING COMPREHENSION PROGRAM* ON READING COMPREHENSION AND CRITICAL THINKING SKILLS OF THIRD, FOURTH, AND FIFTH GRADE STUDENTS IN RURAL, TITLE I SCHOOLS**

**ABSTRACT**

The purpose of this study was to examine the effects of the Jacob's Ladder Reading Comprehension on 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade students' reading comprehension and critical thinking skills in rural, Title I schools.

The *Jacob's Ladder Reading Comprehension Program* was written as a supplemental curriculum targeted toward Title I students in the third, fourth, and fifth grade. The program focus is on scaffolding reading instruction from lower to higher level thinking skills with an emphasis on higher level thinking and textual analysis.

This quasi-experimental study measured the effects of the program on rural Title I students' critical thinking and reading comprehension (N = 495). Within the experimental group, students who were exposed to the *Jacob's Ladder Curriculum* revealed significant and very high practical gains in subject-specific critical thinking behaviors. Between-group analyses suggested that when compared to the basal reader series alone, the *Jacob's Ladder Reading Comprehension Program* produced significant and important gains in students' reading comprehension, as measured by the ITBS, and critical thinking, as measured by the Test of Critical Thinking.

TAMRA L. STAMBAUGH

DEPARTMENT OF EDUCATIONAL POLICY, PLANNING, AND LEADERSHIP

THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA

**EFFECTS OF THE *JACOB'S LADDER READING COMPREHENSION*  
*PROGRAM* ON READING COMPREHENSION AND CRITICAL THINKING  
SKILLS OF THIRD, FOURTH, AND FIFTH GRADE STUDENTS IN RURAL,  
TITLE I SCHOOLS**

## CHAPTER I

### Introduction

Reading instruction in the United States has received much attention in the last decade. Federal initiatives and grant monies are allocated toward reading instruction and literacy, especially for at-risk populations (NCES, 2004). The increased emphasis stems from national research conducted by the National Assessment of Education Progress (NAEP), a congressionally mandated group that studies achievement gains of students in the United States for each of the core content areas. Other federal programs such as Reading First from the federal legislation of the No Child Left Behind Act (2002) emphasize the need for reading scaffolding and programs for Title I students and schools.

The most recent national report card published by the NAEP reported that fewer than one-third of fourth and eighth graders read at the proficient level in 2005. The term “proficient” means that students were able to successfully meet the following objectives based on a standardized reading assessment (Perie, Grigg, & Donahue, 2005, p. 26): Retrieve relevant information to fit description; identify the main theme of the story; identify explicitly stated but embedded text detail; provide explanation of a character’s feeling; recognize fact, supported by text information; infer or identify a lesson based on text information; recognize reason that explains feelings of biographical subject; make inference to identify intent of description; and recognize meaning of specialized vocabulary from context.

Furthermore, fewer than 3% of eighth graders and 7% of fourth graders scored in the “advanced” range, meaning that few students were able to “explain the author’s

use of direct quotations; provide the overall message of the story; explain an author's statement with text information; and make inferences to identify character motivation" (Perie, Grigg, & Donahue, 2005, p. 26).

Students of poverty (e.g., defined as free or reduced lunch eligibility) increased their reading scores by 2 points from 2004 to 2005, based on the subgroup analyses. However, they are still significantly behind their higher socio-economic counterparts. White students continue to score higher than their Black and Hispanic peers, although the gap seems to be narrowing when comparing the trend from 1992 to 2005; still the poverty gap is about the same over time (Perie, Grigg, & Donahue, 2005).

When test examples from the NAEP are juxtaposed with a higher level taxonomy of thinking, such as Bloom's Taxonomy (1956) or Paul's Reasoning Model (1992), the "proficient" criteria from the NAEP assessment are not considered to be at a level commensurate with higher level thinking or reasoning processes.

Remembering and understanding are the two lower levels of the updated Bloom's taxonomy, analogous to the widely known knowledge and comprehension levels (Anderson & Krathwohl, 2001). In order to meet the criteria for the two lowest levels on Bloom's Taxonomy students must "recognize, recall, interpret, exemplify, classify, summarize, infer, compare, and explain" (p. 31). These objectives align closely with the description of "proficient" from NAEP (e.g., identify main theme, recognize fact, explain a character's feelings, or infer to identify intent). Even the "advanced" designation, which few students achieved, does not require complex higher level thinking skills. Students who scored in the "advanced" range were expected to explain and make inferences from text with moderate interpretation. These advanced

expectations mirror those related to the second level of Bloom's Taxonomy: to understand. At this level, students are expected to "construct meaning from instructional message, including oral, written, and graphic communication" (Anderson & Krathwhol, 2001, p. 31).

Higher level thinking skills or components of reasoning include tasks more difficult than word recognition, recalling, or inferring. Students who consistently work at the higher levels of thinking, according to Bloom's Taxonomy (Anderson & Krathwhol, 2001) are expected to examine and determine various points of view, organize or structure evidence to make an argument, implement certain procedures to execute a plan, critique various methods or author's purposes, make judgments about a given passage; or generate, plan, or produce a new plan, idea, or perspective. Likewise, Paul's Reasoning Model (1992) aligns with the higher levels of Bloom's Taxonomy. Paul (1992) suggests that for persons to reason through a situation they must determine points of view, make inferences based on given data, make judgments about a given situation to determine implications and consequences, and determine conceptual relevance or main ideas.

The lack of higher level thinking questions creates the greatest problems for gifted students who master content at a faster pace, deal with complex issues more readily, and conceptually understand advanced content (Clark, 2002). If teachers are teaching to assessments, as recommended by many school leaders, students may not be exposed to the higher level thinking skills required for advanced learning and literary analysis at the upper levels described by Bloom's Taxonomy or Paul's Reasoning Model (Paul, 1992). Without exposure and practice with the levels of

thinking that gifted students are capable of accomplishing, underachievement is more likely to occur. Similarly, gifted students may not make the learning gains expected within a school year (Sanders & Horn, 1998). This exposure to higher level thinking is even more critical for gifted students of poverty who may not have access to an enriched environment or early literacy skills (Hart & Risley, 1995).

Moreover, the emphasis on reading instruction continues to be a focus for those students who are less proficient in reading, especially in the younger grades (Tivnan & Hemphill, 2005). The National Reading Panel, NRP (NICHD, 2000) conducted a meta-analysis of reading studies to determine the strategies most effective for reading instruction. Many of the studies included in the analysis were targeted toward struggling students, learning disabled, and those students just learning to read. Since many gifted students come to school with basic reading skills (Reis, Westberg, Kulikowich, Caillard, Hebert, & Plucker, 1993) and may have been reading since three to four years of age, recommendations of the panel for phonemic awareness and basic reading skills are not as applicable. Regardless, many schools have adopted carte blanche curriculum and strategies recommended by the NRP for use with all learners (VanTassel-Baska & Stambaugh, 2006) even though sub-analyses of the NRP suggest that certain strategies recommended are only appropriate for those students who are learning to read or struggle with reading. Gifted students and others who are already reading and comprehending at levels beyond their grade level peers are at a disadvantage if they are expected to endure phonics and letter instruction through most of the elementary years as exemplified by NRP's recommendations. Access to more

challenging curriculum is denied unless teachers differentiate instruction and provide alternate means for those students who are capable of advanced skills.

Finally, regardless of the societal assumptions, gifted students are not likely to automatically possess the information needed to think at higher levels without exposure to advanced level curriculum (Taylor, Pearson, Peterson, & Rodriguez, 2003; VanTassel-Baska & Stambaugh, 2006), scaffolding of thinking (NRP, 2000; Zemelman, Daniels, & Hyde, 1998), graphic organizers (NICHD, 2000), discussion about conceptual literary topics (Paul, 1992) and meaningful teacher feedback (Taylor, Pearson, Peterson, & Rodriguez, 2003). Additional research-based curriculum is needed to focus on skills for more capable learners, since much of the research-based reading curriculum focuses on students who struggle with reading or who are below “basic” (Tivnan & Hemphill, 2005), especially students in Title I schools.

#### Statement of the Problem

For students to move toward higher levels of thinking and achievement, exposure to more rigorous curriculum is needed (Taylor, Pearson, Peterson, & Rodriguez, 2003; VanTassel-Baska, 2003a). Gifted students, in particular, often are not exposed to the advanced curriculum and higher level thinking skills commensurate with their abilities (Reis et al., 1993; VanTassel-Baska, 2003a). National assessment data suggest that 3 to 8% of students score at advanced levels on reading tests (Perie, Grigg, & Donahue, 2005) although at least 17% of the national school population is identified as gifted (NAGC, 2004). A double dilemma exists for students who are gifted and of poverty. Not only are gifted students typically not considered when school districts make curriculum decisions, but many students of poverty begin school

with a weaker language base and vocabulary (Hart & Risley, 1995), are less likely to have access to advanced literacy curriculum or scaffolding of instruction to move toward advanced skills, must endure unchallenging, rote activities (Taylor et al., 2003), and are assigned to teachers who are less likely to possess the training and skills required to teach literacy at advanced levels (Haycock, 2000).

Funding poses another issue. Poorer districts may be incapable of sustaining programs for advanced students or purchasing quality curriculum. Furthermore, students of poverty are less likely to possess or acquire the necessary tools for advanced thinking skills without support or exposure from other structures, especially in a verbal domain, such as reading (Payne, 1995). Due to this lack of exposure and recognition, scaffolding from lower level to higher level thinking skills in reading may be necessary to promote higher level achievement for students, and especially gifted students who come from poverty.

The *Jacob's Ladder Reading Comprehension Program* (VanTassel-Baska, Stambaugh, & French, 2004) was developed for third, fourth, and fifth grade "promising learners" in Title I schools to provide teachers and students with a framework for improving thinking skills, and providing scaffolding for students who need support for higher level thinking. However, limited data exist regarding the effectiveness of the *Jacob's Ladder* program for gifted learners in Title I schools (French, 2005).

## Conceptual Framework

### *Bloom's Taxonomy of Educational Objectives*

The basic goal of progressing students from lower order to higher order thinking skills is the foundation of the *Jacob's Ladder Reading Comprehension Program*. Bloom's Taxonomy of Educational Objectives (Anderson & Krathwohl, 2001; Bloom, 1956) provides a similar framework of objectives from lower order to higher order skills based on the level of complexity in thinking. This hierarchical framework is most evident in the early work of Bloom (1956). In recent years, Anderson and Krathwohl (2001) reorganized the hierarchical framework, making it two-dimensional. The first dimension is almost identical to that of Bloom's Taxonomy but is now labeled "cognitive processes." The added dimension of "knowledge" is not included in the original Bloom's Taxonomy. Each will be discussed separately.

The "cognitive processes dimension" outlines the specific thinking processes that students progress through as they obtain greater understanding or gain higher levels of complexity. Cognitive processes include the following components, in hierarchical order from lowest to highest complexity:

1. Remember: Remembering requires students to recall knowledge from long-term memory and to recognize, identify, retrieve, or recall specific information.
2. Understand: Understanding involves "making meaning from instructional messages" (p. 67). Specific requirements within this category include restating data in one's own words through clarifying, paraphrasing, representing, or

translating; providing examples or illustrations of an idea or concept; categorizing or classifying information given; summarizing information to more abstract themes; inferring or coming up with logical conclusions based on presented information; comparing and contrasting or determining how well two ideas correspond; and explaining or being able to outline a model to explain a concept.

3. **Apply:** Applying suggests that students use a specific procedure to carry out a task. For example, students may be required to apply what they have read to another situation or use information garnished with a task that is unfamiliar.
4. **Analyze:** In order to analyze, students must be able to “break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose”. This includes differentiating between important and unimportant information or parts of material presented, organizing information in such a way to determine how various parts fit within a structure, attributing or deconstructing what is written in order to determine bias, assumptions, or perspectives.
5. **Evaluate:** Students evaluate when they “make judgments based on criteria and standards” (Anderson & Krathwohl, 2001, p. 68). This means they are able to critique two alternatives or determine internal or external consistencies within a procedure or author’s conclusions as aligned with the other knowledge.
6. **Create:** Creating involves the placement of “elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure” (Anderson & Krathwohl, 2001, p. 68). Students who create would

be asked to hypothesize, generate alternative options to account for an observed situation, or plan or design a new product or way of accomplishing a task.

The second, and new, dimension of the taxonomy outlines the types of knowledge a student might possess for each cognitive process dimension. The knowledge dimension includes factual knowledge, conceptual knowledge, procedural knowledge, and meta-cognitive knowledge. These are not hierarchical among each category but could be, depending on how the knowledge is applied through the cognitive processes dimension.

1. Factual knowledge outlines those specific details, elements, and facts needed within a specific discipline.
2. Conceptual knowledge captures the relationships among different elements, principles, generalizations, or theories.
3. Procedural knowledge includes the methodology of how to do something including the criteria involved in determining appropriate procedures.
4. Metacognitive knowledge is the knowledge an individual has of personal cognition, self-knowledge of strengths and weaknesses, and personal and general strategies used in thinking.

Each of these processes is incorporated to some extent in the *Jacob's Ladder Reading Comprehension Program*, either implicitly or explicitly. Alignment of the conceptual framework to *Jacob's Ladder* is outlined in more detail in Chapter 3.

## Rationale for the Study

The *Jacob's Ladder Reading Comprehension Program* (VanTassel-Baska, Stambaugh, & French, 2004) was developed specifically for promising learners in Title I schools in order to provide teachers and students with a framework for progressing students from lower order to higher order thinking skills. However, few data exist regarding the effectiveness of the program for students in Title I schools, especially in heterogeneous classrooms. One pilot study was conducted to evaluate the efficacy of the *Jacob's Ladder Reading Comprehension Program* among high ability students, finding no significant differences in critical thinking or reading comprehension achievement mean gain scores between the control and experimental groups during an eight week intervention (French, 2005).

French (2005) posits several reasons for this and suggests a need for further study. First, she explained the intervention was only eight weeks in duration and qualitative data suggest that teachers did not have adequate time to devote to the curriculum. If the teachers used the *Jacob's Ladder* curriculum as a supplemental curriculum, which was directed, only half of the program could be implemented and therefore, the full impact of the program would not be evident. The timing of the intervention was also called into question. The intervention took place in the Spring, just before school ended and after state assessments. Consequently, student motivation and special end-of-the-year programs may have interfered with the treatments. Finally, French (2005) questioned the validity of the assessments, especially the critical thinking instrument, based on the duration of the intervention as being too short to see significant effects in critical thinking. The reading

comprehension test was also posed as being problematic since the on grade level version was administered to students who are advanced. There may have been a ceiling effect or regression to the mean.

Regardless of its limitations, the pilot study showed promise, even in the short duration of use (French, 2005). Qualitative data suggest that teachers observed growth in student responses to literature and students reported that they enjoyed the program components. Informal feedback from trainings led by the researcher during the past two years support this finding. Teachers and students believe the program has merit. Many educators provided scenarios of student growth gains and increased personal confidence to support learner needs in reading.

Although the pilot results were not significant and are attributed to sampling error, the post-test means of the experimental group were slightly higher than the control, indicating promise. French (2005) recommended additional study, including a longer duration of the intervention, use of the intervention during the fall semester of the school year - prior to state testing, and a better match of assessments to the curriculum for measuring higher level thinking and reading comprehension.

A broader scope of the study of *Jacob's Ladder* is also warranted. As the curriculum gains popularity throughout the United States, based on copies distributed through the Center for Gifted Education and an increased demand for training on the curriculum, many teachers are reporting that they are using the *Jacob's Ladder Reading Comprehension Program* with their entire classroom, not just with gifted learners. To date no formal data had been collected on the effects of *Jacob's Ladder* for non-gifted learners/promising learners in heterogeneous classrooms.

Additional study of this program is validated. However, in order to build upon the pilot study recommendations and other anecdotal evidence the assessments, targeted population, intervention duration, and sample size will need to be modified to decrease error and increase the validity of results, as suggested by French (2005).

#### Purpose and Research Questions

The purpose of this study is to conduct a more comprehensive follow-up of the study of the *Jacob's Ladder Reading Comprehension Program* to determine academic effectiveness of the program for gifted and non-gifted learners in Title I schools. The program was designed to target students in grades three, four, and five. The following questions will be explored with respect to each of those grade levels:

1. What differential effect does the *Jacob's Ladder Reading Comprehension Program* have on students' reading comprehension?
2. What differential effect does the *Jacob's Ladder Reading Comprehension Program* have on students' critical thinking and higher level thinking skills in literature?
3. Are there differential effects of the *Jacob's Ladder Reading Comprehension Program* on critical thinking and reading comprehension by gender, poverty level, gifted status, and grade level?
4. What are the perceptions of teachers and students regarding the use of the *Jacob's Ladder Reading Comprehension Program* after implementation?

#### Methodology Synopsis

This intervention study utilized a quasi-experimental, pretest/posttest design (Campbell & Stanley, 1969). Intact classrooms of participants were randomly chosen

from a purposeful selection of two Title I school districts in a rural geographic area that housed high concentrations of students who qualify for free and reduced lunch (i.e., >40%). The effects of the *Jacob's Ladder* intervention were measured to determine the to what extent the curriculum impacts students' reading comprehension and critical thinking. Between and within subjects analyses were conducted for the total population. The effects of the curriculum were also assessed on subgroups of participants including those of different gender, socio-economic status, gifted classification, and grade level.

#### Definition of Terms

The subsequent key words or terms will be used within the study. The words and definitions are included to serve as a delimitation of the study and to ensure common understanding of the semantics used.

#### *Gifted and Nongifted*

Gifted is defined by the United States Department of Education as students who “perform or show potential for performing at remarkably high levels of accomplishment when compared with others of the age, experience, or environment” (USDOE, 1993, p. 26). The state of Ohio has adopted the same definition (ORC 3301-51-15). That definition is operationalized in school districts based on the Ohio Revised Code (ORC) 3301-51-15 which identifies gifted students in one of four areas: Superior Cognitive, Specific Academic, Creative, and Visual or Performing Arts.

The state of Ohio also outlines the specific criterion required for students to be identified in one or more of the aforementioned categories. Superior cognitive students are defined as those who score two or more standard deviations above the

mean, minus the standard error of measure on a state approved intelligence test. Most students identified as superior cognitive would possess an IQ at or above approximately 124, depending on the assessment used.

Specific academic students are those who score at or above the 95<sup>th</sup> percentile on a state-approved standardized achievement assessment in any one or more of the following areas: science, social studies, reading, or mathematics. Most of the state approved assessments include widely recognized assessments such as the Iowa Test of Basic Skills or the Stanford Achievement Test.

For the purposes of this study the term “gifted” includes those students who are identified by the state of Ohio as Superior Cognitive (IQ 2 standard deviations from the mean minus the standard error of measure) and/or Specific Academic in reading only (at or above the 95<sup>th</sup> percentile on a standardized achievement test in reading). Other students participating in the study may be identified by the state of Ohio as gifted in math, science, social studies, visual/ performing arts or creativity. However, they will not be counted as gifted for this study since those areas are not the focus of intent. “Nongifted” would then refer to those students who are not identified as gifted in reading or superior cognitive as measured by the State of Ohio and reported by school districts.

### *Scaffolding*

Scaffolding is an instructional technique whereby the teacher models the desired learning strategy or task, then gradually shifts responsibility to the students (Clay & Cazden, 1992). It is defined as “providing support to help learners bridge the gap between what they know and can do and the intended goal” (Graves, Graves, &

Braaten, 1996, p. 53). Within the realms of this study, scaffolding is defined as: a) the teacher processes needed to move students from lower level to higher level thinking including the modeling of desired strategies and behaviors that guide students to discuss ideas and defend answers through justification from the reading selection and b) the construction of the *Jacob's Ladder* curriculum. *Jacob's Ladder* provides “scaffolding” through the use of questions that begin at the lower level, basic knowledge and application questions, and then increase in difficulty to more difficult questions that require analysis, synthesis, and evaluation. These questions are intended to “bridge” the gap from lower level to higher level thinking in reading.

#### *Higher Level Thinking Skills and Lower Level Thinking Skills*

Higher level thinking skills are defined by Bloom's Taxonomy as part of the conceptual framework for this study and Paul's Reasoning Model (1992). The upper rungs of the *Jacob's Ladder Reading Comprehension Program* are aligned with the upper levels of Bloom's Taxonomy and Paul's Reasoning Model. Higher-level skills include implications/ consequences, concept or theme, inference, generalization, and creative synthesis.

Lower level thinking skills are those categories in Bloom's Taxonomy that are more basic in nature, namely the first three levels of Bloom's Taxonomy: remembering, understanding, and applying. Lower level thinking skills included in the *Jacob's Ladder Reading Comprehension Program* are paraphrasing, sequencing, recalling details, and the identification of literary elements. Therefore, higher level thinking skills are the three upper levels of Bloom's Taxonomy and lower level thinking skills are the lower level skills, as defined by Bloom.

### *Title I*

The term Title I is used to represent the poorest schools in the nation or a state that are eligible for special funding due to the poverty levels of students within a district or school. Title I schools are given special designation by the federal government based on free and reduced lunch percentages of students within the school district or sometimes specific school buildings. Funding formulas established by the United States Department of Education are used to determine which districts qualify for Title I services. Generally schoolwide services and larger grants are given to schools with at least 40% of their students labeled as disadvantaged as measured by free or reduced lunch qualifications (United States Department of Education, 2007). Both elementary buildings in this study have percentages of students on free and reduced lunch that exceed 40%.

### *Students of Poverty/Low SES/Free-Reduced Lunch*

These three terms are used synonymously and are defined as those students reported by the school system who have applied for and qualify for free or reduced lunch within a school system based on family income levels. The United States Department of Education (2006) defines the qualifications for free and reduced as follows: free lunch – total household income at or below 130% of the poverty level; reduced lunch – total household income between 130 to 185 percent of the poverty level. The actual income levels to qualify vary based on the poverty rate each year but typically a family of 4 who makes less than \$37,000 per year is eligible for reduced lunch fees (United States Department of Education, 2006).

### *Reading Comprehension*

For purposes of this study, reading comprehension is defined by the Iowa Test of Basic Skills (ITBS) subtest measure. The ITBS publishing company defines reading comprehension as a measure of three types of understanding. Factual questions tap students' literal understanding of what is stated in the text. Inferential/interpretive questions require students to read between the lines to demonstrate their understanding of what is implied. Analysis and generalization questions require students to "step back from" the text to generalize about a passage's main points or ideas or to analyze aspects of the author's viewpoint or use of language (Riverside Publishing, 2007, para 5).

### Conclusion

This study focused on the effects of the *Jacob's Ladder Reading Comprehension Program* on students' critical thinking and reading comprehension using the higher level thinking skills. The specific audience of focus is third, fourth, and fifth grade students in Title I rural schools. There is qualitative evidence to suggest that the *Jacob's Ladder* curriculum may produce value-added gains in students; however, to date the impact of the curriculum on students in heterogeneous classrooms has not been measured quantitatively.

Chapter II provides an overview of the literature for reading as it relates to thinking, reading comprehension, scaffolding, students of poverty, and conditions necessary for innovative practices to be successful.

## Chapter II

### Introduction

The goals of the *Jacob's Ladder Reading Comprehension Program* are four-fold: 1) to scaffold student thinking through the movement of lower level to higher level thinking skills, including teaching strategies that facilitate the processes, 2) to assist students with reading comprehension skills, 3) to use thinking skills as a basis for textual understanding, and 4) to differentiate reading instruction. Therefore, the following areas of study will serve as the basis for the review of literature:

scaffolding, reading comprehension, thinking skills, and differentiation. Since the *Jacob's Ladder Reading Comprehension Program* is an innovative program that was specifically designed for students in Title I schools, an examination of the literature on poverty, including curriculum effectiveness in reading will also be included as will a review of the literature on innovative practices.

Each of the overarching topics for the literature review corresponds to specific curriculum ladders and implementation procedures for the *Jacob's Ladder Reading Comprehension Program*. The goals, by ladder, include reasoning and thinking skills such as judging relationships within the text and determining implications and consequences (Ladder A), conceptual thinking through the examination of details, classifications, and generalizations (Ladder B), comprehension including literary analysis skills by examination of characters, plot, and setting (Ladder C) and creative synthesis after summarizing and sequencing (Ladder D). Teacher stance and collaborative procedures for students as part of scaffolding are embedded in the

teacher guide to implementation and included as part of professional development training.

## Scaffolding

### *Theoretical Basis*

The notion of employing an instructional process that helps students progress from lower level to higher level thinking skills is not a new idea. Bruner (1961) first introduced the word “scaffolding” in an educational context after observing parents helping their children learn to speak. He found that parents unconsciously taught their children how to speak and read by providing predictable experiences such as bedtime reading, playfulness in words and stories, focus on meaning as parents describe different vocabulary of the family and world as it arises, role reversal, when children “read” to their parents favorite stories, modeling, as parents demonstrate reading informally, and nomenclature which is described as commonly shared words among the child and parent. Bruner’s notion has been extended to the educational field over time and is recommended as an important strategy in teaching instruction in general (Daniels, 1994; Zemelman, Daniels, & Hyde, 1998).

Theoretical models such as Vygotsky’s zone of proximal development and Bloom’s Taxonomy of Educational Objectives also apply to scaffolding. Both emphasize frameworks that aid students in building on lower levels of thought and complexity as they move to advanced levels in a given area of development (Bloom, 1956; Vygotsky, 1978, 1986).

The premise of Vygotsky’s theory of the zone of proximal development (1978, 1986) is two-dimensional. Children innately have two different levels in which they

may work on a given task. One is an independent dimension, the other a collaborative dimension with assistance or mediation. The optimal level between the two is dubbed “the zone of proximal development.” The zone of proximal development implies that tasks given to students should be harder than what they can do independently, to make them appealing and not cause boredom, but not so difficult that students are frustrated (Benbow, 1998; Tomlinson, 1999). Vygotsky believed that the student levels of understanding and development could be raised from their current independent level of working to a more sophisticated and higher level of development, given the proper mediation or assistance (i.e., scaffolding). Therefore, the importance of not only scaffolding tasks but providing mediation for those tasks is essential to increasing student development.

Bloom’s Taxonomy (1956), unlike Vygotsky’s theory, focuses on scaffolding the task or educational objectives, and does not explicitly include teacher or mediation factors. Bloom’s Taxonomy was originally designed as a framework for evaluating progress of thinking and educational objectives written to solicit appropriate student thinking skills in a hierarchical manner from lower level knowledge and recall-based skills, to higher level skills of evaluation and synthesizing information. By combining the teacher mediation and zone of proximal development notions of Vygotsky with the hierarchical approach of objectives of Bloom’s Taxonomy, teachers and curriculum writers have a blueprint for teaching and providing meaningful experiences for students.

### *Instructional Strategies for Scaffolding*

Based on the theoretical frameworks, scaffolding is known as an instructional technique whereby the teacher models the desired learning strategy or task, then over time allows more independence with the task (Clay & Cazden, 1992). Scaffolding not only means the teacher modeled behavior is shifted to the learner, but also the structure of the curriculum provides a framework for students to move from lower level thinking to higher level thinking strategies (Fournier & Graves, 2002). Over two decades of national reports on best practices across content areas has been compiled, many of which include scaffolding as an important teacher strategy regardless of the content area. These studies suggest that scaffolding is most effective when teachers help students work collaboratively through problems, provide feedback as necessary, and ask probing questions that help students think about alternative answers, other ideas, or their own metacognitive thinking processes (Perkins, & Salomon, 1989; Vygotsky, 1986; Zelman, Daniels, & Hyde, 1998).

Much of the empirical research base on scaffolding comes from the observation of and interaction with effective teachers of reading. A qualitative study analyzed teacher talk in classrooms of effective teachers (based on previous teacher success with student reading test scores). Teaching sessions were transcribed and over 89 observances and 200 hours of taped discussion were analyzed. Effective teachers were found to scaffold instruction over time by providing less teacher talk and more student talk, with greater complexity of tasks, as students became more proficient (Hansen, 2004). Within this framework, teacher questions were also scaffolded, or moved from lower level to higher level questions, as follows: 26% were

comprehension-related, 11% reflected on the story, 29% percent included discussion about the text, and 34% focused on ways to respond to the story through discussion of literary elements and evaluation.

Other studies on scaffolding also examined teacher behaviors. Effective teachers were observed in a variety of settings while teaching reading and using modeling and feedback processes associated with scaffolding. Effective teachers of reading:

1. provided more coaching throughout the process of reading (Guthrie, 1996; Pressley, Wharton-McDonald, Allington, Block, Morrow, Tracey, Baker, Brooks, Cronin, Nelson, & Woo, 2001; Taylor, Pearson, Clark, & Walpole, 2000),
2. stressed higher order thinking skills in addition to lower level skills, building upon meaning (Knapp, Adelman, Marder, McCollum, Needels, Padilla, Shields, Turnbull, & Zucker, 1995; Pressley et al., 2001; Taylor, et al., 2003),
3. modeled personal thinking processes aloud throughout reading instruction (NRP, 2000; Zelman, Daniels, & Hyde, 1998),
4. included opportunities for independent reading followed by collaborative small and whole group discussion (Campbell, Voelkl, & Donahue, 1997; Chin, Anderson, & Waggoner, 2001; Guthrie, 1996),
5. included opportunities, with feedback, for students to self-regulate and assess their own learning as they increase their knowledge and skill

base (Chin et al., 2001; Pressley, et al., 2001; Roehler & Duffy, 1984; Zemelman, Daniels, & Hyde, 1998), and

6. linked personal experiences to reading, with an emphasis on process (Chin et al., 2001; Roehler & Duffy, 1984; Taylor, et al., 2000).

Many of these instructional strategies are also intertwined with the literature and meta-analyses findings on thinking skills and reading comprehension.

### *Scaffolding Curriculum Framework*

Fournier and Graves (2002) describe a specific scaffolding program that was found to be effective: the SRE (Scaffolded Reading Experience). The program outline for implementation includes broad-based strategies for pre-reading, during reading, and post reading activities as part of the overall reading instruction (Graves & Avery, 1997; Stephens & Brown, 2005). Using the SRE framework, teachers intervene and interact with their students during the reading process to provide cues, questions at increasingly difficult levels of thinking, coaching to solicit appropriate responses, collaborative opportunities, and information provision or background information as necessary. Prior to the pre-implementation, implementation, and post-implementation phases is a planning phase. Planning includes the selection of readings, consideration of student needs, and the overall objectives for teaching.

Though descriptive analyses for SRE showed gains for students, few accommodations are made to scaffold instruction for advanced readers. The authors concluded “with more proficient students, less difficult selections, and less challenging purposes, less scaffolding is needed” (Fournier & Graves, 2002, p. 32).

Teachers of advanced students must provide selections that are engaging, challenging,

and difficult enough for advanced readers. Otherwise, as implied in Vygotsky's theory, there is nothing to scaffold and advanced students are likely to either complete tasks independently without having the opportunity to move to higher levels of thinking or development unless the teacher intervenes through advanced questioning or more difficult reading selections (Tomlinson, 1999).

### *Summary of Scaffolding Literature*

Scaffolding involves the process of teacher modeling and feedback provision to help students progress from lower level to higher-level skills. Scaffolding is loosely defined as a set of processes teachers use to enhance reading comprehension such as coaching, using higher order thinking skill questions, modeling thinking processes aloud, allowing opportunities for discussion with feedback, and linking personal experiences to reading selections. Teachers who use these behaviors systematically are found to produce higher test scores in reading than those who do not.

One specific curriculum, the Scaffolding Reading Experience (SRE), incorporates the teacher processes of scaffolding into a systematic approach to teaching reading in different phases. Pre-post gains in students' ability to comprehend reading instruction are posited based on non-standardized measures.

## Reading Comprehension

### *Introduction*

Empirical research in reading began over five decades ago (NRP, 2000) and to date a plethora of data exist on reading and effective reading strategies. Much of these data were spawned from previous national reports that called for more empirical evidence in reading as well as more focused attention on reading. Federal legislation

such as the No Child Left Behind Act (2001), and subsequent state legislation, has spawned a flurry of policies and initiatives in reading and the promotion of reading comprehension. Federal money has also been allocated to target at-risk groups such as minority populations, students with English as a second language (ESL), and students of poverty, all of whom score significantly lower on achievement assessments in reading (NCES, 2004).

Even with the emphasis on reading and a stronger knowledge base of the effectiveness of specific reading comprehension strategies, the outcome is less than adequate. Statistics on reading in the United States from the NAEP, the National Assessment of Education Progress, over the past decade continues to be disappointing as only a small percentage of students are proficient in reading and even fewer score at advanced levels on national assessments (NCES, 2004). As children progress through grade levels, even the most fluent readers are unable to comprehend in ways that illustrate understanding of text (Caccamise & Snyder, 2005). Despite the plethora of research available, students are not receiving adequate instruction in reading comprehension (Pressley et al., 2001; Taylor et al., 2002).

### *Reading Comprehension Strategies*

Reading comprehension strategies and instruction are central to an effective reading program (Duffy, Roehler, & Mason, 1984; Pressley, 1998; Roehler & Duffy, 1984). Reading comprehension definitions have changed over the past thirty years. Although reading comprehension was once defined as only being able to decode text, the current definitions are far more complex and involve knowledge, experience, thinking, and teaching. “Comprehension inherently involves inferential and evaluative

thinking, not just literal reproduction of the author's words. Most importantly, it can be taught directly" (Fielding & Pearson, 1994, p. 63).

The National Reading Panel (2001) screened almost 100,000 studies for use in a meta-analysis report on reading. Only those juried, empirical studies with an experimental or quasi-experimental design were included in the meta-analysis. Five areas of reading focus emerged: Phonics awareness, instruction, and fluency; comprehension which includes vocabulary and comprehension; teacher education and reading instruction; teacher preparation and comprehension instruction; and computer technology and reading instruction.

The review of the empirical studies on reading comprehension posit statistically significant evidence for the following eight instructional strategies that promote comprehension: cooperative learning, the use of graphic and semantic organizers, understanding of story basics (who, what, when, where, and why), question answering, question generation, summarization, comprehension monitoring, and the use of multiple strategies together (NRP, 2000).

Cooperative learning is one of the most over-generalized findings of the NRP, with the possible exception of phonemic awareness. Much of the cooperative learning literature included in the panel meta-analysis did not include high ability students. The literature on the effectiveness of cooperative grouping on reading comprehension by ability groups is mixed (Rogers, 1998). Johnson and Johnson (1994) suggest that cooperative learning must be deliberately taught to students and include the following conditions: positive interdependence, positive face-to-face interaction, clear responsibility and accountability for each person in a group, the use of appropriate

social skills, and ongoing process of group functioning. They also emphasize heterogeneity across groups. A meta-analysis on cooperative learning found an effect size of .66 when comparing student achievement gains in cooperative versus competitive groups (Johnson & Johnson, 1989). However, few of the studies were disaggregated based on student ability effects (Robinson, 1990; Slavin, 1990), especially for the gifted. Of the studies that did extrapolate gifted data, none suggest academic benefits for gifted students when the groups are heterogeneous (Robinson, 1990). Ability grouping with differentiation is still the academically effective option for gifted students (Kulik & Kulik, 1997; Rogers, 1998).

Graphic organizers are visual representations of text that assist students with reading comprehension and understanding of complex issues within a particular reading selection (Boyle & Weishaar, 1997). Graphic organizers are reported to contribute to the comprehension skills of students as measured by diagnostic reading assessments (Boyle & Weishaar, 1997) and performance-based assessments (Willerman & MacHarg, 1991). Hudson, Lignugaris-Kraft and Miller (1993) conducted a meta-analysis on graphic organizer use and the impact of graphic organizers at different points of instruction. They found that gains were similar in reading comprehension regardless of whether the advanced organizers were introduced prior to introduction to a reading passage or after a reading lesson. The greatest impact, however, related to the use of graphic organizers, occurs when explicit instruction and training on the use of graphic organizers is present through teacher modeling and explicit instruction.

Questioning is also important in reading comprehension gains. Open-ended questions, questions associated with the highest levels of Bloom's Taxonomy, produce higher academic gains than closed-ended questions (aligned with the lower levels of Bloom's Taxonomy) (Allington, 2001). Rosenshine, Meister, and Chapman (1996) studied a specific teaching method, Reciprocal Teaching, and found that questioning in isolation of other factors had significant effects on student achievement. Teachers who use higher-level questions in reading and who encourage more dialogue produce students with higher test scores in reading comprehension (Guthrie, Schafer, & Huang, 2001; Taylor, et al., 2000; Pressley, 1998).

Comprehension monitoring, or teacher stance, involved coaching, modeling, and feedback which enhances reading comprehension and writing growth (Pressley, et al., 2001; Taylor et al., 2002). "Think-alouds" is a specific metacognitive strategy for comprehension monitoring. Teachers reflect aloud about their own thinking regarding a reading passage and model a specific strategy. Comments and discussion are solicited from students, based on the strategy presented (Block & Israel, 2004). Studies suggest that students more readily replicate the strategy after teacher modeling and transfer the strategy in their personal reading (Oster, 2001). Significant increases in reading comprehension test scores, as measured by standardized achievement data, have been documented using comprehension modeling and feedback as part of a "think aloud" (Block, 2004).

### *Reading Practices and the Gifted*

Although these practices align well with the issues of teaching reading to high ability learners (with the exception of mixed ability cooperative learning),

interventions for the verbally gifted must extend basic practice to add depth and complexity based on the student's ability and verbal characteristics as an extension of the grade level standard. In addition to the types of modifications listed, reading programs for the gifted should include research skills, reading a variety of genres, independent research projects, opportunities to pursue areas of interest in depth over a long period of time, and guidance in critical reading and literary analysis (Halsted, 2002). Additionally, the level of reading and reading content should match the child's proficiency. The incorporation of writing, vocabulary and grammar, oral discourse and communication cannot be overlooked and must be integrated with reading as part of a comprehensive program.

A summary of the research findings for advanced readers would suggest that curriculum and instruction include appropriate selection of reading materials, guided critical discussions and advanced organizers for processing, a connection to broad-based themes and issues, the incorporation of writing models, language and oral communication, language studies, and independent research and interdisciplinary connections (VanTassel-Baska & Stambaugh, 2006).

### *Selection of materials*

The gifted child's primary contact with the world of ideas is through literature. Books stimulate thought and provide the knowledge base required for creative thinking and problem solving. Intellectual growth in gifted children depends on their access to and regular involvement in the reading process. From the time of their earliest ability to read, they need access to a rich variety of fiction and nonfiction and opportunities to respond actively and creatively to what they are reading. Students

should have abundant opportunities to discuss, analyze, and share the enjoyment of what they read with parents, teachers, and each other. Moreover, they need to be guided by adults who model the processes of analyzing and discussing reading.

Several authors provide excellent guidance for teachers about good literature for the gifted and how to teach it to optimize learning and love of literature. In Books for the Gifted Child, Baskin and Harris (1980) suggested the following criteria for locating appropriate books for the gifted:

1. The language used in books for the gifted should be rich, varied, precise, complex, and exciting, for language is the instrument for the reception and expression of thought.
2. Books should be chosen with an eye to their open-endedness, their capacity to inspire contemplative behavior, such as through techniques of judging time sequences, shifting narrators, and unusual speech patterns of characters,
3. Books for the gifted should be complex enough to allow interpretative and evaluative behaviors to be elicited from readers.
4. Books for the gifted should help them build problem-solving skills and develop methods of productive thinking.
5. Books should provide characters as role models for emulation.
6. Books should be broad-based in form, from picture books to folktale and myths to nonfiction to biography to poetry to fiction (p. 46).

Gifted students must read books that incorporate playful thinking, utilize visually inventive pictures, cite unusual connections, abstractions and analogies, and encourage connections and patterns within and among books (Halsted, 2002).

### *Guided critical discussions about literature and inquiry processes*

Guided critical discussions are the basis for literary analysis. Students must be engaged in the process of thinking and discussing text to determine meaning, inferences, and multiple interpretations (Little, 2002). The examination of key words, the structure of the text such as repeated words or sentence structure, main ideas based on readings, and implications of a character actions or setting are a basis for conversation about literature among gifted students. Beck and McKeown (2001) emphasized the importance of literary analysis and discussion. They cited specific teacher strategies, including reframing student responses, synthesizing known material, connecting reading to other situations, and modeling thought processes when conducting literary discussion. Students need practice analyzing text and justifying their basis for assumptions and responses by citing sentences, paragraphs or anomalies within the passage. Advanced graphic organizers within this framework are also necessary and should move students beyond factual interpretation to more in-depth learning and grappling with new content (VanTassel-Baska, Zuo, Avery, & Little, 2002).

Another method of inquiry is critical thinking and textual analysis questioning techniques that help students interpret and reason through situations in a novel, analyze characters, identify themes, and determine the author's purpose for writing. Use of discussion groups, workshop techniques for the writing process, panels, and debates are also appropriate strategies that can enhance collaborative learning (VanTassel-Baska & Stambaugh, 2006).

### *Writing about Literature*

The integration of reading and writing tasks has produced learning benefits for students (Henry & Roseberry, 1996; Newell, 1996). Specifically, the combination of incorporating inquiry through advanced questioning, analyzing and responding in writing to literature, pre-writing, and communicating specific criteria as expectations for learners have been found to be effective strategies that produce higher achievement gains in learners (Sadoski, Willson, & Norton, 1997). Feedback based on writing also produces higher achievement gains if specific, instructional objectives are manifest (Appleman & Applebee, 2000). When graphic organizers, exemplary writing models, explicit teaching of expectations through a rubric, and peer and teacher feedback are incorporated for persuasive writing, gifted learners showed significant improvement (VanTassel-Baska, Johnson, & Boyce, 1996; VanTassel-Baska, Zuo, Avery, & Little, 2002).

Writing opportunities for the gifted then should begin early and provide an abundance of opportunities to write. Writing is a thinking process. Thus, through writing experiences the gifted child can develop excellence in the capacity to think as well as to write. Very young children who may lack the motor coordination to write may nevertheless be engaged in writing-related activities through special teaching techniques such as tape recording, illustration, dictation, puppet shows, and plays (Halsted, 2002).

### *Teaching Strategies for Reading Comprehension*

Another set of knowledge gleaned from the literature is that of specific instructional strategies teachers may utilize in the classroom. In a comprehensive

review of empirical literature conducted based only on samples of students who are already reading, typically those above third grade. Biancarosa and Snow (2004) found the following instructional strategies to be effective for improving reading comprehension and overall literacy: direct, explicit, instruction; effective instructional principles embedded in content; opportunities for self-directed learning; cooperative learning, strategic tutoring; extended time for reading; and the incorporation of diverse texts in the learning experience. Two earlier analyses add the following to the instructional strategy list: higher level questioning strategies (Duke & Pearson, 2002; Pressley, 1998; Pressley et al., 2002) and occasions for students to talk to a teacher or one another about their responses to reading with guided feedback (Fielding & Pearson, 1994).

#### *Summary of Literature Related to Reading Comprehension*

The reading comprehension literature is extensive and multiple definitions are available. As with scaffolding, much of the research on reading comprehension is focused on teaching strategies that enhance student performance in understanding reading based on norm-referenced state achievement measures and nationally standardized assessments in reading comprehension. Teachers with high student pass rates over time, or who had significant student pre/post gains in reading comprehension incorporated at least some of the following strategies: use of graphic organizers, cooperative learning, teacher feedback and modeling of specific skills, open-ended higher level questions, discussion about literature, the selection of rich and varied readings with extended time for reading, and the incorporation of multiple opportunities to write about literature.

## Thinking Skills

### *Introduction*

National assessments in reading have shifted from objective or multiple choice questions to an emphasis on open-ended questions or extended response questions to enhance the evaluation of critical thinking (Sarroub & Pearson, 1998). However, students perform even less proficiently when open-ended questions are presented, including those students who are considered to be advanced on less difficult measures (Allington, 2001). Although bright students are capable of higher level thinking, neither they nor their teachers, understand how to bridge the gap between higher and lower level thinking skills without the necessary framework or models to guide the thinking processes (VanTassel-Baska & Bracken, 2005). The gap is even more severe for students of poverty (Taylor, Pearson, Peterson, & Rodriguez, 2003; Wharton-McDonald, Pressley, & Hampston, 1998). Observations in classrooms of exceptional literacy teachers suggest that more can be done to promote literacy, even in the best classrooms (Pressley et al., 2001; Taylor et al., 2002). When comparing the strategies of excellent reading teachers in schools of poverty versus wealthier schools, the largest difference between the two is the frequency and level of higher level thinking processes incorporated in wealthier school districts and the lack of higher level thinking questions and activities in lower poverty schools (Taylor et al., 2003).

Models are obviously needed to provide students with the necessary bridges from lower level to higher level thinking skills and to help teachers in planning instruction that moves students from lower level to higher level skills. The conceptual

framework of Bloom's Taxonomy provides such a model that assists teachers with writing objectives and planning lessons that engage students in moving from lower level to higher levels of understanding in language arts. Two additional models, Paul's Reasoning Model (1991) and the Taba Conceptual Model (1962) also provide a framework for students and teachers to maintain higher levels of thinking, although the necessary scaffolding from lower level to higher level skills is absent, as in Paul's Reasoning Model, or exclusive to only one type of thinking as in Taba's model. Both of these models are incorporated in the *Jacob's Ladder Reading Comprehension Program*.

#### *Literature on Critical Thinking: An Overview*

Much of the literature reviewed on thinking is less empirical and more explanatory, including definitions of thinking, a cry for the incorporation of thinking skills in schools, and posited models for thinking. Empirical data on the impact of thinking is varied and it is difficult to make generalizations because each study is conceptualized by a specific thinking construct, targeted thinking skills, or a specific model (Cotton, 1991). Moreover, like reading comprehension and scaffolding, there are other considerations that impact students' thinking and academic progress including implementation procedures, the quality of teaching, administrative support, match of student and program, and the level of fidelity to a given model (Sternberg & Bhana, 1986).

Cotton (1991) conducted a review of the research on thinking skills and the effects of thinking on student achievement. She found that that regardless of the thinking construct or definition used, most thinking skill programs made a positive

difference in student achievement, especially when specific thinking skills were targeted. Cotton (1991) also found that in most of the quasi-experimental research studies available on critical thinking, students who had exposure to thinking skill strategies or specific programs outperformed the control group in achievement and ability measures. Likewise, several specific programs are effective in producing an increase in thinking skills and achievement of students as long as teachers have appropriate training in thinking and program implementation. Effective teaching behaviors found in the research include asking higher-level questions, increasing wait time after asking questions, and teacher redirection of inappropriate student responses (Cotton, 1991). Moreover, the greatest barriers to teaching thinking skills are the amount of time and a lack of a positive classroom climate that encourages risk-taking.

There is some debate in the literature regarding the issue of whether critical thinking is subject-specific or general. For example, Ennis (1989) described critical thinking as a set of behaviors or dispositions that are consonant across multiple subject areas. In other words, critical thinking is transferable to all disciplines. However, McPeck (1990) suggested that while there may be some critical thinking skills that are consonant among varied disciplines, different fields (or subject areas) require different types of reasoning, argument, and critical thinking. Therefore, applying specific thinking skills to a discipline may be a more valid way of viewing the construct.

The time involved in measuring critical thinking behaviors may also be an issue. Some posit that general critical thinking dispositions develop over time and may not be detected in a shorter intervention (Abrams, 2005; Cotton, 1991). A longitudinal study using an innovative language arts curriculum supports the notion

that general thinking behaviors increase over time with added exposure to high level curriculum that emphasizes critical thinking (VanTassel-Baska & Bracken, 2005; VanTassel-Baska, Bracken, Feng, & Brown, in press).

The next section of the literature review on critical thinking will focus on the various models used for critical thinking and the research-based associated with each.

### *Bloom's Taxonomy*

Bloom's Taxonomy of Educational Objectives (Anderson & Krathwohl, 2001; Bloom, 1956) is a framework of objectives that move from lower order to higher order skills ranked in terms of complexity of thinking and understanding. In recent years, Anderson and Krathwohl (2001) reorganized the hierarchical framework making it two-dimensional. The first dimension is almost identical to that of Bloom's Taxonomy but is now labeled "cognitive processes." The added dimension, not included in Bloom's Taxonomy, is the "knowledge dimension".

Bloom's Taxonomy was selected as the conceptual framework for this study due to the intentional scaffolding from lower-level to higher level skills, as described in Chapter One. Lower level thinking skills, those of which are mostly viewed in classrooms (Pressley et al., 2001; Taylor et al., 2002), include the three lowest levels of Bloom's Taxonomy. The skill emphases are focused on remembering or recalling information, paraphrasing author's messages, comparing and contrasting, and applying skills to previously learned information, all of which are considered to be lower level thinking skills. The less observed skills, which incorporate the higher levels of thinking as explained by Bloom's Taxonomy, require that students organize text in a different way, differentiate between important and unimportant information, evaluate

the author's purpose or make judgments regarding multiple alternatives or perspectives, and hypothesize new alternatives or create a new structure, pattern, or idea not previously observed.

Although Bloom's Taxonomy is widely recognized and implemented internationally (Anderson & Krathwohl, 2001) little empirical data exist on the impact of the model on student achievement. Perhaps this is attributed to the number of additional variables involved when attempting to measure thinking.

Leveled questions and tasks indirectly related to the model have been measured and found to improve student achievement test scores, especially in reading (Anderson & Biddle, 1975; Beck & McKeown, 2001; Kintsch, 2005). For example, Kintsch (2005) designed a software program using higher level thinking skills in vocabulary and boasts higher student achievement based on use of the questioning strategies which move from text-based questions, to summarization, and then inference. Similarly, Beck and McKeown (2001) suggest that higher-level questions like those from Bloom's Taxonomy increase student comprehension skills as measured by student achievement tests in reading. Anderson and Biddle (1975) found when teachers require students to respond to essay questions that are more open-ended (higher level Bloom's) as compared to factual multiple choice questions (lower level Bloom's), the students who had practice with essay questions performed higher on pre-post assessments in specific content areas.

One meta-analysis of experimental research studies examined the relationship between the types of questions teachers asked and student achievement (Redfield & Rousseau, 1981). In each of the studies teacher questions were coded as lower level

(factual and recall) or higher level (manipulation of information and supporting a response). The meta-analysis posited the highest effect in student achievement when teachers used higher-level questions regardless of the content area.

A more recent, larger-scaled qualitative study found similar results. Teachers were ranked as effective or ineffective based on their students' previous test scores. Both groups of teachers were observed. The questions they asked were coded by researchers into higher order and lower order question categories and then analyzed. Researchers found that the most effective teachers (those who produce positive academic gains in student achievement test scores) asked a greater amount of higher level thinking questions than less effective teachers (Taylor et al., 2002).

#### *Paul's Reasoning Model*

Paul's Reasoning Model (1992) focuses on the upper levels of Bloom's Taxonomy, although different verbiage is used to explain similar concepts. Paul ascertains that persons who reason through a situation as a critical thinker follow similar processes. They consider multiple perspectives or points of view, collect evidence, make inferences based on the evidence, examine assumptions of themselves and others, determine implications and consequences, and conceptualize the main ideas or themes, distinguishing important information from non-important information. These skills cut across all disciplines and are not specific to reading. Paul and Elder (2004) have created checklists and questions for a variety of disciplines to assist teachers in teaching and incorporating critical thinking.

Empirical evidence for Paul's Reasoning Model as a way to increase critical thinking in students is limited. Significant achievement gains in reading and social

studies have been documented when Paul's Reasoning Model is used as part of a larger curriculum study in language arts and social studies (Feng, VanTassel-Baska, Quek, Bai, & O'Neill, 2004; VanTassel-Baska, Little, Rogers, & Feng, 2002; VanTassel-Baska, Zuo, Avery & Little, 2002). Another national, longitudinal study found that when Paul's Reasoning Model was used and explicitly taught as one of several models in a language arts curriculum, students who are exposed to the curriculum show significant and practical gains in critical thinking (Bracken, VanTassel-Baska, Brown, & Feng, 2007).

#### *Rosenblatt's Critical Stance Model*

Rosenblatt (2004) posited a model specific to reading that emphasized a teacher's critical stance and targets critical thinking and reasoning as skills to be addressed. Critical stance is defined as what a teacher does to help students move beyond the aesthetics of the text (e.g., personal connections and lived experiences of a writing) to a more efferent stance (e.g., examination of values, beliefs, inferences, and concepts conveyed within a text). She suggests similar dispositions to Paul's Reasoning Model (e.g., determining varied points of view, implications and consequences, concepts), all of which also align with the upper levels of Bloom's Taxonomy. In order for teachers to develop a critical stance and consequently help student develop a critical stance, students must be exposed to complex problems in literature, examine and evaluate multiple perspectives, determine the essence of an author's message, and discuss their personal perspectives in collaboration with others.

As with the other models, Rosenblatt's model emphasizes widely accepted findings based on effective teacher behaviors, but empirical data specific to her model

are limited. Most of the literature focuses on teacher uses of the model and implications and ideas for classrooms (e.g., McLaughlin & DeVogd, 2004).

### *Taba's Conceptual Thinking Model*

Taba (1962) identified a very specific process of conceptual thinking and inductive reasoning that, like Bloom's Taxonomy, incorporates lower level to higher level skills, hierarchically. The first two levels of Taba's model include brainstorming details and classifying those details into categories, both aligned with the recall and understanding stages of Bloom's revised taxonomy. Recalling and listing facts about a specific story, situation, or content area and classifying those into categories become fundamental processes in providing a foundation for students to move to the highest level of abstraction, that of evaluating categories to create principles in a specific content area, discipline, or specific subject matter. The creation of generalizations is the highest level and is typically associated with broad-based principles or laws of a discipline.

### *Summary of the Literature on Thinking*

The literature on thinking is widespread and includes a variety of definitions and practices to enhance thinking in students. Regardless of the thinking model, the research suggests that when thinking skills are taught to students, their level of achievement increases if there is fidelity in teaching the model and professional development. However, the increase in thinking is based on specific and targeted thinking skills and may develop over time. Likewise, students of teachers who use higher-level questions (as defined by the upper levels of Bloom's Taxonomy) show

increased gain in content-area achievement as measured by standardized achievement assessments.

## Disadvantaged Populations: Students of Poverty

### *Introduction*

The literature on students of poverty and reading has increased since the inception of the No Child Left Behind Act, the National Reading Panel Report, and Reading First, which provides grant money for research-based practices in reading, with funding preference given to low-performing and Title I school districts. The data pertinent to reading and students of poverty could be categorized into three major strands: 1) issues common for students of poverty and a need for intervention, 2) strategies of effective teachers and schools, and 3) assessment. Each strand will be discussed separately.

### *Issues and Needs*

There is a wide achievement gap between students of poverty (as defined by free and reduced lunch qualification) and those students who are not in poverty. The National Center for Education Statistics (NCES) suggests that 16.7% of school age children are in poverty. Of the 16.7% in poverty, the majority are Black, followed by Hispanic. These students score much lower on national and state achievement assessments in all core academic areas than their higher socio-economic counterparts.

Many children of poverty have less educated parents and therefore, are exposed to less modeling of appropriate reading, questioning, and literacy. How parents model reading is critical to how students learn to read and process information. Rothstein (2004, p. 4) wrote:

“When working-class [those in poverty] parents read aloud, they are more likely to tell children to pay attention without interruptions or to sound out words or name letters. When they ask children about a story, questions are more likely to be factual, asking for names of objects or memory of events. Parents who are more literate are more likely to ask questions that are creative, interpretive or connective, like “what do you think will happen next?” “Does that remind you of what we did yesterday?” Middle-class parents are more likely to read aloud, to have fun, to start conversations, as an entrée to the world outside.”

Rothstein (2004) also suggested that middle-class parents model problem solving, involve their children in conversation, and exude a strong sense of efficacy, which is mimicked by their children. Lower-class parents, on the other hand, expect children to be “seen and not heard,” allow fewer opportunities for negotiation, problem-solving and conversations, and enrichment.

Children of poverty are less likely to be identified as gifted and served in gifted programs, and more likely to be identified as special needs or having a learning disability (Donovan & Cross, 2002; Ford, 1995; USDOE, 1993). If students of poverty are identified as gifted, serving these students in traditional gifted programs causes several issues of concern if the curriculum is not adjusted to provide remediation and resources to fill in the gaps missing due to a poverty lifestyle (Slocumb & Payne, 2000). It is recommended that students of poverty be grouped together for gifted services and provided earlier intervention that will minimize disadvantages. “The cycle of poverty can be broken through education and meaningful relationships” (Slocumb & Payne, 2000, p.4).

There are several issues that arise when considering teaching students of poverty to read. First, many students of poverty begin school with little exposure to print, reading, and thinking processes needed for success in a school system (Slocumb & Payne, 2000). These students have not had the same opportunities as their middle-class counterparts, yet they are expected to adapt to a different value system and are measured on the same scale as students from middle-class families and must function without the necessary exposure to these values and thinking skills necessary for success in school.

Second, students of poverty are misunderstood by school systems and teachers (Ford & Harris, 1999; Payne, 1995). Teacher views can potentially inhibit the future success of students of poverty by blocking them from appropriate programming or placing them in appropriate settings because their unique gifts are not understood (Donovan & Cross, 2002; Peterson, 1999). Payne (1995) outlined differences in how students of poverty think about time, education, destiny, language, family structure, love, and motivation as compared to their middle and upper class counterparts. She suggested that if students of poverty are to be successful, educators must understand the differences in how students of poverty think about their lives and adapt their teaching accordingly. For example, students of poverty use language as a means of survival. It is intended to be casual and informal. Therefore, these students have not been immersed in a formal language-rich environment needed for successful acclimation to reading and thinking that is celebrated by schools. Due to a lack of resources, many times students do not have access to books, especially rich literature selections. Additional scaffolding is needed to bridge the gap between lower level

and higher level thinking skills and belief system differences beginning as early as possible in a child's school career (Payne, 1995; Rothstein, 2004). Likewise, teachers need to build relationships with students of poverty since relationships are an important part of the culture (Slocumb & Payne, 2000). Without relationship building and the recognition of student strengths important in their particular system, academic success is less likely.

A survey of faculty in teacher preparation courses concurs with Payne's call for increased understanding of students of poverty (Lyman & Villani, 2002). Four hundred eight faculty members responded to a questionnaire based on their understanding of students of poverty and how the issue is addressed in their classes. Results found that 37% of the respondents felt that a deeper understanding of issues of poverty was important. Fewer than 13% incorporated different activities or emphases for students of poverty. Fewer than 20% included poverty or diversity as a topic in any of their education courses. The researchers concluded that more must be done in teacher preparation programs to help educators understand the special needs of students of poverty and how these needs can be addressed. Likewise, faculty members and teachers need to become aware of their own understandings and beliefs regarding students of poverty and incorporate additional opportunities for cross-cultural awareness in teacher education programs (Ford & Harris, 1999). When teachers consider cultural differences of students, and are trained to do so, there is greater chance that student performance will be positively affected (Frasier & Passow, 1994).

Third, students of poverty may not learn the same way as middle-class students (Ewing & Yong, 1993). Many times, students of poverty prefer group work to discuss and share ideas instead of working in isolation or memorizing facts they believe to be unnecessary (Webb, 1998). Interdisciplinary units that allow student choice, options for studying various cultural groups, and the incorporation of different perspectives is one way to accommodate these differences (Banks, 1993; Sleeter, 1990) and recognize similarities across different groups (Gomez, 1991).

A monograph from the National Research Center for Gifted and Talented related to students of poverty or from culturally diverse backgrounds outlines the following strategies as appropriate to reading and incorporation in the language arts classroom (VanTassel-Baska, 2003b):

1. Presence of information on various cultural groups including contributions of minority groups. This emphasis can be achieved through the inclusion of readings representing many cultural groups and bibliotherapy selections that feature minority role models.
2. Use of expressive activities (oral language, movement, artistic) that require creative synthesis. This emphasis can be addressed through providing options for communication approaches about independent group learning projects as well as the type of projects students may do.
3. Use of interactive strategies. This emphasis may be incorporated by the employment of collaborative learning opportunities, frequent use of discussions, and a value placed on group learning.
4. Use of analogical reasoning. This emphasis can be addressed by the direct

teaching of analogies and by building activities into task demands that emphasize key understandings and comparisons within a text.

Finally, alternative assessments should be used for students of poverty so teachers may better monitor achievement in the classroom or place students in special programming (Callahan, 2005). Classroom assessments include product and performance-based options that allow student choice in sharing the information acquired through drama, art, song, or other modes of expression besides writing. When teachers include a variety of options in the classroom, it encourages student interest and allows the teacher to observe a wider range of abilities for students of poverty (Ford & Trotman, 2001).

Performance-based assessments allow opportunities for teachers to view student growth and could serve as an alternative means to identifying promising learners for special programs (Callahan, 2005; Gallagher, 2006). Two examples of successful use of performance-based assessments include Project Athena and Project Star. Project Athena is a language arts-based intervention study with promising learners in Title I schools. One objective of the project is to “develop and implement instrumentation sensitive to low socio-economic learners for purposes of identification and assessment of learning” (VanTassel-Baska & Bracken, 2005, slide 3). The project uses performance-based pre and post testing to measure gains in literary analysis, interpretation, and persuasive writing. After two of the three years of curriculum implementation, significant results were found with students in the treatment group on the performance-based assessments in both literary analysis and persuasive writing (VanTassel-Baska & Bracken, 2005). This replicates earlier studies that revealed

growth gains within the gifted population (Feng et al., 2004; VanTassel-Baska, Zuo et al., 2002).

Performance-based assessments may also be promising in gifted identification of students of poverty (Hadaway & Marek-Schroer, 1992; Han & Marvin, 2000; VanTassel-Baska, Johnson, & Avery, 2002), although limited data exist. The study with the most empirical data comes from a statewide initiative, Project Star, in South Carolina. Performance-based measures that incorporated non-verbal thinking tasks and more creative writing and verbal tasks were created and piloted with technical adequacy of .89. When using performance-based measures, approximately 17% more students from low socio-economic backgrounds were identified as gifted, as compared to more formal measures of identification (VanTassel-Baska, Johnson, & Avery, 2002).

#### *Summary of Poverty*

Students of poverty typically begin school without the necessary skills to be successful. The achievement gap on standardized assessments is large between students of poverty and their wealthier counterparts. Alternative resources, instruction, and assessments are necessary to help these students become successful.

#### Reform Models and Education Innovation in Title I Schools

Federal programs and financial resources have been allotted specifically for Title I schools to encourage systemic reform and success (NCLB, 2002). In addition research on reform models and effective practices in Title I schools are being studied. Michigan State University (2004) conducted a review of the literature on effective elementary schools in poverty areas. It was found that effective schools of poverty

had effective leaders, effective teachers, a strong approach to teaching literacy, and strong home-school relationships. Each is further defined. Effective leaders were found to have a clear school mission, ongoing staff development, communication and collaboration networks among classroom teachers and specialists and among varied grade level teachers, and a more experienced principal. In addition the school leaders allocated appropriate resources for training, materials, and additional instructional time. Effective teachers were said to have an awareness of purpose of what they were teaching, high student expectations, a positive classroom climate, and appropriate use of time for instructional purposes.

When examining effective reading programs in general, effective teachers and their use of instructional strategies (Taylor et al., 2000; Wharton-MacDonald, Pressley, & Hampston, 1998) as well as effective reading programs or reform models (American Federation of Teachers, 1998) are the main catalysts for student and school success.

### *Effective Teachers and Schools*

High poverty schools are those that enroll over 20% of students on free and reduced lunch (NCES, 2004). These schools tend to score lower on achievement tests, especially in reading. However, teacher differences can be instrumental in student success in reading.

A twelve-year study collected on the national Schools and Staffing Surveys (SASS) from the National Center for Education Statistics were collated to determine whether schools of high poverty had high quality teachers as defined by certification in primary teaching area, teaching experience, and teacher attrition. In comparison to

schools with higher SES, the lower SES schools had more teachers not certified in the area they were teaching, had more temporary certifications, had the highest rate of new teachers, and the highest rate of teacher attrition (Shen, 2003). Similarly, in high poverty schools, many lessons are teacher-led with little time for sustained reading, higher level questioning, and application of what was read (Stringfield, Millsap, & Herman, 1997). More time was also spent on student discipline, and more students were retained in the same grade level at the end of the school year.

Regardless of the setbacks, some schools and teachers in high poverty environments are more successful. An analysis of the literature on effective schools, based on the past eight years of data with students of poverty, was conducted by the Center for Public Education (2006). The following generalizations were found:

1. Attitudes of school staff are positive. There is a strong belief that all students “can and will” learn. These schools and teachers set high standards for all students, have more nurturing classroom environments, and boast higher student achievement.
2. Teachers use ongoing assessment and differentiated instruction. Teachers in high poverty schools that are successful, use pre-assessments to determine the strengths and weaknesses of students. They plan and monitor instruction according to the results and ongoing achievement and assessment of students. They also incorporate higher levels of thinking questions in their instruction and they adjust lessons and strategies that did not work.

3. Leaders ensure that the curriculum is aligned with state assessments.  
Principals and administrators ensure that teachers are teaching to the standards and outcomes that students will be accountable for at their grade level.
4. School leaders utilize shared decision-making models. School staff are continually involved in making decision about their school, especially related to curriculum and instruction, for which they are in charge.
5. There is an atmosphere of collaboration and shared responsibility for students. Reading teachers, school counselors, and special education teachers have time to meet in order to discuss student progress and successes or future successes.
6. Teachers are highly qualified. The definition of highly qualified includes “more years of experience, advanced degrees and training, professional development opportunities, and effective instructional skills”.
7. Family involvement is encouraged and acted upon. More students meet grade level expectations when schools and families work together through special initiatives, activities, and participation.

A national study of 400 Title I schools (Puma, Karweit, Price, Ricciuti, Thompson, & Vaden-Kiernan, 1997) found only five of 400 schools to be effective. Of the five effective schools identified, based on student achievement data, similarities included schoolwide Title I programs, experienced leadership, ability grouping in the elementary grades for specific subject areas, lower teacher attrition, better school climate, the incorporation of higher level thinking skills, and parent and community support.

Teacher effectiveness in schools of poverty has also been studied at the classroom level. One longitudinal study of second and third grade reading teachers found that of the 165 teachers studied, approximately one-third were found to be effective based on student achievement in reading (Anderson, Evertson, & Brophy, 1979). This higher achievement was attributed to reading groups, active learning, and follow-up questions and activities with students who had incorrect responses. Other studies with low-SES students previously in Head Start found that teachers who incorporated small group instruction, emphasized basic skills, allowed more time on task, and incorporated supervised time in seatwork were most effective in helping students make reading gains (Stallings & Kaskowitz, 1974). Teachers who teach for meaning when compared to an emphasis on skill and drill also had students who scored significantly higher on reading assessments (Knapp, et al., 1995).

Finally, similar to the findings over the past three decades, Taylor et al. (2000) found that the most effective reading teachers in schools of poverty allowed more time in small group instruction, independent reading, and time on task, and they built stronger relationships with the home. Moreover, these teachers taught phonics within the larger context of reading, instead of in isolation, employed higher level thinking skills activities and questions based on Bloom's Taxonomy, and encouraged students to write about what they read.

#### *Specific Curriculum and Reading Reform Models*

There are several nation-wide curriculum models marketed as research-based for Title I schools, with a focus on effectiveness in reading for students of poverty. In addition, multiple small-scale initiatives have been instituted and found to be effective.

Many of these models are focused on students of poverty who are struggling with basic literacy skills. Few programs focus on building higher level thinking skills for students (Tivnan & Hemphill, 2005). Instead the focus is on the National Reading Panel's recommendations of phonemic awareness, explicit phonics instruction, oral reading and practice, vocabulary instruction, and comprehension strategies (NRP, 2000). This is most likely due to the number of studies in reading, also used by the NRP for the meta-analysis, that focus on struggling readers. However, students of poverty are capable of learning at higher levels given the proper instruction and tools (Ford, 1995; Payne, 1995). The remainder of this section will focus on the seven major reform models, smaller scale models and projects, and curricula specific to advanced or promising readers.

#### *Large Scale Reform Models and Curriculum*

According to the American Federation of Teachers (AFT), there are seven promising reading curricula. Criteria for selection of programs consisted of curricula that showed evidence of high standards, effectiveness, replicability, and support structures for teachers (AFT, 1998). Most of the programs focus on students in grades K-6 or 2-6. The programs include the Cooperative Integrated Reading and Comprehension (CIRC), Direct Instruction (DI), Exemplary Center for Reading Instruction (ECRI), Junior Great Books (JGB), Multicultural Reading and Thinking (McRAT), Success for All (SFA), and Open Court Collections for Young Scholars (OC). Table 1 illustrates the comparison among each of the programs. All programs, except one, have data from quasi-experimental design studies and many also have case study data.

As suggested in the comparison table, few programs focus on higher level thinking skills or emphasize instruction for promising or advanced students. Data collection on effectiveness with the top students is also scarce or consists of small scale case studies. The only program with more generalizable results for effectiveness in promoting thinking skills for advanced students is Junior Great Books. One program, ECRI, has limited data on the effectiveness for advanced students, probably due to the emphasis on mastery learning, scaffolding, and individualized instruction. However, the ECRI results focus on comprehension instead of higher levels of thinking. None of the promising programs have data disaggregated for advanced students in Title I schools.

Another study found similar results regarding the paucity of higher level thinking skill emphasis in major reform models (Tivnan & Hemphill, 2005). Sixteen high poverty schools in Boston were studied. Each school had marked results with one of the major reform models listed to improve reading scores including Building Essential Literacy, Developing Literacy First, Literacy Collaborative, and Success for All. Of the programs studied, only Success for All was listed by the American Federation of Teachers. However, the reading programs in this study also have marked results for improving student achievement in high poverty schools and are encouraged for use as outlined by the Reading First Act. When comparing the four programs

Table 1: Comparison of Effective Reading Curriculum for Use in Title I Schools

	CIRC	DI	ECRI	JGB	McRAT	SFA	OC
Features	Grouping  Basal Reading  Direct Instruction	Grouping  Scripted Lesson Plans  Individual Coaching	Grouping  Scaffolding  Individual Differentiation (Diagnostic-Prescriptive)	Shared Inquiry  Classical reading selections  Reflection	Critical Thinking  Multicultural Literature  Modeling	Direct Instruction  Reading Groups  Ongoing Assessment	Readings arranged around themes  Explicit Phonics  Reading comprehension
Grade Levels of Focus	2-6	PK-6	1-12	K-12	3-8	K-6	K-6
Effectiveness with Students of Poverty	One case study suggests high growth for ESL Title I students in writing	One case study suggests that one school of high poverty has ranked top in their state since using the program	Yes – Effect sizes on standardized reading tests between .28 to 1.32.	Limited (Small descriptive study found an increase of 30 points on state assessments with use of the program)	N/A	Yes – Case study results suggest that high poverty schools outperform other equivalent schools after four years if using the program – Effect Size = .15-.33.	Yes - OC students in Title I schools outperformed 2 other reading programs (Based on pre-post Woodcock Johnson reading results)
Effectiveness with High Achievers	N/A	N/A	Limited (One study found significant results for gifted students)	Yes (Three studies found significant results with high achieving students)	N/A	N/A	N/A
Emphasis on Thinking Skills	No	No	Minimal (One section of the scripted lesson focuses on critical interpretation and textual analysis)	Yes (Four studies found significant results in critical thinking skills when compared to basal programs)	No	No (Focus on word attack skills and basic comprehension)	No
Effectiveness with Lowest Performing Students	Yes (Effect size of .26 in reading and .40 in comprehension)	Not Available	Yes (Effect size of .28-1.32 over 8 sites)	Yes (student gains increased over a year from the 26 <sup>th</sup> percentile to the 50 <sup>th</sup> percentile)	N/A	Yes (Highest effect sizes are with lowest performing students – 1.0)	Case study data suggest that low performing school moved from a 47% pass rate to a 79% pass rate

				when using the program)			with program implementation (especially in phonics)
Caveats	N/A	Must be faithfully implemented or achievement may decline Effect sizes greatest for K-2	Long term professional development required	Teachers need training in shared inquiry and questioning skills	*Only small scale studies available on achievement tests with show significant increase in scores for students of all levels of achievement (n=30). *Significant pre-post gains in writing on company-created assessments.	Very expensive Focus on low-achieving students in Title I schools	Training is critical to success.  Many teachers report that the teacher manuals are difficult to use.

with first graders across the sixteen high poverty schools, all programs were found to produce equally significant results in phonics and word reading. But none of the programs showed gains or potential for higher level thinking skills in reading such as meaning construction and comprehension (Tivnan & Hemphill, 2005). Therefore, “off-the-shelf” programs may need more supplemental materials to enhance higher-level thinking and comprehension skills with students in poverty.

Two additional nationally known models must also be mentioned due to the effectiveness of use with students of poverty. These models focus on the process of instruction instead of a specific curriculum: Payne’s Poverty Framework and Reciprocal Teaching. Payne’s Poverty Framework (1995) operates under the premise that if teachers understand students of poverty and teach to their specific learning styles and needs, then students will be more successful. The Framework consists of specific cognitive strategies including the use of storytelling, visual representations, metacognition, hands-on learning, and patterning; systemic interventions such as diagnostic-prescriptive approaches and mastery learning for meeting benchmarks; professional development; building relationships between home and school; and the use of special projects (Payne, 1995). Large school districts in Texas and Indiana have used this approach, incorporating all aspects of the framework into their existing curriculum. Both have shown positive results in state achievement tests over a three year period with National Curve Equivalents of more than an increase of 1.9 and statistical significance of  $p < .001$  across all grade levels, although students in the upper grades (6-10) showed the highest gains (Swan, 2004).

Reciprocal Teaching is also a process-oriented approach with a focus on reading comprehension (Palinscar & Brown, 1984). There are four major strategies of reading comprehension noted: generating questions, summarizing, clarifying, and predicting. Research on the model suggests improvement in reading comprehension for students of poverty when using standardized achievement tests (Carter, 1997). Additional research suggests that generating questions (Rosenshine, Meister, & Chapman, 1996) and clarifying (Lublinter, 2002) may be used independently for similar results as using the entire program. Data were not disaggregated for high achieving students.

#### *Curriculum & Programs for Advanced Learners of Poverty*

Most of the literature specific to gifted students of poverty focus on identification issues (e.g., Ford, 1995; Joseph & Ford, 2006), case studies specific to a targeted minority group (e.g., Hebert & Beardsley, 2001), social-emotional needs (e.g., Shumow, 1997) or special programs or services (e.g., Robinson, Gaines-Lanzi, Weinberg, Ramey, & Ramey, 2002). Large-scale curriculum efforts and specific curriculum for gifted students is limited (VanTassel-Baska, 2003b) and almost non-existent when considering the added dimension of poverty. One longitudinal study using the William and Mary curriculum units in language arts, focused on the impact of critical thinking and reading comprehension of students in Title I schools, positing gains in both areas for students who use the curriculum (Bracken, VanTassel-Baska, Brown, & Feng, 2007).

A statewide study of over 1,000 low income and minority students was also conducted using the William and Mary curriculum units. Swanson (2006) provided

professional development on models used in the William and Mary language arts and science units and encouraged teacher use of the models over three years. The William and Mary units were used as starting point and then teachers were encouraged to create their own units. Pre-post data were collected using the MAT7 and the South Carolina State Assessment. Minority and low income student achievement scores were significant ( $p < .001$ ) when using the models and curriculum in reading and science.

A national longitudinal study in language arts, SEM-R, is currently being conducted by the National Research Center for Gifted and Talented and includes diverse populations as part of the study (Reis, Gubbins, Briggs, Schreiber, Richards, Jacobs, Eckert, & Renzulli, 2004). The study focus, however, was on differentiated reading strategies used in classrooms where poverty is higher. Preliminary findings suggested that students made significant gains in reading when differentiated strategies were applied; however, they found few teachers differentiated instruction in reading, and school libraries were not equipped with advanced level reading materials necessary to promote higher level reading skills.

#### *Summary on Innovative Practices in Title I Schools*

In order for Title I schools to sustain effort, particularly in reading, the literature suggests that the teachers' use of higher order thinking skills and instructional strategies such as scaffolding, coaching, small group work, and independent reading time are essential. Likewise, many of the reform models in reading that are targeted toward Title I schools show promise in improving phonemic awareness, vocabulary, and reading comprehension. However, higher level thinking

skills and higher order task demands are lacking in the majority of the recommended models. Supplemental reading curriculum may be necessary so that teachers may infuse higher order thinking processes into the reform curricula.

## Differentiation

### *Introduction*

Differentiation is the new “buzzword” in education. Communication from The Association of Supervision and Curriculum Development (ASCD) boasts sales of books on differentiation among the highest and most popular of all products marketed. The most widely accepted definition of differentiation is by Tomlinson (1999). She defines differentiation as the way teachers respond to individual student needs. Teachers differentiate when they adjust the content, process, and products based on student readiness, interest, or learning profiles. Differentiation is a response to a set of beliefs about students, is grounded in theory, and is research-based (Tomlinson & Allan, 2000). Differentiation is intended to be student-centered, fluid, proactive, qualitative, flexible in grouping strategies, and most importantly, rooted in assessment that guides instruction (Tomlinson, 1995; Tomlinson & Allan, 2000).

Other definitions include adjustments to the Tomlinson definition and depend upon the orientation of the researcher. Researchers in special education define differentiation as providing an adequate learning climate and adjustments to the curriculum to enhance learning (Gregory & Kuzmich, 2004; Nordlund, 2003). Researchers in gifted education are more likely to espouse the Tomlinson definition but include a stronger emphasis on modifying the original curriculum and classroom

experiences to make them more challenging and meaningful for advanced learners (Feldhusen & Moon, 1995; Parke, 1995; Renzulli, 1997).

VanTassel-Baska (2003) suggests a slightly different orientation to differentiation and defines it in terms of adjusting the curriculum specifically for gifted learners. Curriculum for the gifted should be differentiated at all levels of design, including the goals of the lessons, the outcomes required of students, the activities and projects for which students engage, the strategies educators employ, the materials used, and the assessments to measure progress. To incorporate each of these levels, educators are instructed to apply the following features of differentiation to their specific content area: a) acceleration, b) complexity, c) depth, d) challenge, e) creativity, and f) abstractness (VanTassel-Baska & Stambaugh, 2006). Since *The Jacob's Ladder Reading Comprehension Program* fits best under this definition, the features of differentiation using the aforementioned model will be discussed in more detail.

1. Acceleration. Gifted students can learn at faster rates than other children with less practice time required (Colangelo, Assouline, & Gross, 2004). Therefore the pacing of the curriculum can be adjusted to decrease the speed of learning and increase the depth or to increase the speed of learning by requiring few tasks to master a standard, which allows the student to pursue advanced content. In order to effectively accelerate the curriculum, a diagnostic-prescriptive approach must be considered. In other words, pre-assessments must be administered in order to determine the skills and concepts a student already knows. Learning objectives are then based on the child's assessment.

By organizing a standard for learning at a higher level of expectation and broadening the scope, gifted learners have multiple opportunities to engage in real problem-solving behavior as opposed to rote procedures. *Jacob's Ladder* includes an accelerative component by using advanced readings for students, thus accelerating the content and expectations, while also scaffolding instruction to help students reach the appropriate levels. Moreover, a self-assessment and teacher assessment system allows students and teachers to examine progress. If patterns of achievement are high on the *Jacob's Ladder* tasks, teachers are encouraged to move students to a more complex curriculum.

2. Complexity. The complexity of a given task is determined by the level of higher order thinking skills a student is required to practice. Additional variables, multiple resources, or more difficult questions may be posed. This does not mean that students will do more of the same type of activity, however. Rather, more variables are added at a more abstract level to render the task as more complex. Complexity within *Jacob's Ladder* is added based on the movement from lower order to higher order thinking skills as defined by Bloom as part of the conceptual framework. As students move up the "ladder", the complexity of the questions increases.
3. Depth. In order to add depth to a given task, students should be required to conduct original research, develop a product of worth, and apply the concept in multiple ways. Many educators mistake adding depth to the curriculum by the creation of any product or research as interpreted by regurgitating facts from a different source. Specifically, the ladder that requires Creative Synthesis

provides the opportunity for creation of a new product of worth. Students must also grapple with conceptual meaning within other ladders, also adding depth.

*Jacob's Ladder* also adds depth to a topic by requiring gifted students to grapple with a concept, rather than a topic, consider the pros and cons of reading passages, and defend their answer through more than just regurgitation of read material.

4. Challenge. Gifted students need more challenge in the curriculum based on the advanced resources employed, the sophistication of the content discussed, the interdisciplinary connections made, and the amount of reasoning required to arrive at a conclusion about a specific topic. The level of challenge is based on the amount of reasoning required to complete a task. Two specific ladders of *Jacob's Ladder* include elements of reasoning as defined by Paul, and include helping students examine evidence from the text, consider inferences, themes or concepts, and implications and consequences given a specific, accelerated passage.
5. Creativity. To incorporate creativity into the curriculum, students should be asked to construct a model based on a concept studied, have opportunities to complete alternate tasks or products of their choosing, or represent new learning in their personal choice mediums, with an emphasis on oral and written communication to real-world audiences. Creativity must include rigorous content as represented through some type of product. Simply asking students to create or make a product is not differentiated for gifted students unless that product in some way allows for new learning that is advanced.

In the *Jacob's Ladder Reading Comprehension Program*, students have choice in the ways they craft the responses in that there is not one right answer to the task within reason.

6. Abstractness. The differentiation feature of abstractness requires students to focus on conceptual thinking within and across disciplines. They may examine the generalizations behind a specific macro-concept such as change, interdependence, systems, or patterns, for example. Students may be asked to formulate their own generalizations regarding their findings and move from concrete applications to more abstract ways of thinking about a concept or discipline. Two of the ladders of *Jacob's Ladder*, include generalizations and conceptual thinking about literature, thus adding abstractness as a component of differentiation.

#### *Empirical Evidence for Differentiation*

It is argued that little empirical evidence for differentiation exists (Schmoker, 1999). However, evidence is dependent upon the definition and whether differentiation is viewed holistically or by its individual components (e.g., pre-assessment, grouping, acceleration, higher level thinking skills).

Holistic data are limited. One qualitative study from Tomlinson (1995) followed a group of middle school teachers who were attempting to differentiate and found positive results in attitudes of students and teachers as well as a change in teaching behaviors from more whole group to more individualized instruction. However, Tomlinson and Allan (2000) report that action research data from classrooms in Alaska and Canada boasted an increase in schoolwide achievement tests

when differentiation was used. Additional qualitative data from individual classrooms has been documented and schoolwide differentiation is found to positively impact teacher change (Tomlinson & Alan, 2000; Tomlinson, Moon, & Callahan, 1998).

Specific strategies used in a differentiated classroom such as acceleration, grouping, and pre-assessment with an appropriate curriculum match posit a strong research base, especially for gifted learners. Even though some researchers suggest that grouping is inequitable (Oakes, 1995; Slavin, 1986) the deciding factor in grouping that determines whether it is inequitable or best practice has more to do with the identification procedures in place, the flexibility of the grouping, and the assurance of a high level curriculum matched to student ability than a fixed, predetermined group assignment. Meta-analyses show the effectiveness of grouping on student achievement, regardless of the ability of the child (Rogers, 1998), as long as the curriculum is differentiated (Kulik & Kulik, 1992; 1997). Without differentiated curriculum, grouping produces no effect (Rogers, 1998; Slavin, 1986).

Pre-assessment and appropriate placement of curriculum is also found to be effective. One such method has longitudinal research: the diagnostic-prescriptive approach (DT/PI = Diagnostic Testing and Prescriptive Instruction). Within this model, students are pre-assessed based on a particular set of skills, concepts, or topics. The outcome of the pre-assessment determines the instructional goals and delivery. This type of assessment is a necessary management tool to best tailor the curriculum to the gifted learner's needs, knowledge, and skills while ensuring there are no gaps in knowledge acquisition throughout the process. Written educational plans and other management documents are used to record student progress and the instructional

strategies employed. Talent search organizations such as the Study of Mathematically Precocious Youth at Johns Hopkins have utilized this type of differentiated approach for over three decades with documented success in several studies (Lynch, 1992; Swiatek, 1993; Swiatek & Benbow, 1991). The long-term effects of educational acceleration of the gifted have also been studied. Meta-analyses show positive results in cognitive development from acceleration, and no negative effects on social emotional development (Colangelo, Assouline, & Gross, 2004).

Regardless of the differentiation strategy used (individually or holistically) regular classroom teachers need intensive professional development, accountability, and support to effectively differentiate instruction. Without (and sometimes even with) this professional development, studies suggest that few regular classroom teachers differentiate instruction on a regular basis for the gifted (Westberg, Archambault, Dobyns, & Slavin, 1993; Westberg, & Daoust, 2003).

#### Summary of Literature Findings

Table 2 summarizes the findings among the different literature strands discussed in this chapter including scaffolding, reading comprehension, thinking skills, disadvantaged populations, innovative reform in Title I schools, and differentiation.

Table 2: Summary of Research Findings by Literature Strand

<b>Strand</b>	<b>Findings</b>
Scaffolding	Scaffolding is most effective in enhancing student achievement when teachers help students work collaboratively through problems, provide feedback as necessary, and ask probing questions that help students think about alternative answers, other ideas, or their own metacognitive thinking processes (Perkins, & Salomon, 1989; Zemelman, Daniels, & Hyde, 1998)

---

Students in Title I schools need more scaffolding of the curriculum in order to achieve success and increase achievement test scores (Payne, 1995; Rothstein, 2004)

Descriptive research shows that the Scaffolded Reading Experience (SRE) curriculum produces significant reading achievement gains in regular education students (Fournier & Graves, 2002; Graves & Avery, 1997; Stephens & Brown, 2005)

Effective instructional strategies associated with scaffolding and the increase in student achievement test scores in reading include the following: 1) coaching throughout the process of reading (Guthrie, 1996; Pressley et al., 2001; Taylor et al., 2000); 2) asking higher level questions in addition to lower level questions to build meaning (Knapp et al., 1995; Pressley et al., 2001; Taylor et al., 2002); 3) modeling personal thinking processes aloud (NRP, 2000; Zemelman et al., 1998); 4) providing opportunities for independent reading followed by small then whole group discussion (Campbell et al., 1997; Chin et al., 2001; Guthrie, 1996); 5) providing opportunities for teacher feedback and time for students to self-regulate and assess their own learning (Chin et al., 2001; Pressley et al., 2001; Roehler & Duffy, 1984; Zemelman et al., 1998); and 6) linking personal experiences to reading with an emphasis on process (Chin et al., 2001; Roehler & Duffy, 1984; Taylor et al., 2000).

Reading Comprehension      Instructional strategies and the teaching of key processes enhance reading comprehension (Fielding & Pearson, 1994).

Meta-analyses from quasi-experimental studies show the following strategies enhance reading comprehension achievement in students (NRP, 2000): 1) cooperative learning; 2) direct instruction on use of graphic organizers; 3) higher level questioning; 4) comprehension monitoring including teaching, coaching, and modeling of appropriate processes; and 5) use of multiple strategies together. In addition, direct, explicit instruction of reading processes, extended time for reading, and student discussions have been found to improve student reading comprehension skills (Biancaroso & Snow, 2004).

Guided student discussion of critical components of literature and higher level thinking questions is important for students to show gains in literary analysis (Beck & McKeown, 2001; Little, 2002; VanTassel-Baska, Zuo, Avery, & Litte, 2002).

---

Students show greater growth in reading comprehension and

---

	<p>understanding when given opportunities to write about literature (Appleman &amp; Applebee, 2000; Newell, 1996; Sadoski, Willson, &amp; Norton, 1997) – especially when using persuasive writing models in response to literature (VanTassel-Baska, Johnson, &amp; Boyce, 1996; VanTassel-Baska, Zuo, Avery, &amp; Little, 2002).</p>
Thinking Skills	<p>Title I school teachers ask fewer higher order thinking skills than non Title I teachers (Pressley et al., 2001; Taylor et al., 2003)</p> <p>Teachers who ask higher level questions produce significant gains in student achievement regardless of the content area, when compared to those teachers who ask recall or lower level questions (Anderson &amp; Biddle, 1975; Beck &amp; McKeown, 2001).</p> <p>Critical thinking skills take time for students to develop (Abrams, 2005; Bracken, VanTassel-Baska, Brown, &amp; Feng, 2007; VanTassel-Baska &amp; Bracken, 2005)</p> <p>When critical thinking skills are taught in a specific content curriculum, student achievement gains are higher in that content area when compared to those students who did not have exposure to thinking skills in the same content area (Cotton, 1991; Kintsch, 2005)</p> <p>The use of Paul’s Reasoning Model shows increased ability in students’ general thinking behaviors over time (VanTassel-Baska, Bracken, Feng, &amp; Brown, in press).</p>
Disadvantaged Populations	<p>Children of poverty are less likely to identified as gifted (Donovan &amp; Cross, 2002; Ford, 1995; Slocumb &amp; Payne, 2000;)</p> <p>Students of poverty do not perform as high on achievement assessments as their higher SES counterparts (NCES, 2001)</p> <p>Performance-based assessments may be a more accurate way to measure achievement performance of disadvantaged populations (Callahan, 2005; Ford &amp; Trotman, 2001; VanTassel-Baska, Johnson, &amp; Avery, 2002; VanTassel-Baska &amp; Stambaugh, 2007) – including students of gifted (Hadaway &amp; Marek-Schroer, 1992; Han &amp; Marvin, 2000; VanTassel-Baska, Johnson, &amp; Avery, 2002).</p> <p>Students of poverty need more opportunities for discussion and communication, hands-on learning and choice of projects and activities (Banks, 1993; Rothstein, 2004; Sleeter, 1990; Slocumb &amp; Payne, 2000; VanTassel-Baska, 2003b).</p>

---

---

**Innovation and Reform in Title I Schools** Successful Title I schools (as ranked by high student achievement on state assessments) show the following attributes (Center for Public Education, 2006): 1) positive attitudes of school staff; 2) ongoing assessment and differentiated instruction; 3) alignment of curriculum to state assessments; 4) shared decision making models; 5) collaboration and communication among staff; 6) highly qualified teachers; 7) family involvement; 8) small group instruction (Puma et al., 1997); 9) skill-based emphasis (Knapp et al., 1995); 10) time on task in reading (Stallings & Kaskowitz, 1974); 11) experienced leadership (MSU, 2004; Puma et al., 1997); 12) flexible ability grouping (Puma et al., 1997); 13) teaching scaffolding and modeling of appropriate behaviors (Taylor et al., 2000); and 14) the use of higher level thinking questions (Taylor et al., 2000).

Reform reading curriculum found to be effective in Title I schools focus on skills such as phonemic awareness and word recognition (Tivnan & Hemphill, 2005). There was limited success found in student achievement in reading comprehension among reform curricula (AFT, 1998). Only two reform curricula showed any achievement data on increased higher level thinking skills: Junior Great Books and Exemplary Center for Reading Instruction (ECRI).

Payne's Poverty Framework (Payne, 1995; Swan, 2004), Reciprocal Teaching (Carter, 1997; Lubliner, 2002; Palinscar & Brown, 1984) and the William and Mary Language Arts Units (Swanson, 2006; Bracken, VanTassel-Baska, Brown, & Feng, 2007; VanTassel-Baska, Bracken, Feng, & Brown, in press;) show promise in student reading achievement and critical thinking gains for students in Title I schools.

**Differentiation** Few teachers use differentiated teaching strategies in reading (Reis et al., 2004; Westberg et al., 1993; Westberg & Daoust, 2003).

When differentiated instruction is used, students show significant gains in reading (Eckert, 2005) and on school-wide achievement tests (Tomlinson & Allan, 2000).

Effective differentiated strategies include flexible ability grouping based on targeted skill development (Kulik & Kulik, 1997; Rogers, 1998); diagnostic-prescriptive approaches (Swiatek, 2002); and acceleration (Colangelo, Assouline, & Gross, 2004).

---

Applications of Literature Findings to *Jacob's Ladder Reading Comprehension*  
Program

Based on the literature review, *Jacob's Ladder Reading Comprehension Program* incorporates many of the research-based strategies as outlined by major national reports, meta-analyses, and studies on reading comprehension, higher level thinking skills, scaffolding, differentiation, and process skills found to be effective with students of poverty in reading, including advanced readers.

Table 3 illustrates the relationship among the features of *Jacob's Ladder Reading Comprehension Program*, the relevant citations based on the features of *Jacob's Ladder* and the major studies or theories that support the curriculum features.

Table 3: *Jacob's Ladder* Features and Research Citations

<b>Features of <i>Jacob's Ladder</i></b>	<b>Relevant Literature</b>
Open-ended questioning and thinking (e.g., concept development, reasoning, and higher level thinking)	Anderson & Biddle, 1975 Feng, VanTassel-Baska et al., 2004 Palinscar & Brown, 1984 Paul, 1992 Taba, 1962 Taylor et al., 2000 VanTassel-Baska et al., 2002
Teacher modeling, direct instruction, stance, and coaching	Guthrie, 1996 Palinscar & Brown, 1984 Taylor et al., 2000 Zemelman, Daniels, & Hyde, 1998
Small group discussion and inquiry (independent reading following by collaborative small and whole group discussion)	Campbell, Voelkl, & Donahue, 1997 Chin, Anderson, & Waggoner, 2001 Guthrie, 1996 Villaume & Brablam, 2002
Ability grouping with differentiation	Kulik & Kulik, 1992; 1997 Rogers, 1998 Tomlinson, 1999
Building on lower level thinking and	Anderson & Krathwohl, 2001

---

moving to higher level thinking	Fournier & Graves, 2002 Knapp et al., 1995 Pressley et al., 2001 Taylor et al., 2002 Tomlinson, 1999 VanTassel-Baska & Bracken, 2005
Self-assessment opportunities and regulation with teacher feedback	Chin et al., 2001 Pressley et al., 2001 Roehler & Duffy, 1984 Zemelman, Daniels, & Hyde, 1998
Writing about reading	Beck & McKeown, 2001 Roehler & Duffy, 1984 Taylor et al., 2000 VanTassel-Baska & Stambaugh, 2006
Advanced or accelerated curriculum opportunities	Benbow and Stanley, 1983 Colangelo, Assouline, & Gross, 2004 Feldhusen & Moon, 1995 Renzulli, 1997 Swiatek & Benbow, 1991 Tivnan & Hemphill, 2005 Tomlinson, 1999 VanTassel-Baska & Stambaugh, 2006 VanTassel-Baska, 2003a
Opportunities for creative synthesis and hands-on learning	Payne, 1995 Payne & Slocumb, 2001 Robinson, 1990 Swan, 2004
Focused professional development (especially in schools of poverty)	Reis et al., 2004 Swanson, 2006 Taylor et al., 2002 Westberg, Archambault, et al., 1993 Westberg, & Daoust, 2003

---

### Conclusion

Literature on reading comprehension, scaffolding, thinking skills, and differentiation imply that student achievement is linked to effective instructional strategies within a curriculum that promote higher level thinking skills and provide

opportunities for student discussion, practice, and teacher modeling in appropriate reading behaviors. Many of the curriculum reform models for reading used in Title I schools fail to incorporate higher level thinking skills and consequently do not produce higher-level achievement gains in reading. Limited data exist related to programs in reading that focus on critical thinking skills and reading comprehension that are found to be effective in Title I schools. The next three chapters focus on the study of an innovative supplementary reading curriculum, *The Jacob's Ladder Reading Comprehension Program*, and its effects on third, fourth, and fifth grade rural Title I students' reading comprehension and critical thinking achievement.

## Chapter III

### Methodology

#### *Introduction*

The purpose of this quasi-experimental study was to conduct a pilot of the *Jacob's Ladder Reading Comprehension Program* to determine the academic effectiveness of the program on students' reading comprehension and critical thinking. The demographic population of interest in this study was 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> grade students in heterogeneous classrooms who attend rural, Title I schools. This chapter provides an overview of the site selection demographics, participants, instrumentation, and procedures for data collection and analysis. Table 4 provides a synopsis of the research questions, instrumentation, and data collection procedures of this study.

#### *Site Selection*

Two rural Title I school districts in Ohio were selected to participate in the study based on their similarities in demographics as outlined by the State of Ohio Department of Education. The Ohio Department of Education reports district statistics and sorts districts into comparison groups based on each school's average daily membership (ADM), poverty level – defined by the percent of those on free/reduced lunch, percent of the professional occupations within the school district, district median income, percent of the population with a college degree, the percent of agriculture property, population density, and percent of minority students. Both school districts posit high poverty rates and a low percentage of minority students. The participating elementary schools in each district are ranked "Excellent" by the state of Ohio. This ranking is the highest a school may achieve based on student

Table 4: Overview of research questions and data collection procedures

<b>Question</b>	<b>Instrumentation</b>	<b>Collection</b>
What differential effects does the <i>Jacob's Ladder Reading Comprehension Program</i> have on students' reading comprehension?	Survey battery reading comprehension subtest of the Iowa Test of Basic Skills (ITBS)	Pre and post test for both experimental and comparison groups
What differential effects does the <i>Jacob's Ladder Reading Comprehension Program</i> have on students' critical thinking and higher level thinking skills in literature?	Test of Critical Thinking (TCT)	Pre and post test for both experimental and comparison groups
	Performance Based Assessment (PBA)	Pre and post test for experimental group
Are there differential effects of the <i>Jacob's Ladder Reading Comprehension Program</i> on critical thinking and reading comprehension by gender, poverty level, gifted status, and grade level?	Survey battery of the reading comprehension subtest of the Iowa Test of Basic Skills (ITBS)	Pre and post test for both experimental and comparison groups
	Test of Critical Thinking (TCT)	Pre and post test for both experimental and comparison groups
What are the perceptions of teachers and students regarding the use of the <i>Jacob's Ladder Reading Comprehension Program</i> after implementation?	Standardized, open-ended interview protocol	One teacher focus group and one student focus group by District

achievement data, growth, attendance, and other factors. Therefore, these districts and buildings are prime candidates for piloting an innovative practice such as *Jacob's Ladder* as they are not at-risk of failing and show a history of effective innovation and success as measured by state achievement tests. Table 5 illustrates the similarities of each district and participating buildings.

*Table 5: Comparison of School Districts and Buildings in the Sample*

<b>District Building</b>	<b>ADM</b>	<b>% Poverty</b>	<b>Academic Ranking</b>	<b>% Minority</b>
District One	1,919	35.5	Effective	5.7
Elementary	751	45.6	Excellent	6.1
Middle	575	33.7	Effective	5.8
District Two	2,462	44.5	Effective	2.4
Elementary	798	42.7	Excellent	3.1

District One is a rural district close to a major Ohio city. The school district is comprised of approximately 1,900 students and includes one elementary school, one middle school, and one high school. The elementary school houses students Pre-K through 4<sup>th</sup> grade and the middle school houses 5<sup>th</sup> through 8<sup>th</sup> grade students. There are four classrooms per grade level in each building. Gifted services at the elementary school are provided in language arts and in mathematics on a daily basis. Gifted students are not cluster grouped in the elementary classroom with the exception of one fourth grade classroom. At the middle school level, fifth grade students who are identified as gifted in language arts or in math are homogeneously grouped in self-contained classrooms for each of those subject areas, respectively.

District Two is located near the foothills of the Appalachian Mountains. The closest major city is approximately 70 miles away. The district is comprised of approximately 2,500 students. There are two elementary schools, two middle schools, and one high school. The largest elementary school was selected to participate in the study because the smaller elementary school did not have enough classrooms per grade level for comparison and experimental groups. The selected elementary building houses students in grades K-5. There are four classrooms per grade level.

One classroom per grade includes the special education students and the other three include a cluster group of gifted students in a classroom with regular education students. Gifted students in grades 3-5 are served in the classroom via cluster grouping and are pulled out of the classroom on a variable basis to work with the teacher of gifted.

### *Selection of Participants*

Principals at each participating building were instructed to follow a purposeful random selection of teacher classrooms by randomly drawing the names of two teachers per grade level (third, fourth, and fifth) to be the comparison teachers. The remaining two teachers per grade level would be experimental. If there were two cluster grouped classrooms of gifted, one was to be randomly selected as comparison and the other experimental. Special education classrooms, which are discussed later in this chapter as a delimitation, were not be considered as part of the study sample.

As with most intact classrooms there were complications to this process that were dealt with on a situational basis. For example, in District One the fifth grade had three regular education classrooms and one self-contained classroom of students who are gifted in language arts. Since gifted was one of the factors in this study, the self-contained fifth grade language arts classroom was automatically selected as one of the experimental classrooms. The remainder of the fifth grade classrooms was randomly selected as either comparison or experimental.

In District Two, grades four and five systematically incorporated departmentalization by content area, meaning that one teacher taught one or two content areas to three of the four grade level classrooms. The teacher who taught the

majority of language arts classes was selected as the experimental teacher. The gifted cluster group classroom was selected as the experimental group. Of the two remaining classrooms, one was randomly selected as the second experimental classroom for pre/post testing. One classroom remained that was not being taught language arts by the same teacher. This classroom became the comparison classroom by default. Therefore, in District Two, there are two experimental classrooms and one comparison classroom each for grade four and for grade five.

Table 6 provides demographic information for the entire sample by condition (e.g., comparison and experimental classrooms) and factors (gifted, those on free/reduced lunch, grade, gender). As discussed in Chapter One, gifted students for this sample included those students who scored at or above the 95<sup>th</sup> percentile in reading on a standardized achievement test and/or an IQ two standard deviations from the test mean less the standard error measure. Free and reduced lunch students are labeled as such based on their family income level. Both the gifted and free-reduced lunch student lists were obtained from each school's database.

*Table 6: Number of Participants by Factor and Condition*

Condition	Free Reduced		Gender		Gifted		Grade		
	N		n		n		n		
	Yes	No	Male	Female	Yes	No	3	4	5
Comparison (n=222)	63	149	116	106	11	201	86	75	61
Experimental (n=273)	72	196	133	140	44	226	91	97	85
Total (n=495)	135	345	249	246	55	427	177	172	146

A total of 495 students participated in the study, including both comparison (n=222) and experimental (n=273). Similar numbers of students by district

participated. In District One, 252 students participated in the study and in District Two, 243 students participated.

### *Instrumentation*

The instrumentation in this study included two standardized assessments with strong technical adequacy (one to measure reading comprehension and another to measure critical thinking), a researcher-developed performance based assessment, a treatment fidelity form, and a focus group open-ended standardized question protocol. The two standardized assessments were used as pretest/posttest measures for comparisons between and within the comparison and experimental groups. The performance-based assessment was also used in a pretest/posttest format but with the experimental group only. The treatment fidelity form was used as a tool to observe experimental teachers twice during the study implementation. The focus group questionnaire was used with all experimental teachers in both districts and with two students per experimental classroom per district. Each is discussed separately.

### *Reading Comprehension Measure*

The Iowa Test of Basic Skills (ITBS) Survey Battery Reading Comprehension subtest was used to measure pre/post student performance in reading comprehension. The ITBS is a standardized assessment that has been in existence for over seventy years and is widely used throughout the United States to assess student performance in core content areas. Separate tests are designed for students in grades K-8. It is also one of the standardized tests placed on the Ohio approved test list to identify gifted students.

The reading portion of the assessment is administered in approximately 30 minutes. Multiple-choice questions are used as the basis for assessing student performance. Students read fiction and nonfiction passages including myths, fables, poetry, biographical entries, indexes, and core-content area nonfiction selections. Students are asked to make literal interpretations as well as inferential generalizations about the passages and make determinations of the most appropriate answer given four choices.

Technical adequacy of the ITBS is strong. Reliability coefficients for each of the subtests range from .85-.92 as determined by the Kuder-Richardson Formula 20. The reading comprehension reliability coefficient (re-normed for Fall and Spring, 2003) for Levels 9 (grade three), 10 (grade four), and 11 (grade five), are between .86 and .89 with a standard error of measurement of approximately 2.3, depending on the level (Hoover, Dunbar, Frisbie, Oberley, Bray, Naylor, Lewis, Ordman, & Qualls, 2003). There has been some concern that the ceiling level of the ITBS is too low and may not be an adequate assessment for gifted students (French, 2005) unless out-of-grade level testing occurs. However, since this study includes gifted and nongifted students, the regular grade level assessments were used.

#### *Critical Thinking Measure*

The Test of Critical Thinking (TCT) was used to assess the critical thinking ability of students in grades 3 through 5. According to the test authors, the theoretical framework of the assessment relied on Paul's Reasoning Model and The Delphi Report both of which outlined core skills for thinking (Center for Gifted Education, n.d.). The operational definition of thinking provided by the test authors is "the

process of making reasoned judgments or inferences about issues or problems based on evidence available, with recognition of the influence of point of view, assumptions, and context” (Center for Gifted Education, n.d., p. 6).

The test includes 10 scenarios and 3-6 corresponding multiple choice test items about the scenario. Students are allotted 45 minutes to complete the assessment. After 45 minutes, the students must stop the test, even if the test is not completed. The same test for grades 3-5 is used for both the pre and post assessment. The authors suggest that the test has an adequate floor and ceiling ( $\pm 2z$ ), making it appropriate for lower functioning third graders and gifted fifth graders (Center for Gifted Education, n.d.).

Technical adequacy is strong based on pilot results of the assessment. Internal consistency reliability coefficients are reported at .89 for the total sample and between .83 and .88 by grade level and gender. Content dependence was measured by asking individuals to randomly select answers without reading the scenarios. In the final version of the TCT content dependence was deemed sufficient by the authors. Age/grade progression were also measured by determining the mean scores by grade level to test the assumption that third graders would have the lowest mean score and fifth graders the highest (Center for Gifted Education, n.d.).

#### *Performance-based Assessment*

Performance-based assessments have been used successfully to assess literary analysis and higher-level thinking skills in similar, but larger-scaled studies (see VanTassel-Baska, Bracken, Feng, & Brown, in press; VanTassel-Baska, Zuo, Avery, & Little, 2002). Likewise, Gallagher (2006) suggests that performance-based

measures are more likely to show a true curriculum effect than standardized measures due to curriculum match. Other researchers postulate that performance-based measures are better indicators of true performance, specifically for students of poverty (Callahan, 2005; VanTassel-Baska, Johnson, & Avery, 2002).

The performance-based assessment (PBA) in this study is aligned with the higher-level tasks and readings incorporated into the *Jacob's Ladder Reading Comprehension Program*. Six separate assessments (one pre-assessment and one post-assessment per grade level) were created. Each consists of a reading selection and four open-ended questions based on the higher level thinking required for each of the ladders: generalizations/ concepts, creative synthesis, inference, and implications/ consequences. The assessments and rubric are included in Appendix A. The format of each of the pre/post assessments by grade level is similar with questions modified slightly to match the given reading. A single rubric was also designed to measure the responses of all assessments, patterned after a rubric used for similar purposes in earlier studies (VanTassel-Baska, Bracken, Feng, & Brown, in press; VanTassel-Baska, Zuo, Avery, & Little, 2002). The rubric scale is 0 to 4, with 0 being inappropriate or no response and 4 exemplifying an insightful, detailed response with evidence from the text.

Content validity for the pre and post assessment was determined based on qualitative reviews from three content experts with terminal degrees in gifted education and published studies or research experience in reading, curriculum, and/or assessment. Each reviewer was asked to provide feedback regarding the following: 1) readability of the reading selections for each respective grade level, 2) question match

to the higher level thinking skill intended, 3) the content and clarity of the rubric, and 4) appropriateness of the relevant reading selections. Reviewer comments were positive regarding the pre and post assessments by grade level with the exception of the implications and consequences questions and some clarification to the rubric wording. Recommended changes from the reviewers were incorporated and evidenced in the current version of the assessment.

The pre and post assessments were piloted in a local school district in Virginia to assess the equivalence of the pretest and posttest forms within and across grade levels. Three classrooms participated in the pilot (N=51), one per grade level (3<sup>rd</sup> n=15; 4<sup>th</sup> n=17; 5<sup>th</sup> n=19). The sample included gifted and regular education students in Title I schools who had no exposure to the curriculum. Teachers were instructed to administer the pre and post measures to their class on the same day or consecutive days, depending on the amount of time allotted for reading instruction. The assumption of this procedure is that since students have not been exposed to teaching of the curriculum between assessments, student pre and post assessment responses should be similar if the forms are equivalent (Creswell, 2002).

Reliability coefficients between the pretest/posttest measures were assessed using Cronbach's Alpha to determine the amount of shared variance and equivalence between the pretest and the posttest forms. Although there is debate regarding an acceptable level of reliability coefficients, the level of acceptability is usually dependent upon the intended use of the assessment (Pedhazer & Pedhazer-Schmelkin, 1991). In exploratory research, such as this study, a reliability coefficient of .70, or lower, could be considered acceptable as long as the use of the assessment is for

exploratory purposes that are not highly consequential in outcome (Pedhazer & Pedhazer-Schmelkin, 1991). Assessments with higher consequential validity (e.g., special education placement) would require higher coefficients ( $>.9$ ) (Pedhazer & Pedhazer-Schmelkin, 1991).

This PBA pilot study yielded reliability coefficients above .70 for both fourth grade ( $\alpha=.76$ ) and fifth grade ( $\alpha=.73$ ). The third grade results yielded an unacceptable negative coefficient ( $\alpha=-.18$ ). After another content expert review of the third grade assessment was conducted with a different assessor, the third posttest story and questions were rewritten and the third grade assessment was repiloted in a separate school. An alpha coefficient for third grade of  $\alpha=.69$  resulted. The overall results for all grade levels yielded an acceptable level of  $\alpha=.76$ .

Reviewers from the Center for Gifted Education were secured and trained to score the assessments individually using the designed rubric. Inter-rater reliability was measured by comparing the consistency between each rater's student score and conducting a two-way mixed intraclass correlation alpha coefficient (Grimm & Yarnold, 1995). The overall inter-rater reliability was ( $\alpha=.81$ ) meaning that the raters were consistent in their scoring approximately 81% of the time. Grade level inter-rater coefficients were also analyzed. Fourth grade calculations revealed the highest rater consistency ( $\alpha=.83$ ), followed by third grade ( $\alpha=.75$ ), and then fifth ( $\alpha=.70$ ).

#### *Jacob's Ladder Treatment Fidelity Form*

The *Jacob's Ladder Treatment Fidelity Form* (French, 2005) outlines nine key components of the *Jacob's Ladder* curriculum that could be observed in an experimental classroom, based on the processes inherent to the *Jacob's Ladder*

curriculum. It was not expected that all nine key processes would be observed during one classroom setting. The *Treatment Fidelity Form* is intended to note behaviors and processes listed in the curriculum as “observed” or “not observed”.

### *Focus Group Protocol*

An interview protocol consisting of four open-ended questions was used to measure perceptions of the *Jacob’s Ladder* curriculum among experimental teachers and their students. The following four questions were asked of each of the participants: 1) What are the strengths of the *Jacob’s Ladder Reading Comprehension Program*? 2) What are the weaknesses or barriers to the implementation of the *Jacob’s Ladder Reading Comprehension Program*? 3) What evidence do you have that the curriculum was effective or ineffective? 4) What changes would you recommend to the curriculum or the implementation process? Student questions were modified slightly for clarity with the population. For example, students were asked what they liked best, what they liked least, what they learned and how they knew they were learning, and what changes they would make.

### Treatment Curriculum

#### *Premise of Jacob’s Ladder*

The *Jacob’s Ladder Reading Comprehension Program* was developed by the Center for Gifted Education at the College of William and Mary as part of a federally funded grant. The curriculum targets third, fourth, and fifth grade students in Title I schools. The curriculum is intended to be supplemental and was designed to enhance reading comprehension skills, build reading skills from lower level thinking to higher level thinking, and to enhance student discussion of the meaning about text

(VanTassel-Baska, Stambaugh, & French, 2004). The curriculum design mimics that of climbing a ladder. Students move from lower rungs of the ladder, or lower level thinking skills, to higher level thinking skills at the upper rung of the ladder. There are four ladders, each with specific objectives and related lower level tasks that prepare students for the higher-level tasks at the upper rung of the ladder. Ladder examples are included in Appendix C.

Ladder A is designed to help students judge relationships among data provided. The highest level of the ladder, assuming the most complex task, is for students to determine implications and consequences from text or application in the real world. Lower levels of the ladder include sequencing events and analyzing cause and effect relationships.

The conceptual framework for Ladder B is the Taba Model for conceptual thinking (Taba, 1962). Students develop deductive reasoning skills by brainstorming or providing details, examples, or illustrations from the written text and then classify the details, story, or text into distinct categories. Students then make generalizations about the story or text based on the established categories from the lower level ladder rungs.

The focus of Ladder C is on literary elements. The lower level thinking skills, on the lower ladder rungs, focus on character, plot, setting, and literary device identification. The middle rung of the ladder focuses on making inferences based on the data from the reading passage and the literary elements practice on the lowest rung. Finally, students identify the main ideas or theme from the passage based on the evidence and inferences made.

Ladder D is found in the fourth and fifth grade curriculum and focuses on synthesizing information from text. Students summarize events from the story or passage at the lowest rung of the ladder, recalling events from the story and retelling in their own words. The second rung of the ladder is paraphrasing. Students synthesize larger passages and report the main ideas based on their interpretation. The highest level of complexity is creative synthesis. Students use knowledge gained from the reading passage and create something new based on their understanding.

Three reading genres are included in *Jacob's Ladder*, but differ slightly by grade level. The third grade curriculum includes myths and fables, poetry, and nonfiction. The fourth grade curriculum includes short stories and essays, poetry, and nonfiction. The fifth grade curriculum includes essays or short stories – mostly in the form of primary documents, poems, and nonfiction selections. Ten passages of each genre, for a total of thirty passages, are incorporated for each grade level. At least two ladders for each reading passage genre are included in the curriculum for each grade level, with the exception of third grade poetry, which has one ladder per poem.

#### *Jacob's Ladder Alignment to Bloom's Taxonomy*

*Jacob's Ladder Reading Comprehension Program* aligns with the conceptual framework of Bloom's Taxonomy. Table 7 shows the alignment of Bloom's Taxonomy to each of the ladder rungs. As noted, the lower level of Bloom's Taxonomy (i.e. remembering, understanding, and applying) aligns with the two lower rungs of *Jacob's Ladder*. The more complex levels of Bloom's Taxonomy (i.e. analyzing, evaluating, and creating) align with the highest ladder rung of *Jacob's Ladder*.

Table 7: Alignment of Jacob's Ladder and Bloom's Taxonomy

<b>Ladder</b>	<b>Jacob's Ladder Skills</b>	<b>Bloom's Taxonomy</b>
Ladder A	Implications and Consequences	Analyze
	Cause and Effect	Understand and Apply
	Sequencing	Remember
Ladder B	Generalizations	Synthesize
	Categories	Understand and Analyze
	Details	Remember
Ladder C	Main Idea or Theme	Synthesize
	Inference	Analyze
	Identification of Literary Elements	Remember and Understand
Ladder D	Creative Synthesis	Create
	Summarize	Understand and Apply
	Paraphrase	Understand and Apply

The cognitive processes of Bloom's Taxonomy provide a one-dimensional framework. Anderson and Krathwohl (2001) added a new, second dimension of the types of knowledge required of students. These types of knowledge are: factual, conceptual, procedural, and metacognitive. The knowledge types are on a continuum or hierarchy of Bloom's taxonomy. Each of the knowledge dimensions are also incorporated within the *Jacob's Ladder* framework as process skills.

Factual knowledge includes the basic elements students must know within a discipline including terminology, specific details, and reliable or justifiable sources of information. Within the framework of *Jacob's Ladder* students must use appropriate literary terminology and justify their answers using specific details and sources from the text.

Conceptual knowledge, defined as interrelationships among elements, is gleaned by the cumulative *Jacob's Ladder* program, as each ladder focus is on a different higher level thinking skill and requires students to understand the interrelationship among the lower level to higher level ladders and literary elements.

Ladder B, however, provides the best illustration with the conceptual knowledge component of generalizing and classifying as outlined by Anderson and Krathwohl (2001).

The procedural knowledge component of the taxonomy includes criteria for using skills, how to do something, and methods of inquiry. Many of the nonfiction questions as well as the second rung of Ladder D and Ladder A, determining cause and effect, require students to apply their procedural knowledge of a given passage to make inferences or determine the author's purpose.

Finally, the metacognitive knowledge dimension of the revised Bloom's Taxonomy includes awareness of one's own knowledge, knowledge about others, and strategic knowledge about how one thinks. The self-assessments, collaborative emphasis, and specific ladder questions that require students to reflect on oneself or response as a result of the passage are all embedded components of the overall *Jacob's Ladder* curriculum.

#### Comparison Curriculum

The curriculum used by the comparison teachers in each of the districts varied by teacher and grade level and could not be quantified. Each of the district principals reported that the district purchased reading anthologies from known publishers (e.g., Harcourt Brace and McGraw Hill), but teachers were not required to use the curricula in its entirety. Many teachers supplemented the reading text with novel studies, guided reading, and literature circles. Both districts required teachers to have training in assessing the reading levels of students. Teachers were then encouraged to select reading materials based on students' reading levels. Principals also reiterated that

teachers were required to teach to the Ohio standards and many times one curriculum will not meet all standards; therefore teachers were free to select other materials as appropriate. A standard curricula may only be required if the teacher is not meeting state goals.

Since the comparison curriculum is based on the Ohio Reading Content Standards, relevant standards were analyzed. An abridged list of the reading content standards by grade level cluster is included in Appendix D. The Ohio curriculum standards are divided into six reading strands including: phonemic awareness, vocabulary, reading, writing, research, and communication (Ohio Department of Education, 2001). Each strand is further divided by indicators and subsequent benchmarks. Indicators are specific objectives outlined by grade level, whereas benchmarks are overarching goals that include multiple grade level clusters (e.g., K-4, 5-7, 8-10, 11 & 12). Indicators are the mechanism for which benchmarks are met. For example, a specific writing indicator at the third grade level is “Write responses to novels, stories, and poems that demonstrate an understanding of the text and support judgments with specific references to the text” (p. 22). The third through fourth grade commensurate benchmark is: “Write responses to literature that summarize main ideas and significant details and support interpretations with references to the text” (p. 22).

### Study Procedures

#### *Study Timeline*

The pre and post assessments were administered approximately three months apart to prevent students from learning the test and to avoid pre-post test sensitization (Gall, Borg, & Gall, 1996). Within that timeframe it is unlikely that students would

remember test items such that it would become a threat to the validity of the study. There was some discrepancy in the actual amount of time of the intervention due to Thanksgiving holiday breaks and District Two conflicts that required that post test occur one week earlier than anticipated due to previously scheduled state test practices.

Administrators and experimental teachers were given a timeline of the study that included tasks, testing dates, implementation dates, and researcher visits. The innovation lasted between 10-12 weeks. A copy of the timeline is outlined in Table 8. Specific procedures listed on the timeline including teacher training and assessment administration are explained in the next section.

### *Teacher Training*

Prior to the curriculum implementation, experimental teachers in each district were trained on the use of the *Jacob's Ladder Reading Comprehension Program*. Training was one day in duration and lasted approximately six and a half hours. In District One all experimental teachers (n=12) attended the training in addition to the principals and assistant principals for each building (n=4). The superintendent of schools was also in attendance. In District Two the curriculum director, teacher of gifted, coordinator of gifted, and experimental teachers (n=10) attended the training. Each district hired substitute teachers so that training could occur during the school day. Ongoing professional development occurred through classroom observations and feedback by the researcher, e-mail groups, and modeling of lessons by the teacher of gifted in District 2. In addition, a Blackboard site was established for ongoing teacher

*Table 8: Timeline of Tasks for Study Implementation*

Check	Task	Dates/Deadlines	Person Responsible
	Submit permission slips to designee	Friday, Sept. 8	Comparison & Experimental Teachers; School Leader
	Administer the ITBS, TCT, & CBA (comparison and experimental) PBA: Experimental Only	Week of Sept. 11	Comparison & Experimental Teachers (with assistance of school designee)
	Score PBA according to the rubric	Week of Sept. 11	Experimental Teachers
	Submit all completed answer sheets to designee	Friday, Sept. 15	Experimental & Comparison Teachers and School Leader
	Provide list of identified gifted students & areas of id (e.g., SC and reading) for comparison & experimental classes	Week of Oct. 9-18 (or before, if possible)	School Leader or designee
	Provide list of comparison and experimental students on free/reduced lunch		
	Visit Classrooms and Collect Answer Sheets	District One: week of October 9 (10-13) District Two: week of Oct. 16 (18 <sup>th</sup> or 19 <sup>th</sup> )	School Leader & Researcher
	Begin curriculum implementation: 2 readings and ladders per week Complete curriculum by Dec. 1 Visit Classrooms	Sept. 18 – December 1	Experimental Teachers
	Administer the ITBS, TCT, & PBA (exp. only) (comparison & experimental)	Week of December 4-8	School Leader & Researcher Experimental and Comparison Teachers (school designee to remind everyone)
	Score PBA according to rubric	Friday, Dec. 8	Experimental Teachers
	Submit all completed answer sheets/test booklets to designee	Friday, Dec. 8	Experimental and Comparison Teachers
	<ul style="list-style-type: none"> <li>• Focus group arrangements and feedback (teachers and random group of students)</li> <li>• Collect Answer Sheets and Test Booklets</li> </ul>	Week of Dec. 11	Researcher and School Leader
	Train additional teachers as desired	TBD after January, 2007	Researcher and School Leader

discussion and the sharing of ideas. However, teachers preferred the group e-mail approach instead.

The same PowerPoint presentation and handouts were used in each district to ensure continuity in training. The training for experimental teachers and administrators consisted of the following topics: a) Background of *Jacob's Ladder* Design, b) Purpose of the Study, c) Study Timeline and Procedures, d) Reading Research, e) Modeling of Curriculum Processes by Researcher, f) Practice with the Curriculum by Grade Level, g) Assessment Procedures and Performance-Based Rubric Scoring, h) Practice Scoring Assessments with Feedback, i) Classroom Management and Grading, f) Communication mechanisms with Researcher (e.g., Blackboard, e-mail, phone).

#### *Implementation Guidelines and Process Skills Discussed in the Training*

Teachers were instructed to first administer pre-assessments, prior to curriculum implementation. Principals arranged the pre-testing in the comparison classrooms. All teachers administered the pre-assessments the second full week of September. The experimental teachers were asked to implement the *Jacob's Ladder Reading Comprehension* curriculum by first modeling a sample of each ladder of the curriculum for the entire class, moving from the lowest ladder rung to the highest. This whole-group instruction format allowed students to understand the verbiage of the various ladders and provides a necessary first step for scaffolding that builds metacognitive skills and reading comprehension (Clay & Cazden, 1992). Modeling also enforced an expectation for how students will work and conduct discussions when using the curriculum in a small group format. After each of the specific ladders (A, B,

C, & D) has been modeled and students have practiced each ladder in a whole group modeling session, teachers were asked to provide opportunities for students to work in pairs to complete the ladders and curriculum, after independently reading the selection and answering their personal ladder set, either in class or as homework.

The next step was to pair students with a partner to discuss the ladder answers and share ideas, working toward consensus on each question, using either a combination of answers between partners or coming up with a new idea. Teachers were then instructed to provide closure to the lesson with the entire group by: asking certain pairs to share, discussing the most interesting components of paired discussions, or requiring one person from each pair to write responses on chart paper. This individual, dyad, whole group method was the process teachers were to follow for implementing the remainder of the curriculum.

Teachers were also provided instruction regarding the assessment procedures throughout the implementation of *Jacob's Ladder*. Students were to independently write their answers for each corresponding ladder section on the provided answer sheet or their own paper. Then students were to complete a self-assessment and checklist for each ladder, staple all papers together, and turn them in to the teacher. The teacher completed the teacher assessment section and the record-keeping forms provided in the curriculum guide. The assessments were used for guiding reading selection sections, grouping students for the next reading selection, or planning further instruction based on trouble areas for individual students or groups.

Since the *Jacob's Ladder Reading Comprehension Program* is a supplemental program, teachers were asked to teach two reading selections and corresponding

ladders per week, spending two or three reading classes per week on *Jacob's Ladder*. The remainder of the time during each week, they continued using their "normal" reading program. Teachers were permitted to select which *Jacob's Ladder* readings were used each week, as long as a variety of genres were included during the implementation phase.

#### *Pre- Assessment*

Prior to the *Jacob's Ladder* implementation and testing, permissions for comparison and experimental students were secured using the outline and forms approved by the College of William and Mary Human Subjects Committee. Comparison and experimental teachers then administered the ITBS reading comprehension subtest and the TCT. The PBA pretest was also administered to the experimental group only. Building principals or a designee coordinated the testing implementation and procedures for the comparison teachers. All teachers participating in the study administered the assessments during the same week, the second full week of September, 2006.

Experimental teachers scored the PBA using the provided rubric as practiced during the professional development day. The TCT and the ITBS were collected for scoring by the researcher. Both the TCT and the ITBS were machine-scored. Random tests were selected for rescoring to ensure machine accuracy. The ITBS scores were accurate; however, the TCT scores were inaccurate and all TCT's were re-scored by the researcher and other assistants.

### *Treatment Fidelity*

To ensure treatment fidelity, the researcher visited each experimental classroom two times during the course of the intervention, once during the beginning of the implementation and once toward the end of the implementation. The classroom observation for each teacher lasted approximately forty-five minutes. The *Jacob's Ladder Treatment Fidelity Form* (see Appendix B) was used to record what was happening in the classroom and to ensure that the teachers were implementing the curriculum appropriately. Upon completion of the observation, each experimental teacher was given a feedback card from the research with accolades and suggestions for continued successful implementation of the curriculum.

### *Post-Assessment*

Upon conclusion of the treatment period, both comparison and experimental teachers again administered the Reading Comprehension section of the ITBS in addition to the TCT. Experimental teachers also administered and scored the posttest PBA. Principals or a designee oversaw the post assessment testing procedures for both the experimental and comparison teachers. District Two administered the post assessments one week earlier than scheduled, thus reducing the implementation period by one week. This was done because the scheduled week for the *Jacob's Ladder* post-testing was also the same week as the district-wide Ohio achievement practice testing.

One week after all assessments were administered, four focus groups were conducted by the researcher: one teacher focus group and one student focus group per district. All experimental teachers participated in the teacher focus groups by district,

and two students per experimental classroom were selected by their teacher to participate. Teachers were asked to select two students (one male and one female) who would be willing to openly discuss pros and cons of the curriculum in a focus group setting.

During the focus group, each participant was given four index cards, one per question. Question One was read aloud and then participants responded by writing on the index card. Whole group discussion then occurred as led by the researcher. Follow-up questions based on participant responses were asked for clarification purposes. The index cards for Question One were collected and the same process occurred for Questions Two through Four. Upon completion of the focus group, teachers and students had an opportunity to ask questions about the study and next steps.

### Data Analysis

The data analysis for this study will be discussed by question. For each question requiring quantitative statistical analyses, alpha was set at  $p \leq .05$ . Table 9 outlines the analyses by research question.

Question One: The pre and post test scores of the ITBS for both the comparison and the experimental group were analyzed using an Analysis of Covariance (ANCOVA) to control for initial mean pretest differences between comparison and experimental groups (Weinfurt, 1995). Within-subjects analysis were also analyzed for both the comparison and the experimental groups, as measured by a paired-samples t-test, to determine the mean differences between the pretest and posttest for each group (Kiess, 2002).

Question Two: The TCT pretest and posttest scores were analyzed using an ANCOVA to determine mean differences between the comparison and experimental group. The PBA was analyzed by conducting a paired samples t-test to compare growth between pre and posttest means within the experimental group.

*Table 9: Statistical analysis by research question and corresponding data*

<b>Question</b>	<b>Instrumentation</b>	<b>Analysis</b>
What differential effects does the <i>Jacob's Ladder Reading Comprehension Program</i> have on students' reading comprehension?	ITBS pretest and posttest scores converted to IQ measures to make comparisons across grade levels	ANCOVA Paired samples t-test
What differential effects does the <i>Jacob's Ladder Reading Comprehension Program</i> have on students' critical thinking and higher level thinking skills in literature?	TCT pretest and posttest raw scores PBA pretest and posttest total scores (experimental group only)	ANCOVA Paired samples t-test
Are there differential effects of the <i>Jacob's Ladder Reading Comprehension Program</i> on critical thinking and reading comprehension by gender, poverty level, gifted status, and grade level?	TCT pretest and posttest raw scores ITBS pretest and posttest scores converted to IQ measures	2 x 2 x 2 x 2 x 3 MANCOVA (gender, gifted, SES, condition, and grade)
What are the perceptions of teachers and students regarding the use of the <i>Jacob's Ladder Reading Comprehension Program</i> after implementation?	Teacher and student written and verbal comments by question	Holistic coding for emergent themes/patterns

Question Three: A 2 (gender) x 2 (socio-economic status) x 2 (condition: experimental or comparison) x 2 (gifted) x 3 (grade level) Multivariate Analysis of

Covariance (MANCOVA) assessed the differential effects of *Jacob's Ladder* by factor for both of the standardized, dependent variables: the TCT and the ITBS. The MANCOVA was selected to protect from familywise error of the use of more than one dependent variable (Weinfurt, 1995).

Question Four: An open-ended interview protocol was used for each focus group (Patton, 2002). Follow-up questions were asked on a variable basis and determined by participant responses. Responses for each group were holistically coded and emergent themes across all groups were determined (Patton, 2002).

### Limitations and Delimitations

#### *Limitations*

Limitations of this study included the possible lack of generalizability to certain school districts whose demographics may vary, possible diffusion of treatment, teacher selection, and PBA technical adequacy.

The sample was drawn from small rural Title I school districts of less than 3,000 students. The generalizability is hence, limited to rural or small school districts of similar demographics and geography. Likewise, many of the rural districts in Ohio may not approximate the total population of rural districts since the overwhelming majority of students in this sample are White, non-Hispanic.

Treatment fidelity or the level of assurance that the experimental curriculum was actually being taught as intended poses an ongoing threat (Gall, Borg, & Gall, 1996). Researchers, as guests in schools, do not have control over how the curriculum will be implemented. This is even more evident when conducting research in a different state. In order to ensure as much treatment fidelity as possible, training

sessions about *Jacob's Ladder* were conducted, communication mechanisms were established, and classroom visitations by the researcher were analyzed using the Treatment Fidelity Form twice during the study. Feedback on teacher lessons were given by the researcher, and in District 2 liaisons assisted teachers with curriculum implementation as requested. Teachers also self-reported the number of ladders and readings they implemented within their classroom. Regardless of these precautions, fidelity could still be an issue as the researcher had to rely upon teacher self-reporting and scheduled visitations to determine fidelity.

Diffusion of treatment may also be a potential limitation or threat to the validity of the study (Gall, Borg, & Gall, 1996). Since this study's focus is on rural districts, which are typically smaller, two different school buildings could not be secured (one for experimental and one for comparison). Therefore diffusion of treatment was plausible. To prevent this as much as possible, the researcher emphasized the importance of maintaining confidentiality regarding the treatment curriculum from the comparison teachers until the conclusion of the study.

Comparison teachers were also assured that they would receive the same training and materials as the experimental teachers upon conclusion of the study. This was done to further promote equity and lack of diffusion. Despite these precautions, there is little researcher control over what students may have discussed on the playground or what experimental teachers may have inadvertently discussed in a meeting or what comparison teachers may have overheard from an experimental classroom next door.

Another limitation of this study is that no observations occurred in the comparison teachers' classrooms. This poses a problem when discussing the results of

the study as teaching behaviors, practices, and curriculum could not be compared. Moreover, the monitoring of possible diffusion of treatment would be less evident since comparison classrooms were not observed.

History could have posed a threat to validity and is a limitation of this study. Students in District Two were administered state of Ohio achievement tests as well as this study's posttests within two weeks. This may have caused a lack of motivation, test anxiety, or test exhaustion among participants. Moreover, testing also occurred after the Thanksgiving Break, which could have ramifications for students' motivation, and test preparedness after a long holiday break.

A final limitation of the study is the technical adequacy and subjectivity of the performance-based assessment. Even though it may be argued that the technical adequacy is acceptable (Pedhazer & Pedhazer-Schmelkin, 1991), this type of assessment is subjective. Experimental teachers participating in the study were trained to score the assessments individually, given a provided rubric. However, since this study is the first to use the pre-post assessment as intended, no exemplars could be given to guide assessment scoring. Moreover, reliability between pretest and posttest forms is moderate (.69 to .76) and could be stronger. The assessment needs further refinement, stronger technical adequacy, and additional reliability and validity information.

### *Delimitations*

*Jacob's Ladder* has been implemented across many school districts in the United States in a variety of ways, as informally reported by teachers. For example, some teachers have used the intended third grade *Jacob's Ladder* curriculum with

extremely precocious first and second graders while other teachers have used the intended fifth grade curriculum with struggling readers in middle and high school. Likewise, the curriculum has been used in urban and suburban schools, including Title I and non-Title I schools. Even though these uses may have merit, the purpose of this study was to determine the effectiveness of the curriculum with the intended audience of third, fourth, and fifth grade gifted and non-gifted regular education students in rural, Title I schools. Therefore, other grade levels and special education students were not included as part of this study.

### Conclusion

This quasi-experimental study was designed to assess the differential effects of the *Jacob's Ladder Reading Comprehension Program* for improving reading comprehension and thinking skills for students in heterogeneous classrooms and low-income, Title I schools. Pretest and posttest scores were collected on two standardized tests and a performance-based measure to determine to what extent the *Jacob's Ladder Reading Comprehension Program* impacted students' reading comprehension and critical thinking. In addition, this study intended to determine the impact of the curriculum for subgroups of students including those on free and reduced lunch, male and female, gifted and nongifted, and by grade level. Data were also collected on students' and teachers' perceptions of the program after implementation.

The next chapter will discuss the findings of each of the research questions explained in this chapter and present specific data related to each.

## Chapter IV

### Results

#### *Introduction*

This chapter presents and summarizes findings from the study on the *Jacob's Ladder Reading Comprehension Program* and its impact on reading comprehension and critical thinking for students in grades three, four, and five. Perceptions of teachers and students were assessed qualitatively; pre/post standardized and performance-based measures were used to quantify students' academic growth. Treatment fidelity was measured by conducting two classroom observations per experimental teacher during the course of the study. The length of the study was approximately 11 weeks. This quasi-experimental study included 495 students within 22 classrooms (10 comparison, 12 experimental) across two rural Title I school districts in Ohio. Results from a third grade comparison teacher were omitted from the study due to accusations of post-test practice. Findings for this study will be addressed by research question.

#### *Findings Related to Question One*

*Research Question One: What differential effects does the Jacob's Ladder Reading Comprehension Program have on students' reading comprehension?*

This question probed the effects of the *Jacob's Ladder Reading Comprehension Program* on student reading comprehension scores as measured by the Reading Comprehension section of the survey battery of the ITBS. The ITBS reading comprehension section of the assessment was administered to comparison and

experimental classrooms prior to and after treatment. ITBS forms employed corresponded to each grade level. Approximately 12 weeks passed between pre and post testing. All schools and grades administered the pre-test the second full week in September. Post testing for District Two occurred one week earlier than District One due to mandated state practice testing previously scheduled to be administered in District 2.

Once raw scores on the ITBS were obtained, they were converted to a standard IQ metric, with a mean of 100. An independent samples *t*-test was conducted to compare comparison and experimental pretest means to determine the initial equivalence between the two groups. Means and standard deviations for both groups are listed in Table 10. The *t*-test results revealed a significant difference between the ITBS Reading Comprehension pre-test means for the comparison and experimental groups ( $t(3.87)=470, p < .001$ ).

*Table 10: Means and Standard Deviations of Student ITBS Scores*

<b>Sample</b>	<b>N</b>	<b>ITBS Pre-test Mean (SD)</b>	<b>ITBS Post-test Mean (SD)</b>
Comparison	172	101.96 (14.80)	108.11 (14.23)
Experimental	243	107.92 (14.94)	114.12 (15.04)

A one-way analysis of covariance, ANCOVA, was selected to assess the post-test mean differences between the two groups because they differed significantly on the ITBS pretest measure. The ITBS Reading Comprehension post-test was used as the dependent variable, and the ITBS pretest was the covariant. Levene's Test of Equality of Error Variance was also conducted to test the assumption of equal variances across groups. The Levene's test revealed nonsignificant results, supporting the assumption of equal variance and the use of an uncorrected ANCOVA.

With alpha set at  $p \leq .05$ , the results from the ANCOVA yielded non-significant differences on the ITBS reading comprehension subtest between the comparison and experimental groups ( $F_{(1,437)} = .081$ , ns).

Since significance between the experimental and comparison group was not ascertained, a paired-samples  $t$ -test was also conducted to determine whether there were significant pre-post gains within each group on the ITBS Reading Comprehension subtest. The results of the paired samples  $t$ -test yielded statistically positive and moderate gains for both the comparison group ( $t_{(171)} = 6.71$ ,  $p < .001$ ;  $d = .37$ ) and the experimental group ( $t_{(242)} = 9.42$ ,  $p < .001$ ;  $d = .40$ ).

Upon further examination of the dataset, the ITBS pretest revealed a ceiling effect in that several gifted students scored nearly perfect scores, leaving little room for posttest gains. The ITBS manual explains ceiling and floor effects for the various levels of the assessment and caution that “to measure high-ability students accurately, the test must have enough ceiling to allow such students to demonstrate their skills. If the test is too easy, a considerable proportion of these students will obtain perfect or near-perfect scores” (Hoover et al., 2003, p. 31). Indices of ceiling effects reported and based on the ITBS pilot data reveal that less than five percent of students obtain a perfect or perfect less one score (Hoover et al., 2003).

Frequency distributions for the ITBS pretest scores were conducted for the gifted sample, both the experimental and comparison groups. The frequency data revealed that approximately 30% of the third graders, 26% of fourth graders, and 30% of fifth graders earned scores within three points of a perfect score. Therefore, due to possible ceiling effects, a separate ANCOVA was conducted to determine the

differences between comparison and experimental groups without gifted students included. Significant gains favoring the experimental group were evident ( $F_{(1, 355)} = 6.86, p=.009, d=.41$ ).

### *Findings for Question 2*

Research Question Two Asked: What differential effects does the *Jacob's Ladder Reading Comprehension Program* have on students' critical thinking and higher level thinking skills in literature? The intent of Question Two was to address the effects of the *Jacob's Ladder Reading Comprehension Program* on students' higher level critical thinking skills. Two measures were used to examine the effects on a students' thinking: the TCT and a PBA. The TCT was administered in both the comparison and experimental classrooms prior to and after the experimental treatment. This test is a general measure of students' ability to comprehend what is read and to think critically about the text as defined in part by the elements of Paul's Reasoning Model (Paul, 1992.).

A pre and post curriculum-based assessment (PBA) was also used to assess specific content-based learning for experimental students only. The questions in this assessment were open-ended and measured students' synthesis of reading content to create meaning, determine implications and consequences, make inferences, and create or apply generalizations after reading a short passage. Teachers were trained to score the assessment during the professional development training prior to curriculum implementation.

Both the TCT and PBA were analyzed separately because the TCT was administered to both the experimental and comparison group, and the PBA was administered to the experimental group only.

*Findings Related to the Test of Critical Thinking*

A total of 45 points can be accrued on TCT. The means and standard deviations for each group are listed in Table 11. As can be seen in the table, the pretest and posttest means are well within the 0 to 45 raw score test range and ceiling effects are not evident.

*Table 11: Means and Standard Deviations of Student TCT Scores*

<b>Sample</b>	<b>N</b>	<b>TCT Pre-test Mean (SD)</b>	<b>TCT Post-test Mean (SD)</b>
Comparison	186	13.91 (6.36)	17.22 (6.87)
Experimental	263	17.35 (7.45)	20.57 (6.61)

An independent samples t-test was conducted to analyze the pretest means of the comparison and experimental groups and to determine the equivalency between groups on the TCT. The *t*-test results revealed a statistical difference in the pre-test means between the comparison and experimental group:  $t_{(-5,12)}=447, p=.001$ ).

Based on the significantly different pre-test means, an analysis of covariance, ANCOVA, was selected to analyze the between group mean differences and to control for unequal pretest means with the TCT post-test as the dependent variable and the TCT pretest as the covariant. Levene’s Test revealed no significant differences in variance between groups, supporting the use of the ANCOVA analysis. With an alpha level set at  $p \leq .05$ , the results from the ANCOVA yielded significant mean

differences on the TCT between the comparison and the experimental group, favoring the experimental group ( $F_{(1,421)}=4.21$ ,  $p=.019$ ,  $d=.53$ ).

*Findings Related to the Performance-Based Assessment*

The PBA was administered to the experimental group only, pre and post treatment. For this descriptive analysis, a paired-samples t-test was conducted to examine pretest and posttest means on the PBA to assess the effects of the *Jacob's Ladder Reading Comprehension Program* on students' ability to apply higher level thinking skills to a reading passage after experiencing the curriculum. The paired-samples t-test was analyzed for the entire experimental group and also by grade level. Results of the analysis are listed in Table 12. The maximum score possible on the performance-based assessment was 16.

*Table 12: Paired-Samples t-Test Results for PBA by Condition*

	<i>N</i>	<b>Pretest M(SD)</b>	<b>Posttest M(SD)</b>	<i>t</i>	<i>d</i>
Total Group	242	6.54 (3.52)	10.05 (3.85)	19.16**	1.00
Third Grade	84	4.46 (2.04)	8.32 (2.60)	15.64**	1.89
Fourth Grade	87	6.75 (3.42)	10.76 (4.32)	12.39**	1.17
Fifth Grade	71	8.75 (3.65)	11.24 (3.80)	6.92*	.68

\* $p<.05$     \*\* $p<.001$

The PBA yielded significant results for the total group and for each grade level. The effect size of the intervention based on this assessment was very large for the total group and by grade level ( $d=1.00 - 1.89$ ) with the exception of fifth grade which showed positive moderate gains ( $d=.68$ ) (Cohen, 1988).

Since the PBA assessment scoring was subjective, two assessments per classroom posttest (approximately 10% of the sample) were randomly selected, scored

and analyzed by the researcher to determine the level of agreement between the researcher and the teachers' scores and fidelity of scoring. The intraclass alpha coefficient showed strong agreement between the teachers and the researcher on the assessments ( $\alpha=.89$ ).

### *Findings Related to Question 3*

Question Three: Are there differential effects of the *Jacob's Ladder Reading Comprehension Program* on critical thinking and reading comprehension by gender, poverty level, gifted status, and grade level? This question addresses the extent to which the *Jacob's Ladder Reading Comprehension Program* posttest means on the ITBS and the TCT differ between experimental and comparison groups by gender, socio-economic status, gifted status, and grade level. Table 13 shows the number and percentage of participants for each factor by condition. To determine the differential effects, a 2 (gender) x 2 (socio-economic status) x 2 (gifted) x 3 (grade) MANCOVA was conducted.

*Table 13: Number and Percent of Participants by Factor and Condition*

Condition	Free/Reduced		Gender		Gifted		Grade		
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
	Yes	No	Male	Female	Yes	No	3	4	5
Comparison	53 (27.6)	132 (62.3)	103 (52.6)	94 (47.4)	11 (5.6)	175 (89.3)	60 (30.6)	75 (33.8)	61 (27.5)
Experimental	72 (26.4)	196 (71.8)	133 (48.7)	140 (51.3)	44 (16.1)	226 (82.8)	91 (33.3)	97 (35.5)	85 (31.1)
Total	125	453	236	234	55	401	151	172	146

*Note: Not all percents will equal 100 due to unreported data*

The Levene's Test of Equal Variance was also conducted as part of the analysis to test the assumption of unequal variances between groups. The test yielded no significant results, thus supporting the use of MANCOVA. A model was designed prior to conducting the MANCOVA by covarying the TCT and ITBS pretest and then

adding main effects for condition, free and reduced lunch, gifted, and grade.

Interaction effects were then included for condition and each of the listed factors.

The MANCOVA yielded significant main effects using Wilk's Lambda for two factors: gifted ( $\Lambda=4.5$ ,  $p=.012$ ) and grade ( $\Lambda=13.99$ ,  $p<.001$ ). An interaction between condition and grade was also detected ( $\Lambda=3.68$ ,  $p=.006$ ). No significant gender or free and reduced lunch effects were revealed.

Univariate follow-up statistics assessed mean differences between the control and experimental group by dependent variable (i.e., TCT and ITBS) for the main effect of gifted and the interaction between condition and grade. The main effect of grade was not examined due to the interaction with grade and condition. Significant differences were revealed for gifted on the TCT ( $F_{(1, 386)}=9.00$ ,  $p<.01$ ) and significant interactions by grade and condition were evident on the ITBS ( $F_{(2, 386)}=7.38$ ,  $p=.001$ ).

An independent samples t-test was conducted to examine the significant main effect of gifted on the TCT. The t-test results revealed that the gifted group significantly outperformed the nongifted group on the TCT ( $t_{(429)}=9.78$ ,  $p<.001$ ).

Follow-up analyses on the significant grade level by condition interaction were also conducted to further examine the effects of the ITBS. A one-way ANOVA revealed significant differences with moderate effect sizes between the experimental and comparison groups for both third grade ( $F_{(1, 137)}=13.87$ ,  $p<.001$ ,  $d=.59$ ) and fourth grade ( $F_{(1, 160)}=9.62$ ,  $p=.002$ ,  $d=.54$ ), favoring the experimental groups. Fifth grade students' results between the experimental and control group means on the ITBS were nonsignificant ( $F_{(1, 119)}=.16$ ,  $p=.694$ ).

Because the ITBS was one of the dependent variables used for the MANCOVA analysis, ceiling effect for gifted students was again a concern as explained in Question One. Therefore, a separate MANCOVA was conducted excluding gifted students' test scores on the ITBS. The MANCOVA revealed similar findings regarding significance when the gifted group was excluded (i.e., significant grade main effect and significant interaction by grade and condition on the ITBS). Follow-up analyses were conducted to test for mean differences on the ITBS by grade level and condition. The one-way ANOVA analysis again yielded significant mean differences between comparison and experimental groups, favoring the experimental group for both third ( $F_{(1,131)}=13.85, p<.001$ ) and fourth grade ( $F_{(1,141)}=7.00, p=.009$ ). For fifth grade the mean difference results on the ITBS between the comparison and experimental groups were still nonsignificant ( $F_{(1,100)}=2.95, p=.09$ ) when the gifted group scores were excluded.

#### *Treatment Fidelity*

The *Jacob's Ladder* Treatment Fidelity form was used to assess key observed behaviors in each of the experimental classrooms associated with the treatment curriculum. Two observations were conducted during the course of the study. The fidelity form included a rating that indicated if the specific teaching behaviors were "observed" or "not observed" based on nine different criteria. It is not expected that all nine observed behaviors would be viewed during one observation. The observed behaviors listed on the Treatment Fidelity Form are inclusive of any process or behavior that could be observed during use of the curriculum. Table 14 shows the

frequency of behaviors observed across all experimental teachers' classrooms for each scheduled observation.

*Table 14: Treatment Fidelity Form: Frequency of Observed Behaviors*

<b>Behavior</b>	<b>Observation One</b>	<b>Observation Two</b>	<b>Total</b>
Students complete answers individually	5	5	10
Students are grouped in dyads for discussion	6	8	14
Students complete self evaluation	0	0	0
Students complete record sheets	0	0	0
Teacher differentiates reading selections based on student needs	3	3	6
Teacher provides student feedback	10	9	19
Students are completing readings from each genre	6	7	13
Students are discussing literature as a whole group	9	9	18
Student and teacher are conferring on readings	9	9	18

Based on these observation data, experimental teachers used the core processes inherent to the curriculum, with the exception of the assessment procedures. They were consistent using the key processes of soliciting discussion, providing feedback, asking open-ended questions, reminding students to turn back in the text for evidence to support an answer, and providing a combination of individual, small, and whole group discussion opportunities. Teachers devised their own assessment measures, not

provided in *Jacob's Ladder*, to monitor student progress and therefore, the assessment procedures used in the *Jacob's Ladder* curriculum were not observed.

#### *Findings Related to Question 4*

Question Four: What are the perceptions of teachers and students regarding the use of the *Jacob's Ladder Reading Comprehension Program* after implementation?

#### *Teacher Focus Group Findings*

Two experimental teacher focus groups were held, one in each district (District One n=6; District Two n=4). All experimental teachers participated.

Before beginning each focus group, teachers were asked to report the number of *Jacob's Ladder* lessons they taught by genre. Table 15 conveys the number of stories within the curriculum each teacher self-reported as implemented. Teachers most frequently implemented poetry, followed closely by short stories/essays/fables. Nonfiction was the least used across groups. Moreover, teachers' use of the curriculum varied in implementation of the full curriculum by 33 to 80 percent.

Teachers in each focus group were asked a series of four questions. Question One focused on the strengths of the *Jacob's Ladder Reading Comprehension Program*. Both District One and District Two teachers noted the variety of genres and readings, the level of questioning and higher level thinking solicited in students, the "rich" discussion elicited, the amount of discussion elicited, and use of higher-level

Table15: Number and Percentage of Readings by Genre and Teacher

Class	Grade	District	Short Stories (n=10)	Poetry (n=10)	Nonfiction (n=10)	Percentage of Readings by Teacher (n=30)
Teacher 1	3	1	7	8	1	53
Teacher 2	3	1	8	6	0	47
Teacher 3	4	1	10	8	2	67
Teacher 4	4	1	10	10	2	73
Teacher 5	5	1	9	10	5	80
Teacher 6	5	1	7	8	3	60
Teacher 7	3	2	5	5	2	40
Teacher 8	3	2	5	5	2	40
Teacher 9	4	2	4	6	0	33
Teacher 10	5	2	4	5	1	33
Percent by Genre			69	71	18	

vocabulary that transferred to other subject areas. One teacher wrote that Jacob’s Ladder “allowed for individual expression, higher level, thinking, and processing of a variety of reading skills.” Another suggested that “I saw my students’ way of thinking change over time. The students in my class answer questions in a whole new way, often times making sure they have evidence to back it up.” Another teacher mentioned that she liked the short readings that “lend to deepened understanding.” *Jacob’s Ladder* made it easier for students to return to the text to find evidence for their

answers. “Students were really thinking about what they read and re-reading. They really enjoyed the discussions and sharing their ideas in groups. Their ideas were theirs and not the teachers.” *Jacob’s Ladder* “forces you to teach differently”.

Question Two focused on the weaknesses and challenges of the curriculum itself and the challenges of implementation of the curriculum. Each teacher mentioned the issue of time as the major barrier. When questioned what they meant by time teachers responded in different ways. Some felt frustrated trying to “cover” two ladders per week during the experimental period, while others felt they did not have enough class time allotted to Language Arts to effectively implement the ladders. As one teacher wrote, “I would have loved to implement this over a more lengthy time span. Many times I felt rushed to do a ladder, than to have used it when needed for my schedule. Sometimes it felt like a chore.” Another teacher said “Time was a factor. I allowed only for open-full discussion and very little writing. My time was limited to 45 minutes.” Another teacher reiterated the challenge by saying that “My biggest issue about *Jacob’s Ladder* was the burn out that my kids experienced. Doing so many [ladders] per week caused them to become less enthused with the program.”

Teachers also struggled with how to “fit” *Jacob’s Ladder* into an already packed curriculum. One teacher wrote, “It was difficult to decide what had to give when time was short.” Another goes on to say “It was hard to teach *Jacob’s Ladder* and the basal reader.”

Another factor related to time was in regards to the nonfiction selections, which were reported as being too long and time consuming. One teacher wrote “Nonfiction took us a long time to complete. This was in part due to the fact that I was

looking for quality instead of quantity. However, by the end of the second ladder they had lost interest and quality was diminishing.”

Although time was the main weakness discussed, a few teachers also noted the difficulty of the ladders. For example, as one teacher explained “for me, the selections were way over my students. We did not have a lot of students able to work in smaller groups as few understood the content without teacher leading.” A different teacher agreed that “It [*Jacob’s Ladder*] was very difficult for students who are struggling readers.” While another teacher mentioned that the “vocabulary in the questions took a lot of explanation and many times I found it difficult to word definitions in ‘third grade language’ – ex: implications.”

One teacher also mentioned that her students were getting burned out on the curriculum because of too much writing. She explained that some reading passages included two ladders that were “heavy in writing” while other passages and commensurate ladders did not require as much writing.

Question Three probed the evidence teachers observed that suggested the curriculum was either effective or ineffective. All teachers said the curriculum was effective with one caveat: “this curriculum is not appropriate for struggling students,” especially at the fifth grade level as one teacher explained “a few students were frustrated at the higher level of reading required.” Another suggested that “some lower students couldn’t read well and shut down at times.”

Evidence of effectiveness reported by teachers included both positive changes themselves and positive changes in their students. Although no teacher wrote about this on the provided index card, during the whole group sharing of this question, one

teacher said that her teaching improved during the curriculum implementation time. She began asking higher-level questions and re-evaluating how she taught reading. Other teachers agreed and added to the first teacher's sentiments. One teacher said she started wondering how often she expects one right answer instead of allowing different responses. She said she became more cognizant of the types of questions she asked and is working to make her questions more open-ended. Another said "there's not going back to the old way of teaching."

Teachers also noted changes in their students' responses and behaviors after using *Jacob's Ladder*. Teachers reported that their shy or reticent students were more willing to share and that it was difficult to get conversations about reading to "come to a close". "All students were involved. I had students offering answers and joining in that maybe I hadn't expected". Another teacher agreed and said "typically quiet students were sharing ideas and thoughts. Students were challenging one another to support their answers"; while another said "I have one student who is quiet and not confident with his answers but through *Jacob's Ladder* I saw him blossom and encourage others".

Teachers also noted improvement in their students' ability to work effectively in groups, answer two-part questions with in-depth responses and justification from the text, and use advanced vocabulary introduced in *Jacob's Ladder* across content areas. One teacher summarized this by writing "[*Jacob's Ladder* is] effective. Through conversations with my students they shared how awesome, challenging, and interesting their *Jacob's Ladder* activities were. They admitted that they felt that their 'brain was stretched' and understood what 'reading between the lines meant'. Their

vocabulary has also shown growth.” Another said, “Students are much better at answering two-part questions where they must defend or prove their answer using the selection. Students backed up their answers better.” Still another teacher wrote that [*Jacob’s Ladder* is] “effective. Observing and listening to student discussion and written work show that they are using higher level thinking. When they are given time in class to discuss anything, they talk more and are more eager to share their responses.”

Students’ interests in reading, writing, and discussion of reading also increased. Teachers reported that students were more eager to talk about stories and share their responses. Students “smiled and laughed more in class” when using the curriculum and would ask when it was “*Jacob’s Ladder* time.” One teacher also noted how amazed she was when examining pre-post assessment growth in student writing as measured by the performance-based assessment.

Finally, teachers reported that parents also noticed a change in classroom instruction and their child’s behavior. One teacher said that this was the first year parents of her brighter students “were not asking for more challenging work for their children. Students weren’t bored.” Other teachers reported that parents were commending teachers because their children have not complained of being bored this school year and enjoy school and reading more.

The fourth and final question focused on key changes that teachers would recommend regarding the curriculum and implementation processes. There were few common responses to this question. Among those discussed, teachers asked for more specific direction on correct versus incorrect answers, a better mix of long versus short

writing assignments within one reading selection, a teacher glossary of terms used in the curriculum with suggestions on how to best explain the vocabulary to students, more poetry, shorter nonfiction pieces, differentiated reading levels, and master copies of graphic organizers, charts, and other classroom visuals and tools that could be posted on a bulletin board in the classroom.

### *Student Focus Group Findings*

Two students per experimental classroom within each District were selected to participate in the focus group. A total of 12 students per District (N=24) participated (District One: six males, six females; six gifted students; District Two: five males, seven females; six gifted students). Students were asked the same four questions as the teachers to ascertain their perceptions regarding the strengths, weaknesses, level of challenge and learning, and recommended changes to the *Jacob's Ladder* curriculum.

Students reported three main strengths of the *Jacob's Ladder Reading Comprehension Program*: 1) interesting story selections, 2) the level of challenge, and 3) the group work. Students said the stories kept their attention and made them “dig deeper and think harder.” They cited specific passages they liked, especially poetry. One student said “I didn’t like poetry as much. Now I have my own poetry book and I write stories.” Other students reported that it was interesting to hear different answers and work in groups sharing ideas and defending their answers.

When questioned about the weaknesses of the program, or things they liked least, student answers were diverse. Many students commented that they didn’t like all of the writing involved in the program, especially when they had to write summaries. Some students explained that “the writing made my hand cramp.” Others commented

that some of the stories were “kind of dry and not as interesting as the rest”. Students also said there was inconsistency between the level of questions and stories, some being too easy and others too hard. When probed further about this, students referred to the climbing of the ladder, moving from lower level to higher-level questions in some instances, while other students thought there was a significant difference in the level of challenge between stories and ladder questions. Some believed the questions were unclear, and they were not sure what to do.

Students were also asked to discuss how they would gauge the level of challenge in the curriculum and to discuss what, if anything, they learned. As with Question Two on weaknesses, answers on this question were mixed. Some students said the curriculum was too easy, while others said it was too hard. Some said it was “just right”. One student commented “I used to think that doing easy work was fun but then I realized *Jacob’s Ladder* was harder, and harder was better. It makes my brain go far”. Many said the curriculum made them think “harder” or “outside the box”. Others noted that they realized they could do more and “it [*Jacob’s Ladder*] taught you perseverance.” Students also found the curriculum and level of challenge was “off and on” depending on the story and the questions. As one student commented “some were hard and some easy; it varied. It was a nice mix – a little hard, but a good hard.” Students also said they learned new vocabulary and more definitions from the *Jacob’s Ladder* words (i.e., generalizations, implications, personification). Others responded concretely and discussed a lesson they learned after discussing a moral or concept from a particular story. One student said she liked learning the morals to stories because “I used to cheat at games but now I don’t.”

Question 4 focused on the recommended changes for *Jacob's Ladder*. Again, student answers were mixed. Some students suggested that stories could be more interesting while others said the stories were perfect and nothing should be changed. Several students suggested creating a website where they could acquire more stories and ladders. Others recommended adding more activities and projects to the curriculum. A few students also recommended that less writing and summarizing should be included in the ladder format. In some classes, students were asked to use the ladder worksheets and suggested that the answer sheet space to write in should be larger. District Two students felt that the *Jacob's Ladder* format should include more multiple choice questions since the *Jacob's Ladder* posttests were multiple choice.

#### *Emergent Themes from All Focus Groups*

The findings from the standardized, open-ended focus group format were inductively analyzed with no predetermined themes (Patton, 2002). All themes emerged from the data based on teacher and student responses. The data were holistically coded (Patton, 2002). The following major themes emerged from the focus groups results: higher level of challenge and questioning; value-added gains and transfer; solicitation of effective conversations about literature; time; and inconsistency. Each of the themes with relevant comments from teachers and students is included in Appendix E and will be discussed separately.

*Higher level of challenge and questioning.* The majority of students and teachers discussed how *Jacob's Ladder* incorporated higher level thinking skills and was challenging to students. Teachers said that after teaching *Jacob's Ladder* they

became more cognizant of how they taught and worked to incorporate more open-ended questions and higher levels of challenge in other reading activities.

*Value-added gains and transfer.* Value-added gains refer to the skills students and teachers reported from student learning after using the curriculum. Transfer refers to the use of the skills in other learning activities that were emphasized in *Jacob's Ladder* by teachers and students and how those skills were applied to other content areas or life situations. Students and teachers both noted an increase in student vocabulary after using *Jacob's Ladder*. Participants also reported an increase in students' writing and their ability to defend their answers, using evidence from a story. Moreover, some students' reported that their interest in reading and writing also increased and *Jacob's Ladder* stories helped them learn life lessons and historical events.

*Solicitation of effective conversation about literature.* One of the key processes emphasized in the *Jacob's Ladder* curriculum is the use of literary discussion groups. Teachers found that students' ability and willingness to appropriately discuss literature was evident when using *Jacob's Ladder*. Likewise, more reticent and shy students felt comfortable discussing their ideas and talked more than normal when discussing *Jacob's Ladder* questions. Students reported that they enjoyed the group work, liked defending their answers, and enjoyed hearing their partner's ideas about the story.

*Time.* Time was an emergent theme for teachers only. Time refers to the amount of time necessary to implement *Jacob's Ladder* and the amount of time for planning and matching *Jacob's Ladder* with other units and school curriculum

requirements. Teachers needed more time to effectively implement and plan the *Jacob's Ladder* curriculum. They suggested that the curriculum could be implemented more effectively if the implementation period occurred throughout a school year when they could consistently match *Jacob's Ladder* to other standards.

*Inconsistency.* Inconsistency was mentioned by only two of teachers but was an emergent theme for students. Inconsistency refers to the unevenness of certain activities and challenge levels required for selected reading selections. Students perceived that the ladder requirements were inconsistent. Some were too easy and some were too hard. There were also too many lengthy writing assignments in some ladders associated with the same reading and very few writing activities in other ladders and commensurate readings. Students also felt that the curriculum and corresponding assessments for the project were inconsistent (i.e., multiple choice versus extended response).

### Summary of Findings

The research findings from this *Jacob's Ladder Reading Comprehension Program* study are divided into two categories: curriculum intervention effectiveness as measured by standardized and performance-based assessments, and curriculum perceptions from teacher and students who used the curriculum.

#### *Key Findings on Curriculum Intervention Effectiveness*

##### Reading Comprehension

1. After controlling for pretest mean differences, nonsignificant differences were detected on the ITBS Reading Comprehension Subtest between the experimental and comparison group.

2. A paired-samples t-test assessed the pretest/posttest mean scores both within the comparison and within the experimental group. Significant gains and moderate effect sizes for both the experimental and comparison groups were revealed.
3. Due to ceiling effects on the ITBS Reading Comprehension Subtest within the gifted sample, analyses were conducted without the gifted students' ITBS Reading Comprehension Subtest scores included. After controlling for pretest mean differences on the ITBS, results revealed significant gains and a moderate effect size, favoring the experimental group.

#### *Critical Thinking*

1. After controlling for unequal pretest differences on the TCT, significant results and a moderate effect size were revealed between the comparison and experimental means, favoring the experimental group.
2. Within the experimental group, a paired-samples t-test revealed significant gains and very high effect sizes on the PBA between pretest and posttest mean scores for the overall sample. By grade level, third and fourth grade analyses revealed significant gains and very high effect sizes. Fifth graders showed a significant gain and a moderate effect size.

#### *Analysis on Critical Thinking and Reading Comprehension by Factor*

1. A MANCOVA was conducted to determine the impact of *Jacob's Ladder* on the following factors: condition (experimental and control), gifted, SES, gender, and grade. There were statistically significant main effects for gifted students revealed on the TCT after controlling for pretest differences. Further

analyses revealed that gifted students' mean scores on the TCT were significantly higher than the nongifted students' TCT mean scores.

2. There were no significant gender, condition, or SES main effects as measured by the MANCOVA.
3. There was an interaction effect by grade and condition after controlling for pretest differences. By grade level, the experimental group of third and fourth graders significantly outperformed the comparison group on the ITBS. There were nonsignificant differences between the ITBS mean scores of the experimental and the comparison groups of fifth graders.

#### *Treatment Fidelity*

1. Experimental teachers most frequently implemented poetry, followed closely by short stories/essays/fables. Nonfiction was the least used across groups. Moreover, teachers' use of the curriculum varied in implementation of the full curriculum by 33 to 80 percent, depending on the classroom.
2. The most observed behaviors within the experimental classroom included soliciting discussion, providing feedback, asking open-ended questions, reminding students to turn back in the text for evidence to support an answer, and providing a combination of individual, small, and whole group discussion opportunities.
3. The only curriculum feature that was not used by teachers in the classroom was the assessment forms. Teachers devised different assessment measures, not provided in *Jacob's Ladder*, to monitor student progress.

### *Key Findings on Teacher and Student Perceptions of the Curriculum*

1. Focus group data revealed that teachers and students were positive about the curriculum and recommended few changes. The level of challenge varied but overall both teachers and students believed that the curriculum was challenging but appropriate, with easier and more difficult readings and questions. Teachers reported positive changes in themselves and their students as a result of using the curriculum, and students reported learning new information and increased enjoyment of discussing literature in groups.
2. Emergent themes across the focus groups included: 1) the solicitation of higher level thinking skills and challenge from the curriculum; 2) the value-added gain provided by the curriculum, including the increased use of advanced vocabulary, the transfer of learning in *Jacob's Ladder* to other content areas, and new skills and lessons learned from the readings and discussions; 3) in-depth and targeted conversations about literature; 4) time constraints in curriculum implementation within the duration of the study; and 5) inconsistency in some aspects of the program.

### *Conclusion*

The next chapter discusses the findings presented in this chapter and provides relevant connections to the literature on reading comprehension and critical thinking as related to these findings. Conclusions will be drawn based on the data presented and suggestions for implications for practice and additional research will be discussed.

## Chapter V

### Discussion, Conclusions, and Implications

#### *Introduction*

The purpose of this quasi-experimental study was to pilot the *Jacob's Ladder Reading Comprehension Program* in third, fourth, and fifth grade heterogeneous classrooms in rural, Title I schools. The effects of the curriculum for students' reading comprehension and critical thinking were measured by a pretest/posttest purposeful random quasi-experimental design. Effects were determined between and within comparison and experimental groups. Subgroup performance was also measured. Specific subgroups of interest included: gender, ability, socio-economic status, and grade level. Perceptions of teachers and students were also examined after they used *Jacob's Ladder* for at least ten weeks.

The findings related to the effects of *Jacob's Ladder* were discussed in Chapter Four. This chapter builds upon the findings and explains the conclusions that can be drawn to guide practice and future research.

#### *Discussion*

##### *Effects of Jacob's Ladder on Thinking Skills*

The analysis of critical thinking gains for students who participated in the *Jacob's Ladder Reading Comprehension Program* suggests that after controlling for pretest differences, *Jacob's Ladder* produces significantly and practically important gains in students' ability to think critically when compared to those who did not have exposure to the curriculum, as measured by the TCT. However, the MANCOVA results did not show significant differences between groups by condition in critical

thinking when adding the ITBS as an additional dependent variable and increasing the number of factors to be considered (i.e., free/reduced lunch, gender, gifted, grade). This discrepancy between the ANCOVA and MANCOVA results could be due to the number of cells required for the analysis, which reduced the sample size for each cell, and in turn reduced variance, increased error variance, decreased power, and further protected against Type I error, while possibly increasing a Type II error, when analyzing multiple dependent measures and factors (Kiehl, 2002).

Another probable explanation for the discrepant analyses between the MANCOVA and ANCOVA findings on the TCT could be the notion that general critical thinking dispositions develop over time and may not be detected within a short intervention (Abrams, 2005; Cotton, 1991). Instead, critical thinking dispositions develop over a person's life span. Studies using similar curriculum emphases and critical thinking measures found that critical thinking behaviors increase over time with added exposure to high level curriculum that emphasizes critical thinking (VanTassel-Baska & Bracken, 2005; VanTassel-Baska, Bracken, Feng, & Brown, in press). Therefore, significant growth on general critical thinking measures during a 10-12 week intervention may not be as evident, when other factors are considered.

#### *Performance-Based Assessment and Critical Thinking*

Gallagher (2006) explains the importance of performance-based assessments for determining a true measure of the impact of a curriculum since standardized assessments may not fully measure what was being taught. In addition, for students in Title I schools, which is the demographic context of this study, performance-based measures have been found to be better indicators of true performance than more

widely used standardized measures (Callahan, 2005; VanTassel-Baska, Johnson, & Avery, 2002).

The PBA findings for this study, which target specific content-based critical thinking behaviors, showed significant and important mean pretest to posttest gains within the experimental group. No comparisons between groups were made. By grade level, both the third and fourth grade groups showed extremely large practical gains, and fifth grade showed moderate gains. These findings indicate that *Jacob's Ladder* does produce significant gains in targeted and focused critical thinking skills as applied to reading comprehension within the experimental group. However, these findings should be used with caution as the PBA assessment reliability coefficients, while acceptable for a pilot study, should be higher before making consequential decisions (Pedhazur & Pedhazur-Schmelkin, 1991).

The PBA findings of this study align with other study results that also targeted thinking skills in reading. Two different studies using the College of William and Mary language arts curriculum assessed students' knowledge in literary analysis skills by administering a similar performance-based instrument. As with this study's findings, data from these curriculum intervention studies found significant and practical gains in students' growth of literary analysis and persuasive writing after using the curriculum (VanTassel-Baska, Johnson, Hughes, & Boyce, 1996; VanTassel-Baska, Zuo, Avery, & Little, 2002), that held up over three years (Feng, et al. 2004).

The *Jacob's Ladder Reading Comprehension Program* shows promise in increasing students' general critical thinking when compared to other reading

programs and supplemental curriculum based on the Ohio reading standards, but more time may be necessary to determine actual between-group effects, especially when other factors are considered. However, the use of *Jacob's Ladder* does produce significant and important results in targeted critical thinking behaviors applied to reading comprehension for those students who use the program.

#### *Effects of Jacob's Ladder on Reading Comprehension*

The initial findings as assessed by the ANCOVA suggest that *The Jacob's Ladder Reading Comprehension Program* does not produce significant gains in students' reading comprehension. However, two issues with these findings emerged when assessing reading comprehension in this study: 1) ceiling effect for gifted students, and 2) a lack of significant results between subjects for the fifth grade experimental group.

The ITBS manual explains ceiling and floor effects for the various levels of the assessment and caution that "to measure high-ability students accurately, the test must have enough ceiling to allow such students to demonstrate their skills. If the test is too easy, a considerable proportion of these students will obtain perfect or near-perfect scores" (Hoover et al., 2003, p. 31). Indices of ceiling effects are reported, based on the ITBS pilot data, and less than five percent of students obtain a perfect or perfect less one score (Hoover et al., 2003). However, pretest scores from this study sample show that within the gifted sample (comparison and experimental) approximately 30% of the third graders, 26% of fourth graders, and 30% of fifth graders scored within three points of a perfect score. Therefore, there was little room to show growth in reading comprehension on a posttest. These findings coincide with a pilot study of

*Jacob's Ladder* effects on gifted and promising learners (French, 2005). French (2005) found that positive gains were not evident for the high ability group and recommended out-of-grade-level testing if administering the ITBS. Out-of-level testing was not incorporated into this study since the majority of students were nongifted.

Once the gifted sample was removed from the larger group (both comparison and experimental), results indicated that nongifted students posited significant gains in reading comprehension when compared to the comparison group. The MANCOVA supported these results for the third and fourth grade students, but not fifth grade when other factors were considered. The third and fourth grade students showed significant and practical gains on the ITBS reading comprehension subtest when compared to the comparison group. However, fifth grade results were nonsignificant between the two groups.

There are several plausible causes for the experimental fifth graders' lack of significant results on the ITBS in relation to the comparison fifth grade group. One explanation may be the differences in the *Jacob's Ladder* curriculum at the fifth grade level. The short stories genre in the fifth grade curriculum contain primary documents as the major source of reading material whereas the third and fourth grade curriculum contain myths, fables, and short stories. Some of the primary documents within the fifth grade curriculum use unfamiliar language and may be perceived by students as uninteresting or too difficult. During the focus groups, one teacher suggested that the fifth grade readings were quite difficult for lower functioning regular education students, in particular, and more direct instruction was necessary for these students to

complete certain readings. Likewise, one fifth grade teacher in this study implemented only 30% of the curriculum because she said her students were struggling with the vocabulary and content. A previous study (French, 2005) found similar problems with the level of story selections within the fifth grade curriculum. Moreover, if fifth grade teachers did not differentiate the curriculum and the material was too difficult, it is possible that students became frustrated and “shut down” (Tomlinson, 1999; Vgotsky, 1986).

Another explanation for the lack of significant results between the fifth grade experimental and comparison groups may be that the curriculum standards at the fifth grade level for all students were more aligned with the same emphases of the *Jacob’s Ladder* curriculum and therefore, no between group differences at this grade level were detected. As explained in Chapter Three, the Ohio curriculum standards are divided into grade level clusters. Third and fourth grade are clustered together and grades 5-7 are in a separate, advanced cluster of standards. The fifth grade standards are intended to “feed” into the seventh grade standards, which means that the standards at the fifth grade level may be more aligned to the processes and components of the *Jacob’s Ladder* curriculum. Many of the reading comprehension processes may have been taught in the comparison classrooms as well. For example, the content standards at the fifth to seventh grade level include more of the higher rung emphases of *Jacob’s Ladder* such as making inferences, comparing and contrasting, determining causation, and analyzing textual features; whereas the third and fourth grade standards require students to master more basic skills such as identifying central ideas, demonstrating comprehension by responding to a variety of questions,

summarizing details in a text, and explaining how authors word choices impact the meaning of a text.

Additionally, one cannot rule out teacher effect as a plausible explanation at the fifth grade level. The importance of teacher effects cannot be underestimated (Rivers, Sanders & Horn, 1998; Rivkin, Hanushek, & Kain, 2000; Rowan, Correnti, & Miller, 2002), especially in reading (Taylor et al., 2000, 2003). One of the limitations of this study is that classroom visitations were not conducted in the comparison classrooms. Therefore, no comparisons between the experimental and the comparison teachers could be made to determine teacher effectiveness by condition. It is conceivable that the fifth grade comparison teachers were superior to the experimental group in the use of effective reading comprehension and higher level thinking strategies.

Overall, the *Jacob's Ladder Reading Comprehension Program* produced significant and practical gains in reading comprehension for third and fourth grade nongifted students. The impact of *Jacob's Ladder* on the reading comprehension of gifted students cannot be determined by this study due to ceiling effects on the ITBS within this subgroup. Alternative explanations previously discussed may account for the lack of significance between the two study groups at the fifth grade level.

#### *Perceptions of Jacob's Ladder*

The perceptions of students and teachers support the statistical findings from the standardized assessments and the performance-based assessment, suggesting that *Jacob's Ladder* produces significant growth in regular education students' thinking and reading comprehension. Teachers believed that the curriculum and instructional

processes allowed them to provide more in-depth learning and instruction through open-ended questions. They reported that their students increased their advanced vocabulary as applied to critical thinking (i.e., implications, consequences, generalizations, inferences) and they were more willing to discuss answers and provide evidence from the text.

These teacher observations align with the research on innovation and reform as transferred to classroom practice and changes in teacher beliefs (see Guskey, 1986, 2002). Teachers are more likely to use an innovation if they view it as successful and see their students “attaining higher levels of achievement, becoming more involved in instruction, or expressing greater confidence in themselves and their ability to learn” (Guskey, 1986, p.7).

Students believed the program made them think harder, use better vocabulary, and equipped them to discuss stories in depth. This transferability of vocabulary and the ongoing use of the language of a discipline could be an indicator of the development of critical thinking dispositions (Ennis, 1989; Facione, 2007; McPeck, 1990).

Although, not directly measured, the process of the curriculum implementation and systemic structures within the school building are other factors that may have impacted the curriculum implementation and effects. These include professional development, instructional and leadership support, treatment fidelity and implementation, and teacher effect.

Ongoing professional development is one of the identified best-practices of effective schools (Bellanca, 1995; Michigan State University, 2004; Muijs, Harris,

Chapman, Stoll, & Russ, 2004). For this study, the experimental teachers were provided ongoing professional development, in multiple forms, both direct and indirect. One formal professional development training session was conducted with the entire group, and two subsequent follow-up sessions within each teacher's individual classroom were conducted by classroom observations. E-mail communication with the researcher was also established so that the researcher could immediately answer questions or provide feedback. In District Two mechanisms were put into place for a teacher of gifted and a coordinator of gifted to assist the experimental teachers as needed by modeling lessons. When teachers are provided the appropriate resources, ongoing professional development training, and feedback mechanisms with coaching, monitoring and accountability (i.e., classroom observations) the implementation of an innovation is more likely to transfer to classroom practice (DuFour & Eaker, 1998; Joyce & Showers, 1988; Sparks and Hirsch, 1997). *Jacob's Ladder* may have been successful in the classrooms due to these professional development dimensions being in place.

Instructional and supportive leadership is another factor of effective schools (Marzano, 2003; Michigan State University, 2004). Although this was not specifically assessed by this study, this factor should be considered. All district leaders were supportive of the curriculum implementation as indicated by resource allocation, verbal support, and evidence of ongoing communication with the researcher and school building teachers. For example, resources were allocated for substitute teachers during professional development; and all leaders attended the training sessions, including the superintendent in District One. Principals would call or e-mail

the researcher to ask questions about the curriculum and reported that they visited classrooms to view the curriculum implementation.

Linkages between teacher quality, classroom practices (e.g., higher-order thinking skills and hands-on learning), and professional development in higher order thinking skills have also been found directly to influence student achievement (Wenglinsky, 2002). In this study, teachers were taught how to lead discussions that encourage higher level thinking skills as part of the professional development process. In addition, discussion and hands-on learning were processes embedded within the *Jacob's Ladder* curriculum, as was scaffolding from lower-order to higher-order thinking. Additional studies specific to the development of reading comprehension support the effects of teacher behaviors and their impact on student learning. Teachers who were most effective in teaching reading provided more coaching (Taylor et al., 2000), stressed higher order thinking skills in addition to lower order skills to building upon meaning (Knap et al., 1995); modeled thinking processes (NRP, 2000), incorporated small and whole group literary discussions (Chin, Anderson, & Waggoner, 2001; Guthrie, 1996); and included opportunities for feedback, self-regulation, and students' personal assessment of their reading progress. Treatment fidelity forms indicated that teachers used small group and whole group literacy discussion, included opportunities for feedback, and stressed higher order thinking skills as part of the curriculum implementation. The only research-based behavior not observed in any of the treatment classrooms was the students' personal assessment of reading progress. Teachers were not observed using the self-assessment forms that

came with the curriculum, which were intended to enhance personal assessment and monitoring of reading progress.

Another issue related to treatment fidelity was the percentage of curriculum implemented within each classroom. Approximately 30 to 80 percent of the entire curriculum was implemented, depending on the teacher. This range is quite large and could have impacted the study results. The two teachers who implemented 30% of the curriculum and the teacher who taught 80% of the curriculum were individually questioned about their use of the curriculum. The teachers who taught 30% of the curriculum reported that their students struggled with the curriculum and more whole-group time was necessary at the beginning of the implementation period to build student understanding. They also said that they taught the *Jacob's Ladder* curriculum during two to three reading periods per week but due to their students' need for additional assistance, they erred on the side of depth instead of breadth because they wanted to ensure understanding. The teacher who implemented 80% of the curriculum taught students gifted in language arts in a self-contained classroom. She felt she could move at a faster pace with her students. This type of differentiation, while confounding from a study perspective, suggests that the teachers were attuned to their students' needs and modified the time and level of independence necessary to promote understanding (Tomlinson, 1999). This level of sophistication in teaching may be found in expert teachers who have the ability to examine a curriculum, take ownership of it, and modify that curriculum based on the needs of their students to ensure learning (Bellanca, 1995).

Even with some of the modified changes as reported, the experimental teachers who participated in the study exhibited many of the curriculum features and intended processes according to the training as observed by the researcher. They also welcomed researcher feedback, and attempted to incorporate recommended changes in their teaching for the next researcher visitation. As previously discussed, comparison teacher classrooms were not observed, and data were not collected on the perceived effectiveness of the implementation, just on the fidelity of the implementation. Therefore it is difficult to know if the control teachers were just as or more capable and willing to implement the curriculum according to the training parameters and curriculum guidelines.

#### *Implications for Practice*

Findings from this study provide several implications for practice. Appropriate and high level curriculum is an important consideration for any school district (VanTassel-Baska & Stambaugh, 2006). The selection of textbooks and curriculum may directly influence what students learn (AFT, 1998; Tivnan & Hemphill, 2005). Likewise, teacher influence and instructional processes are important for student reading comprehension and critical thinking (NRP, 2000; Taylor et al., 2002). *Jacob's Ladder* incorporates many research-based practices that are similar to other effective critical thinking programs (see Cotton, 1991; Sternberg & Bhana, 1986 for a review of programs) and encourage strategies that enhance reading comprehension, especially for students in Title I schools (NRP, 2000; Taylor et al., 2001).

Other factors, some of which are beyond this study's scope, cannot be ignored as potential variables that may have impacted the outcome and should be considered

before a school district would endorse or encourage implementation of this program. Sternberg and Bhana (1985), after studying five different critical thinking curricula programs, suggested that consumers of research should be cautioned when considering results because other factors such as the teacher quality, administrative support, curriculum match to the population, and treatment fidelity are critical to a program's success. The purposeful, random selection of participating districts, included rural Title I schools that were ranked as successful by state standards. Cumulative research on effective Title I schools found several similar aspects for effective implementation of an innovation such as supportive leadership, ongoing professional development, teacher efficacy, and effective school-home relationships (Michigan State University, 2004; Muijs, et al., 2004).

Another consideration for practice is the issue of grade level standards required by each state. The comparison group curriculum was derived from the mandated state standards and the basal reader. States that have different standards from Ohio or different comparison curricula will need to examine their current standards and practices and compare them to the *Jacob's Ladder* curriculum to determine how it might enhance student achievement as a value-added curriculum feature.

In addition, ongoing professional development and support structures were in place for the use of the curriculum for this study. District leaders who consider adopting the *Jacob's Ladder* curriculum for use in their schools, need to be sure that teachers are provided the necessary resources and support structures for effective implementation. This includes teacher training, principal support, the provision of appropriate resources to implement the curriculum, accountability and support

structures such as classroom observations and feedback, and a school building environment that is conducive to implementing an innovative curriculum such as those schools selected for this study that were not at risk for failing but ranked “excellent” by the state of Ohio.

Reading teachers and regular classroom teachers who teach reading should be aware of the need for supplemental curriculum and the effects on student achievement, but the practices the processes inherent to good teaching of reading and critical thinking as incorporated within the *Jacob’s Ladder* curriculum. These include student coaching (Taylor et al., 2000), scaffolding from lower level to higher level thinking skills (Pressley et al., 2001), open-ended questioning (Beck & McKeown, 2001), modeling of thinking processes (Zemelman et al., 1998), literary discussion (Guthrie, 1996), linkages to personal experiences (Chin et al., 2001), and guided teacher feedback (Fielding & Pearson, 1994). These skills are even more important for the teaching of students in Title I schools due to a paucity of higher level thinking skills solicited by the teachers in Title I schools (Pressley et al., 2001), less exposure to appropriate language for students of poverty (Rothstein, 2004), and fewer opportunities for exposure to high level curriculum (AFT, 1998).

#### *Implications for Future Research and Next Steps*

Many questions arise from this study that warrant future research. First, this study was conducted in effective rural, Title I schools. Does this curriculum produce similar results in non Title I schools or in urban Title I schools? What about less effective Title I schools? Therefore, studies of the effectiveness of *Jacob’s Ladder* in other types of school districts are suggested.

Future research is also warranted in other states and additional school districts due to the difference in state standards and the comparison group curricula. Since the comparison group curricula in this study was connected to the state standards and no set curriculum was predominantly mandated, it is uncertain whether the *Jacob's Ladder* curriculum would be effective in states that use different language arts standards or use a specific reading curriculum or initiative.

Other variables such as teacher effectiveness and treatment fidelity may warrant future research. For example, this study did not include classroom observations of comparison teachers as part of the data collection effort. Therefore it is difficult to ascertain how much the teacher effect (or lack thereof) contributed to the overall success of the curriculum implementation. In addition teachers were not ranked based on effectiveness but were only assessed based on whether or not specific behaviors were observed. A treatment fidelity scale that includes rankings of effectiveness rather than bifurcated "observed" or "not observed" categories, could be created to assess research-based reading comprehension and critical thinking behaviors inherent in *Jacob's Ladder* in order to compare experimental and comparison classrooms quantitatively.

Since *Jacob's Ladder* was shown to be effective in reading comprehension with third and fourth graders, in particular, it may be interesting to determine the effects of the curriculum on different grade levels. For example, does the fifth grade curriculum also produce gains with lower-functioning middle school students or higher functioning fourth grade students? Similarly, is a new curriculum warranted

and would that new curriculum produce significant achievement gains for students in higher or lower grades?

In regard to specific findings, there are still questions about the impact of the *Jacob's Ladder* curriculum on reading comprehension for gifted learners. Future studies should include a separate off-level (at least one grade higher) standardized achievement measure for gifted students in order to accurately measure growth in reading comprehension and determine the differential effects of the program in this dimension.

Additional research regarding the effects of *Jacob's Ladder* on critical thinking is also warranted. The ANCOVA results revealed significant mean score gains between groups favoring the experimental group. However, the MANCOVA results did not reveal the same positive effects when other factors were considered. Studies with a larger sample size, more specific attention to the effects of critical thinking, and an extended length of treatment across multiple years would be warranted before conclusions about value-added gains in multiple aspects of critical thinking as measured by the TCT, could be definitively drawn.

In addition, the PBA assessment needs further refinement and testing. This assessment should be continued based on the research on the effectiveness of performance-based assessments that are matched to the curriculum (Gallagher, 2006) and may show a truer picture of performance for students of poverty (Callahan, 2005). However, modifications to the reading samples and specific pretest/posttest questions listed in the assessment could be made and further assessed for stronger reliability and validity.

Finally, additional research is needed to examine the effects of the *Jacob's Ladder Reading Comprehension Program* on fifth graders' reading comprehension, and possibly the fifth grade curriculum. Findings from this study are inconclusive regarding why *Jacob's Ladder* did not produce significant findings in reading comprehension for the fifth grade experimental group when compared to the comparison group. When conducting additional studies, it may also be helpful to observe and compare the exact practices of the comparison teachers to determine the differences in the curriculum and its implementation.

### *Conclusion*

*Jacob's Ladder* was found to be a promising program that produced significant and practical gains in rural Title I students' critical thinking ability and reading comprehension in the non-gifted group when compared to similar groups that used a variety of different reading curricula. While the program showed significant and very large practical gains within the experimental group on targeted critical thinking skills, results between groups is less definitive when assessing a broader range of critical thinking behaviors with other factors. For reading comprehension, *Jacob's Ladder* is most effective for third and fourth grade nongifted students. Additional research regarding the impact of the curriculum on gifted students' and on other students with diverse profiles appears warranted.

# Appendices

**Appendix A:**

**Pre and Post Performance-Based Assessments with Rubric**

## Rubric

### Implications and Consequences

0=Provides no response or response is inappropriate to the task demand

1=limited response, inaccurate, confusing, or copies from text

2=response is accurate and makes sense but does not adequately address all components of the question or provide rationale from the text

3=response is accurate, answers all parts of the question, provides a rationale that justifies the answer

4=response is well written, specific, insightful, answers all parts of the question, offers substantial support, and incorporates evidence from the text

### Inference

0=Provides no response or a response inappropriate to the task demand

1=limited, vague, inaccurate; rewording of the prompt or copies from text

2=accurate response but literal interpretation with no support from the text

3=interpretive response with limited support from the text

4=insightful, interpretive, well-written response with substantial support from the text

### Theme/Generalization

0=provides no response or a response inappropriate to the task demand

1=limited, vague, inaccurate; only uses quotes from the story

2=literal description of the story without explaining the moral; no reasons why

3=valid interpretive moral with no reasoning

4=provides an insightful, interpretive moral, with substantial justification or reasoning

### Creative Synthesis

0=Provides no response or a response inappropriate to the task demand

1=limited, vague, or inaccurate title, no reasoning

2=appropriate but literal title with no attempt to support

3=interpretive title with limited reasoning or justification

4=insightful title, interpretive, and extensive justification or reasoning

*The Fox and the Leopard (Aesop)*

*The Fox and the Leopard disputed which was the more beautiful of the two. The Leopard exhibited one by one the various spots which decorated his skin. But the Fox, interrupting him, said, "And how much more beautiful than you am I, who am decorated, not in body but in mind."*

### Third Grade Pre-Test: The Fox and the Leopard

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the story, *The Fox and The Leopard*. Answer the four questions related to the story.

Who do you think is more beautiful, the fox or the leopard? Why? Provide evidence from the story to defend your answer.

What does the fox mean when he says “And how much more beautiful than you am I, who am decorated, not in body, but in mind”? Provide evidence from the story to defend your answer.

| What is the moral of this story? Give a reason why you think so.

Create a new title for this fable. Give a reason why your title is better than the original title.

*The Frogs and the Well*  
(Aesop)

*Two Frogs lived together in a marsh. But one hot summer the marsh dried up, and they left it to look for another place to live in: for frogs like damp places if they can get them. By and by they came to a deep well, and one of them looked down into it, and said to the other, "This looks a nice cool place. Let us jump in and settle here." But the other, who had a wiser head on his shoulders, replied, "Not so fast, my friend. Supposing this well dried up like the marsh, how should we get out again?"*

### Third Grade Post-Test: The Frogs and the Well

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the story, *The Frogs and the Well*. Answer the four questions related to the story.

What do you think would have happened if the frogs jumped in the well? Provide evidence from the story to defend your answer.

The wise frog said, "Not so fast, my friend. Supposing this well dried up like the marsh, how should we get out again?" What made this question so important for the frog to ask? Provide evidence from the story to defend your answer.

What is the moral of this story? Give a reason why you think so.

Create a new title for this fable. Give a reason why your title is better than the original title.

*The Old Woman and the Physician*  
(Aesop)

*An Old Woman having lost the use of her eyes, called in a Physician to heal them, and made this bargain with him in the presence of witnesses: that if he should cure her blindness, he should receive from her a sum of money; but if her infirmity remained, she should give him nothing. This agreement being made, the Physician, time after time, applied his salve to her eyes, and on every visit took something away, stealing all her property little by little. And when he had got all she had, he healed her and demanded the promised payment. The Old Woman, when she recovered her sight and saw none of her goods in her house, would give him nothing. The Physician insisted on his claim, and as she still refused, summoned her before the Judge. The Old Woman, standing up in the Court, argued: "This man here speaks the truth in what he says; for I did promise to give him a sum of money if I should recover my sight: but if I continued blind, I was to give him nothing. Now he declares that I am healed. I on the contrary affirm that I am still blind; for when I lost the use of my eyes, I saw in my house various chattels and valuable goods: but now, though he swears I am cured of my blindness, I am not able to see a single thing in it."*

**Fourth Grade Pre-Test: The Woman and the Physician**

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the story, *The Woman and the Physician*. Answer the four questions related to the story.

What do you think the judge will do in this case? Why? Provide evidence from the story to defend your answer.

What does the old woman mean when she says “but now, though he swears I am cured of my blindness, I am not able to see a single thing...”? Provide evidence from the story to defend your answer.

What is the moral of this story? Give a reason why you think so.

Create a new title for this story. Give a reason why your title is better than the original title.

## ***The King and the Shirt***

*A king once fell ill. "I will give half my kingdom to the man who can cure me," he said. All his wise men gathered together to decide how the king could be cured. But no one knew. Only one of the wise men said what he thought would cure the king. "If you can find a happy man, take his shirt, put it on the king – and the king will be cured." The king sent his emissaries to search for a happy man. They traveled far and wide throughout his whole kingdom, but they could not find a happy man. There was no one who was completely satisfied: if a man was rich he was ailing; if he was healthy he was poor; if he was rich and healthy he had a bad wife; or if he had children they were bad – everyone had something to complain of. Finally, late one night, the king's son was passing by a poor little hut and he heard someone say: "Now, God be praised, I have finished my work, I have eaten my fill, and I can lie down and sleep! What more could I want?" The king's son rejoiced and gave orders that the man's shirt be taken and carried to the king, and that the man be given as much money as he wanted. The emissaries went in to take the man's shirt, but the happy man was so poor that he had no shirt.*

### Fourth Grade Post-Test: The King and the Shirt

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the story, *The King and the Shirt*. Answer the four questions related to the story.

What do you think will happen since the happy man has no shirt? Why? Provide evidence from the story to defend your answer.

Why do you think the man without a shirt was happy when no one else was? Provide evidence from the story to defend your answer.

What is the moral of this story? Give a reason why you think so.

Create a new title for this story. Give a reason why your title is better than the original title.

## Poem by Emily Dickinson

*THIS is my letter to  
the world,  
That never wrote to  
me,—  
The simple news that  
Nature told,  
With tender majesty.*

*Her message is  
committed  
To hands I cannot  
see;  
For love of her, sweet  
countrymen,  
Judge tenderly of  
me!*

### **Fifth Grade Pre-Test: Emily Dickinson Poem**

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the poem by Emily Dickinson. Answer the four questions related to the poem.

What does the author think about the world? Provide evidence from the story to defend your answer.

What did the author mean when she wrote “The simple news that Nature told with tender Majesty”? Provide evidence from the story to defend your answer.

What do you think this poem is about? Give a reason why you think so.

Create a title for this poem. Give a reason why your title is good.

*Poem by Emily Dickinson*

*There is no frigate like a book  
To take us lands away,  
Nor any coursers like a page  
Of prancing poetry.  
This traverse may the poorest take  
Without oppress of toll;  
How frugal is the chariot  
That bears a human soul!*

**Fifth Grade Post-Test: Emily Dickinson Poem**

Name \_\_\_\_\_ Grade \_\_\_\_\_

Please read the poem by Emily Dickinson. Answer the four questions related to the poem.

What does the author think about books? Provide evidence from the story to defend your answer.

A frigate is a small warship. Why does the author compare a book to a frigate? Provide evidence from the story to defend your answer.

What one word best describes what this poem is about? Give a reason why you think so.

Create a title for this poem. Give a reason why your title is good.

**APPENDIX B:**

**Treatment Fidelity Form**

**Jacob's Ladder Treatment Fidelity Form**

**Teacher:** \_\_\_\_\_ **Date:** \_\_\_\_\_ **Observation #:** \_\_\_\_\_

<b>Implementation</b>	<b>Observed</b>	<b>Not Observed</b>	<b>Comments</b>
Students complete initial answers individually			
Students are grouped in dyads for discussion			
Students are completing self-evaluations			
Students are completing record sheets			
Teacher is differentiating reading selections based on student strengths and weaknesses			
Teacher is providing student feedback			
Students are completing reading selections from each genre			
Students are discussing literature as a whole group			
Students and teacher are conferring on readings			

Used with permission from French, 2005.

**APPENDIX C:**  
**Ladder A-D Templates**

A 3	<i>Consequences and Implications</i>	<i>Title of Reading Selection:</i>
A 2	<i>Cause and Effect</i>	
A 1	<i>Sequencing</i>	

B 3	<i>Generalizations</i>	
B 2	<i>Classifications</i>	
B 1	<i>Details</i>	

*Title of Reading Selection:*

C 3	<i>Theme/Concept</i>	<i>Title of Selection:</i>
C 2	<i>Evidence/Inference</i>	
C 1	<i>Characterization</i>	

D 3	<i>Creative Synthesis</i>	<i>Title of Reading Selection:</i>
D 2	<i>Summarizing</i>	
D 1	<i>Paraphrasing</i>	

Appendix D:

*Ohio Content Standards by Grade Level Cluster*

*Ohio Content Standards by Grade Kindergarten to Grade Four*

By the end of grade four, the following expectations are to be met (Ohio Department of Education, 2001, pp. 157-268):

- 1) apply structural analysis skills to build and extend vocabulary and to determine word meaning,
- 2) use context clues to determine the meaning of new vocabulary,
- 3) establish a purpose for reading and use a range of reading comprehension strategies to understand literary passages and texts,
- 4) draw conclusions from information in text,
- 5) apply reading skills and strategies to summarize and compare and contrast information in text, between text, and across subject areas,
- 6) demonstrate comprehension by responding to questions (literary, informational, and evaluative),
- 7) identify central ideas and supporting details of informational text,
- 8) ask clarifying questions concerning essential elements of information text,
- 9) use text features and structures to organize content, draw conclusions, and build text knowledge,
- 10) identify a theme of literary text,
- 11) use supporting details to identify and describe main ideas, characters, and setting, and
- 12) explain how an author's word choice and use of methods influences the reader.

*Ohio Content Standards Grades Five to Seven*

By the end of the seventh grade, students are expected to:

- 1) use context clues and text structures to determine the meaning of new vocabulary,
- 2) infer meaning of words using context clues,
- 3) make meaning through asking and responding to a variety of questions related to the text,
- 4) apply effective reading comprehension strategies including summarizing and making predictions and comparisons using information in text, between text, and across subject areas,
- 5) apply self-monitoring strategies to clarify confusion about text and to monitor comprehension,
- 6) use text features and graphics to organize, analyze, and draw inferences from content and to gain additional information,
- 7) recognize the difference between cause and effect and fact and opinion to analyze the text,
- 8) explain how main ideas connect to each other in a variety of sources,
- 9) identify arguments and persuasive techniques used in informational text,
- 10) explain the treatment, scope and organization of ideas from different texts to draw conclusions about a topic,
- 11) determine the extent to which summary accurately reflects the main ideas, critical details, and underlying meaning of original text,
- 12) describe and analyze the elements of character development,

- 13) analyze the importance of setting,
  - 14) identify elements of plot and establish a connect between an element and a future event,
  - 15) differentiate between different points of view in narrative text,
  - 16) demonstrate comprehension by inferring themes, patterns and symbols,
  - 17) identify similarities and differences of various forms and genres, and
  - 18) explain how figurative language expresses ideas and conveys mood
- (Ohio Department of Education, 2001, pp. 157-268).

**Appendix E:  
Focus Group Emergent Themes**

<i>Theme</i>	<i>Supporting Comments from Teachers</i>	<i>Supporting Comments from Students</i>
Higher Level of Challenge and Questioning	<ul style="list-style-type: none"> <li>• Parents were not asking for more challenging work for their children (this year). Students weren't bored.</li> <li>• Questioning (in JL) pushes a child's thinking process to a higher level</li> <li>• (I am) Expecting more in-depth thinking through open-ended questions</li> <li>• Students were really thinking about what they read and re-read</li> <li>• JL challenges me as a teacher; I've had to re-evaluate. Do I always expect one right answer?</li> <li>• My expectations are higher - no going back</li> <li>• The selections were way over my students. We did not have a lot of students able to work in smaller groups as few understood the content without the teacher leading</li> </ul>	<ul style="list-style-type: none"> <li>• What I liked least about <i>Jacob's Ladder</i> is that the questions would sometimes get really challenging and I don't like really challenging questions.</li> <li>• I can understand more big words like perseverance and moral. Now I know what they mean. The questions were higher and at my level. The story "The Cottage" helped me understand perseverance.</li> <li>• I learned to work more hard and everything isn't always easy so I learned to challenge myself more. It wasn't very hard, but more hard than I usually work.</li> <li>• I like the challenge it (JL) gives me. It makes me dig deeper and think harder. It words the stories good. The poetry has nice words and I like discussing it out loud.</li> </ul>
Value-Added Gains and Transfer (vocabulary, transfer of learning to other content areas, new skills and lessons learned and practiced)	<ul style="list-style-type: none"> <li>• I use ladders in other content areas.</li> <li>• I saw my students' way of thinking change over time. The students in my class answer questions a whole new way, often times making sure they have "evidence" to back it up</li> <li>• Students are much better at answering 2-part questions where they must defend or prove their answer using the selection. Students backed up their answers better.</li> <li>• Students would reference</li> </ul>	<ul style="list-style-type: none"> <li>• The poetry inspired me to write poetry based on the poem, Delilah. My friends have me write about them. I keep my own poetry book.</li> <li>• What I like about <i>Jacob's Ladder</i> is that the stories can get very interesting. <i>Jacob's Ladder</i> got me to start liking reading a lot better.</li> <li>• I like <i>Jacob's Ladder</i> because it is more fun than testing. Also it has selections about historical events and both nonfiction and fiction stories. It makes</li> </ul>

	<p>reading passages throughout the week</p> <ul style="list-style-type: none"> <li>• Parent commented that her daughter is now reading independently at home</li> <li>• I used the JL model with <i>Charlotte's Web</i> and students are using more discussion; no more yes/no questions (for me)</li> <li>• I was amazed at the level of vocabulary used in <i>Jacob's Ladder</i>. My students used it in other subject areas.</li> <li>• Vocabulary, questioning skills, and more extended responses became visible in other content areas such as math, science, and social studies.</li> </ul>	<p>me more enjoyed reading. I loved it!!</p> <ul style="list-style-type: none"> <li>• I had to look back in the book a lot.</li> <li>• I learned about space and ocean animals.</li> <li>• I learned to write more complete sentences and I learned new words. It was hard at first but then it was easy.</li> <li>• Before <i>Jacob's Ladder</i> I didn't know some of the words and now I know a lot more words. I know this because there are a lot more interesting words in the texts.</li> <li>• I would like to do it (JL) with the class more. You could also do different subjects such as math, science, social studies, etc.</li> <li>• I learned how to write summaries and they were a little harder than what I thought it would be.</li> </ul>
<p>Solicitation of Effective Conversation About Literature</p>	<ul style="list-style-type: none"> <li>• JL brought back the lost art of conversation into the classroom"</li> <li>• "Students enjoyed challenging others to defend their answers"</li> <li>• Their ideas were theirs and not the teachers; they really enjoyed the discussions and sharing; more eager to get in to it</li> <li>• Discussions became more in-depth and more students would participate</li> <li>• Typically quiet students were sharing their ideas and thoughts; through <i>Jacob's Ladder</i> I saw students who</li> </ul>	<ul style="list-style-type: none"> <li>• I liked it best when we did the stories' ladders and when we did some in partners and groups...</li> <li>• I liked the group discussions about the poems and stories we read.</li> <li>• The stories and poems were very interesting and the questions made me think more. I liked being partners so we could share ideas.</li> <li>• The thing I liked best about <i>Jacob's Ladder</i> was reading all the poems and getting into groups to discuss.</li> <li>• I like the activities best about <i>Jacob's Ladder</i></li> </ul>

---

	<p>were not confident blossom and encourage others...</p> <ul style="list-style-type: none"> <li>• I now trust students to actually work in small groups due to whole group modeling</li> <li>• I have a hard time closing down class discussion</li> </ul>	<p>because we get to work with partners.</p>
<p>Time Constraints</p>	<ul style="list-style-type: none"> <li>• Nonfiction selections are too long for just a 1-2 day ladder</li> <li>• Implementation time constraints was an issue. I would have loved to implement this over a more lengthy time span. Many times I felt “rushed” to do a ladder. It felt like a chore.</li> <li>• Doing so many per week, caused them (students) to become less enthused with the program.</li> <li>• It was hard to teach JL and the basal reader.</li> <li>• This was a top-down decision with no buy in. (We) could have had a longer time span before implementation to figure out how it (JL) best fits our current units</li> </ul>	
<p>Program Inconsistency</p>	<ul style="list-style-type: none"> <li>• Some of the ladders had too many similar activities or too much writing for one story. Students became burned out.</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Jacob’s Ladder</i> was easy and hard</li> <li>• What I recommend you change is some of the extended response to multiple choice so we could get used to it for the <i>Jacob’s Ladder</i> test.</li> <li>• I recommend only two or three things – one thing is I would change some of the extended response questions to multiple choice questions.</li> </ul>

---

---

Also, I would get a book or website to tell all people about the *Jacob's Ladder*

- It took a long time to answer the hard questions.
  - Some of the stories were too long. The thing I liked least was some were really long stories at least 4 pages long like "The Fisherman and His Wife."
  - I really like *Jacob's Ladder* but sometimes the really long answers make my hand cramp from writing
  - Maybe you could make some of the stories more interesting.
-

## References

- Abrams, Z. (2005). Asynchronous CMC, collaboration and the development of critical thinking in a graduate seminar in applied linguistics. *Canadian Journal of Learning and Technology*, 31(2).
- Allington, R.L. (2001). *What really matters for struggling readers: Designing research-based programs*. New York: Addison Wesley Longman.
- American Federation of Teachers. (1998). *Building on the best, learning from what works: Seven promising reading and english and language arts programs*. Washington DC: Author.
- Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. Bower (Ed.), *Psychology of Learning and Motivation*, (pp. 98-111). Burlington, MA: Academic Press, Inc.
- Anderson, R., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first grade groups. *Elementary School Journal*, 28, 183-223.
- Anderson, L.W., & Krathwohl, D.R. (Eds.). (2001). *A taxonomy of learning, teaching, and assessment: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Appleman, D., & Applebee, A. (2000). *Critical encounters in high school English: Teaching literacy theory to adolescents*. New York: Teachers College Press.
- Banks, J.A. (1993). Multicultural education: Development, dimensions, and challenges. *Phi Delta Kappan*, 75(1), 22-28.
- Barton, J., & Sawyer, D.M. (2003). Our students are ready for this: Comprehension instruction in the elementary school. *The Reading Teacher*, 57(4), 334-347.

- Baskin, B. & Harris, K. (1980). *Books for the Gifted Child*. New York: R. R. Bowker Company.
- Beck, I.L., & McKeown, M.G. (2001). Inviting students into the pursuit of meaning. *Educational Psychology Review*, 13(3), 225-241.
- Bellanca, J. (1995). *Designing Professional Development for change*. Arlington Heights, IL: IRI/Skylight Training and Publishing, Inc.
- Benbow, C. P. (1998). Grouping intellectually advanced students for instruction. In J. VanTassel-Baska (Ed.), *Excellence in educating gifted and talented learners (3<sup>rd</sup> ed.)* (pp. 261-278). Denver, CO: Love.
- Benbow C. P., & Stanley, J.C. (Eds.). (1983). *Academic precocity: Aspects of its development*. Baltimore: Johns Hopkins University Press.
- Biancarosa, G., & Snow, C.E. (2004). *Reading next: A vision for action and research in middle and high school literacy*. A report to the Carnegie Corporation of New York. Washington, DC: Alliance for Excellent Education.
- Block, C. (2004). *Teaching comprehension: The comprehension process approach*. Boston: Allyn & Bacon.
- Block, C., & Israel, S. (2004). The ABCs of performing highly effective think-alouds: Effective think-alouds can build students' comprehension, decoding, vocabulary, and fluency. *The Reading Teacher*, 58, 154-169.
- Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: David McKay Company.
- Boyle, J., & Weishaar, M. (1997). The effects of expert-generated versus student-generated cognitive organizers on the reading comprehension of students with

learning disabilities. *Learning Disabilities Research & Practice*, 12(4) 228-235.

Bracken, B.A., VanTassel-Baska, J., Brown, E.F., & Feng, A. (2007). Project Athena: A tale of two studies. In J. VanTassel-Baska & T. Stambaugh (Eds.), *Overlooked Gems: A national perspective on low-income promising learners* (pp. 63-67). Washington DC: National Association of Gifted Children.

Bruner, J. (1961). *The process of education*. Cambridge, MA: Harvard University Press.

Caccamise, D., & Snyder, L. (2005). Theory and pedagogical practices of text comprehension. *Topics in Language Disorders*, 24(1), 5-31.

Callahan, C. (2005). Identifying gifted students from underrepresented populations. *Theory into Practice*, 44, 98-105.

Campbell, D., & Stanley, J. (1969). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Campbell, J.R., Voelkl, K.E., & Donahue, P.L. (1997). *NAEP 1996 trends in academic progress*. Washington, DC: National Center for Education Statistics.

Carter, C.J. (1997). Why reciprocal teaching? *Educational Leadership*, 54(6), 64-69.

Center for Gifted Education (n.d.). *Test of Critical Thinking examiners manual*. Williamsburg, VA, College of William and Mary: Author.

Center for Public Education. (2006). *High-performing, high-poverty schools: Key lessons learned from research*. Alexandria, VA: Author.

Chin, C.A., Anderson, R.C., & Waggoner, M.A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly*, 30, 278-411.

- Clark, B. (2002). *Growing up gifted* (6th ed.). Columbus, OH: Merrill Prentice Hall.
- Clay, M., & Cazden, C. (1992). A Vygotskian interpretation of reading recovery. In L.C. Moll (Ed.), *Vygotsky and education: Instructional implications and applications of socio-historical psychology* (pp. 206-222). New York: Cambridge University Press.
- Colangelo, N., Assouline, S., & Gross, M.U.M. (2004). *A nation deceived: How schools hold back America's brightest students*. Iowa City, IA: Belin-Blank Center.
- Cotton, K. (1991). *Teaching thinking skills*. Retrieved January 22, 2007 from Northwest Regional Educational Laboratory's School Improvement Research Series: <http://www.nwrel.org/scpd/sirs/6/cu11.html>
- Daniels, H. (1994). *Literature circles: Voice and choice in the student centered classroom*. York, ME: Stenhouse Publishers.
- Donovan, M.S., & Cross, C.T. (Eds.). (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.
- Duffy, G. G., Roehler, L. R., & Mason, J. (1984). *Comprehension Instruction: Perspectives and Suggestions*. New York, NY: Longman.
- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Duke, N., & Pearson, D. (2002). Effective Practices for Developing Reading Comprehension. In A.E. Farstrup & S.J. Samuels (Eds.), *What Research Has To Say About Reading Instruction (3rd ed.)* (pp. 205–242). Newark, DE:

International Reading Association.

- Ennis, R. (1989). Critical thinking and subject-specificity: clarification and needed research. *Educational Researcher*, 18, 4-10.
- Ewing, N.J., & Yong, F.L. (1993). Learning style preferences of gifted minority students. *Gifted Education International*, 9, 40-44.
- Facione, P.A. (2007). *Critical thinking: What it is and why it counts (Revised)*. Millbrae, CA: California Academic Press.
- Feldhusen, J.F., & Moon, S.M. (1995). The educational continuum and delivery of services. In Genshaft, J.L., Bireley, M., & Hollinger, C.L. (Eds.), *Serving gifted and talented students: a resource for school personnel* (pp. 103-121). Austin, TX: Pro-Ed.
- Feng, A., VanTassel-Baska, J., Quek, C., Bai, W., O'Neill, B. (2004). A longitudinal assessment of gifted student's learning using the integrated curriculum model: Impacts and perceptions of the William and Mary language arts and science curriculum. *Roeper Review*, 27(2), 78-83.
- Fielding, L., & Pearson, D. (1994). Synthesis of Research / Reading comprehension: What works. *Educational Leadership*, 51(5), 62-68.
- Ford, D.Y. (1995). *A study of underachievement among gifted, potentially gifted, and general education students*. Storrs, CT: University of Connecticut, The National Research Center of the Gifted and Talented.
- Ford, D.Y., & Harris, J.J. (1999). *Multicultural gifted education*. New York: Teachers College Press.
- Ford, D.Y., & Trotman, M.F. (2001). *Teachers of gifted students: Suggested*

- multicultural characteristics and competencies. *Roeper Review*, 23, 235-239.
- Fournier, D., & Graves, M.F. (2002). Scaffolding adolescents' comprehension of short stories. *Journal of Adolescent & Adult Literacy*, 46(1), 30-40.
- Frasier, M.M., & Passow, A.H. (1994). *Toward a new paradigm for identifying talent potential* (RM 9412). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.
- French, H. (2005). Pilot study of Jacob's Ladder Reading Comprehension Program. Dissertation, Williamsburg, VA: College of William and Mary.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction (6<sup>th</sup> ed.)*. White Plains, NY: Longman.
- Gallagher, J. (2006). How to shoot oneself in the foot with program evaluation. *Roeper Review* 28(3), 122-124.
- Graves, M., & Avery, P. (1997). Scaffolding students' reading of history. *The Social Studies*, 88(3), 134-139.
- Graves, M., Graves, B., & Braaten, S. (1996). Scaffolded reading experiences for inclusive classes. *Educational Leadership*, 53(5), 14-17.
- Gregory, G.H., & Kuzmich, L. (2004). *Data driven differentiation in the standards-based classroom*. Thousand Oaks, CA: Corwin Press.
- Grimm, L., & Yarnold, P. (1995) (Eds.). *Reading and Understanding multivariate statistics*. Washington, DC: American Psychological Association.
- Gomez, R.A. (1991). *Teaching with a multicultural perspective*. Urbana, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. (ERIC Document Reproduction Service No. ED339548)

- Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*(5), 5-12.
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8*(3), 381-391.
- Guthrie, J.T. (1996). Educational contexts for engagement in literacy. *Reading Teacher, 49*(6), 432-445.
- Guthrie, J.T., Schafer, W.D., & Huang, C. W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *Journal of Educational Research, 94*, 145-162.
- Hadaway, N., & Marek-Schroer, M.F. (1992). Multidimensional assessment of the gifted minority student. *Roeper Review, 15*, 73-77.
- Halstead, J. (2002). *Some of my best friends are books: Guiding gifted readers from preschool to high school*. Scottsdale, AZ: Great Potential Press.
- Han, K.S., & Marvin, C. (2000). A five year follow-up study of the Nebraska Project: Still a long way to go... *Roeper Review, 23*, 25-33.
- Hansen, C. (2004). Teacher talk: Literacy development through response to story. *Journal of Research in Childhood Education, 19*(2), 115-129.
- Hart, B.H., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.
- Haycock, K. (2000). No more settling for less. *Thinking K-16: A publication of the Education Trust, 4*(1), 6-12.
- Hebert, T., & Beardsley, T. (2001). Jermaine: A critical case study of a gifted black child living in rural poverty. *Gifted Child Quarterly, 45*(2), 85-103.

- Henry, A., & Roseberry, R. (1996). A corpus-based investigation of the language and linguistic patterns of one genre and the implications for language teaching. *Research in the Teaching of English, 30*, 472-489.
- Hoover, H.D., Dunbar, D.A., Frisbie, D.A., Oberley, K.R., Bray, G.B., Naylor, R.J., Lewis, J.C., Ordman, V.L., & Qualls, A.L. (2003). *Iowa tests of basic skills survey battery: Norms and score conversions*. Itasca, IL: Riverside Publishing.
- Hudson, P., Lingnugaris-Kraft, B., & Miller, T. (1993). Using content enhancements to improve the performance of adolescents with learning disabilities in content classes. *Learning Disabilities Research & Practice, 8*(2), 106-126.
- Johnson, D., & Johnson, R. (1989) *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company.
- Johnson, R.T., & Johnson, D.W. (1994). *An overview of cooperative learning*. Retrieved February 18, 2006, from <http://www.cooperation.org/pages/overviewpaper.html>
- Joseph, L., & Ford, D. (2006). Nondiscriminatory assessment: Considerations for gifted education. *Gifted Child Quarterly, 50*(1), 42-51.
- Joyce, B.R., & Showers, B. (1988). *Student achievement through staff development*. New York: Longman.
- Kelley, M.F. (2004). Research into practice. *Journal of Research in Childhood Education, 19*, 189-194.
- Kiess, H.O. (2002). *Statistical concepts for the behavioral sciences (3<sup>rd</sup> ed.)*. Boston: Allyn and Bacon.
- Kintsch, W. (2005). *Scalable and sustainable technologies for reading instruction and*

*assessment*. Retrieved November, 14, 2005 from <http://drdc.uchicago.edu/community/project.phtml?projectID=50>

- Knapp, M.S., Adelman, N.E., Marder, C., McCollum, H., Needels, M.C., Padilla, C., Shields, P.M., Turnbull, B.J., & Zucker, A.A. (1995). *Teaching for meaning in high-poverty classrooms*. New York: Teachers College Press.
- Kulik, C., & Kulik, J. (1997). Ability grouping and gifted students. In Colangelo & G. Davis (Eds.), *Handbook of gifted education*, (pp.178-196). Boston, MA: Allyn & Bacon.
- Kulik, J., & Kulik, C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly*, 36, 73-77.
- Little, C. A. (2002). Reasoning as a key component of language arts curricula. *The Journal of Secondary Gifted Education*, 13(2), 52-59.
- Lubliner, S. (2002, April). *The power of clarifying: A comparative analysis of strategies that strengthen comprehension*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Lyman, L.L., & Villani, C.J. (2002). The complexity of poverty: A missing component of educational leadership programs. *Journal of School Leadership*, 12(3), 246-280.
- Lynch, S. J. (1992). Fast paced high school science for the academically talented: A six-year perspective. *Gifted Child Quarterly*, 36, 147-154.
- Marzano, R.J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

- McLaughlin, M., & DeVogd, G. (2004). Critical literacy as comprehension: Expanding reader response: critical literacy helps teachers and students expand their reasoning, seek out multiple perspectives, and become active thinkers. *Journal of Adolescent & Adult Literacy*, 48(1), 52-63.
- McPeck, J. (1990) (Ed.). *Teaching Critical Thinking: Dialogue and Dialectic*. New York: Routledge.
- Michigan State University (2004). Characteristics of effective elementary school in poverty areas. *Best Practice Briefs*, 29, 1-4.
- Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving school in socioeconomically disadvantaged areas – a review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149-175.
- National Center for Education Statistics. (2004). *The nation's report card: Reading highlights 2003*. Washington, D.C: Institute of Education Sciences, U.S. Department of Education.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Health and Human Development.

- Newell, G. (1996). Reader-based and teacher-centered instructional tasks: Writing and learning about a short story in middle-task classrooms. *Journal of Literary Research, 28*, 147-171.
- No Child Left Behind Act of 2001*. (2002). Pub. L. No. 107-110, paragraph 115, Stat. 1425.
- Nordlund, M. (2003). *Differentiated instruction: Meeting the educational needs of all students in your classroom*. Lanham, MD: Scarecrow Education.
- Oakes, J. (1995). Two cities' tracking and within-school segregation. *Teacher College Record, 96*(4), 681-690.
- Ohio Department of Education. (2001). *Academic content standards: K-12 english language arts*. Columbus, OH: Author.
- Ohio Revised Code 3301-51-15. *Identification and services for children who are gifted*.
- Oster, L. (2001). Using the think-aloud for reading instruction. *The Reading Teacher, 55*, 65-69.
- Palinscar, A.S., & Brown, A.L. (1984). Reciprocal-teaching of comprehension-monitoring activities. *Cognition and Instruction, 1*, 117-175.
- Parke, B.N. (1995). Developing curricular interventions for the gifted. In J.L. Genshaft, M. Bireley, & C.L. Hollinger, (Eds.), *Serving gifted and talented students: a resource for school personnel* (pp. 123-134). Austin, TX: Pro-Ed.
- Patton, M. (2002). *Qualitative research & evaluation methods (3<sup>rd</sup> ed.)*. Thousand Oaks, CA: Sage Publications, Inc.
- Paul, R. W. (1992). *Critical thinking: What every person needs to know to survive in*

- a rapidly changing world* (2<sup>nd</sup> ed.). Santa Rosa, CA: Foundation for Critical Thinking.
- Paul, R., & Elder, L. (2004). *The miniature guide to the art of asking essential questions* (2<sup>nd</sup> ed.). Dillon Beach, CA: The Foundation for Critical Thinking.
- Payne, R.K. (1995). *A framework: Understanding and working with students and adults from poverty*. Baytown, TX: RFT Publishing.
- Pedhazur, E., & Pedhazur-Schmelkin, L. (1991) *Measurement Design, and Analysis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.
- Perie, M., Grigg, W., & Donahue, P. (2005). *The Nation's Report Card: Reading 2005* (NCES 2006-451). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Perkins, D., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, 18(1) 16-25.
- Peterson, J.S. (1999). Gifted through whose cultural lens? An application of the postpositivistic model of inquiry. *Journal for the Education of the Gifted*, 22(4), 354-383.
- Pressley, M. (1990). *Cognitive strategy instruction that really improves children's academic performance*. Cambridge, MA: Brookline Books.
- Pressley, M. (1998). *Reading instruction that works: The case of balanced teaching*. NY: Guilford.
- Pressley, M., Wharton-McDonald, R., Allington, R., Block, C.C., Morrow, L.,

- Tracey, D., Baker, K., Brooks, G., Cronin, J., Nelson, E., & Woo, D. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5*, 35-58.
- Pressley, M., Wharton-McDonald, R., Raphael, L.M., & Roehrig, A. (2002). Exemplary first grade teaching. In B.M. Taylor & P.D. Pearson (Eds.), *Teaching reading: Effective schools, accomplished teachers* (pp. 73-88). Mahwah, NJ: Erlbaum.
- Puma, M. J., Karweit, N., Price, C., Ricciuti, A., Thompson, W., & Vaden-Kiernan, M. (1997). *Prospects: Final report on student outcomes*. Washington, DC: Planning and Evaluating Service, U.S. Department of Education.
- Rawson, K.A., Dunlosky, J., & Thiede, K.W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition, 28*(6), 1004.
- Redfield, D.L., & Rousseau, E.W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research, 51*, 237-245.
- Reis, S., Gubbins, E., Briggs, C., Schreiber, F., Richards, S., Jacobs, J., Eckert, R., & Renzulli, J. (2004). Reading instruction for talented readers: Case studies documenting few opportunities for continuous progress. *Gifted Child Quarterly 48*(4), 315-338.
- Reis, S.M., Westberg, K.L., Kulikowich, J., Caillard, F., Hebert, T., & Plucker, J. (1993). *Why not let high ability students start school in January? The*

- curriculum compacting study* (Research Monograph 93106). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.
- Renzulli, J. (1997). *Five dimensions of differentiation. Keynote presentation at the 20<sup>th</sup> Annual Confratute Conference*, Storrs, CT.
- Riverside Publishing. (2007). *Iowa Test of Basic Skills, Forms A & B*. Retrieved January 22, 2007 from [www.riverpub.com/products/itbs/details.html](http://www.riverpub.com/products/itbs/details.html)
- Rivkin, S.G., Hanushek, E.A., & Kain J.F. (2000). *Teachers, schools, and academic achievement*. National Bureau of Economic Research, Working Paper No. 6691.
- Robinson, A. (1990). Cooperation or exploitation? The argument against cooperative learning for talented students. *Journal for the Education of the Gifted*, 19(1), 14-17.
- Robinson, N., Lanzi, R., Weinberg, R., Ramey, S., & Ramey, C. (2002). Family factors associated with high academic competence in former Head Start children at third grade. *Gifted Child Quarterly*, 46(4), 278-291.
- Roehler, L.R. & Duffy, G.G. (1984). Direct explanation of comprehension processes. In G.G. Duffy, L.R. Roehler, & J. Mason (Eds.), *Comprehension Instruction: Perspectives and suggestions* pp. 265-280. New York: Longman.
- Rogers, K.B. (1998). Using current research to make “good” decisions about grouping. *NASSP Bulletin*, 82, 38-46.
- Rosenblatt, L. (2004). The transactional theory of reading and writing. In R. Ruddell, M. Muddell, & H. Singer (Eds.), *Theoretical models and processes of reading*

- (pp. 1363-1398). Newark, DE: International Reading Association.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*(2), 181-221.
- Rothstein, R. (2004). *Even the best schools can't close the race achievement gap*. Poverty and Race Research Action Council: Washington DC.
- Rowan, B., Correnti, R., & Miller, R.J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects student of elementary schools. *Teachers College Record*.
- Sadoski, M., Willson, V., & Norton, D. (1997). The relative contributions of research-based composition activities to writing improvement in the lower and middle grades. *Research in the Teaching of English, 31*, 120-147.
- Sanders, W.L., & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.
- Sarroub, L. K., & Pearson, P. D. (1998). Two steps forward, three steps back: The stormy history of reading comprehension assessment. *The Clearinghouse, 72*, 97-105.
- Schmoker, M. (1999). *Results: The key to continuous school improvement (2<sup>nd</sup> ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Shen, J. (2003). *Have minority students had a fair share of quality teachers? Results from a national longitudinal study (1987-88 to 1999-2000)*. Poverty and Race

Research Action Council: Washington, DC.

- Shumow, L. (1997). Daily experiences and adjustment of gifted low-income urban children at home and school. *Roeper Review*, 20(1), 35-39.
- Siegel, H. (1990). McPeck, informal logic, and the nature of critical thinking. In J.E. McPeck (Ed.), *Teaching critical thinking* (pp. 75-85). London: Routledge.
- Slavin, R.E. (1986). *Ability grouping and student achievement in elementary schools: A best-evidence synthesis*. Baltimore: Center for Research on Elementary and Middle Schools.
- Slavin, R.E. (1990). *Cooperative Learning: Theory, Research, and Practice*. Englewood Cliffs, NJ: Prentice Hall.
- Sleeter, C.E. (1990). Staff development for desegregated schooling. *Phi Delta Kappan*, 72(1), 33-40.
- Slocumb, P.D., & Payne, R.K. (2000). *Removing the mask: Giftedness in poverty*. Highlands, TX: aha! Process.
- Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Stallings, J., & Kaskowitz, D. (1974). *Follow through classroom observation evaluation, 1972-1973*. Menlo Park, CA: Stanford Research Institute.
- Stephens, E., & Brown, J. (2005). *A handbook of content literacy strategies: 125 practical reading and writing ideas (2<sup>nd</sup> ed.)*. Norwood, MA: Christopher-Gordon Publishers, Inc.
- Sternberg, R.G., & Bhana, K. (1986). Synthesis of research on effectiveness of intellectual skills programs: Snake-oil remedies or miracle cures? *Educational*

*Leadership*, 44(2), 60-67.

Stringfield, S., Milsap, M.A., & Herman, R. (1997). *Urban and suburban/rural strategies for educating disadvantaged children*. Washington, DC: Planning & Evaluation Service, U.S. Department of Education.

Swan, W.W. (2004). *Impact of Ruby Payne's Instructional Framework on student achievement in East Allen County Schools, Indiana 2001-03*. Athens, GA: University of Georgia.

Swanson, J. (2006). Breaking through assumptions about low-income, minority gifted students. *Gifted Child Quarterly* 50(1), 11-24.

Swiatek, M.A. (1993). A decade of longitudinal research on academic acceleration through the study of mathematically precocious youth. *Roeper Review*, 15, 120-123.

Swiatek, M.A., & Benbow, C.P. (1991). Ten-year longitudinal follow-up of ability-matched accelerated and unaccelerated gifted students. *Journal of Educational Psychology*, 83, 528-538.

Taba, H. (1962). *Curriculum development: Theory and practice*. New York: Harcourt, Brace, & World.

Taylor, B.M., Pearson, P.D., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: Lessons about primary grade reading instruction in low-income schools. *Elementary School Journal*, 101, 121-166.

Taylor, B.M., Pearson, P.D., Peterson, D., & Rodriguez, M.C. (2002). Looking inside classrooms: Reflecting on the "how" as well as the "what" in effective reading instruction. *The Reading Teacher*, 56(3), 270-279.

- Taylor, B.M., Pearson, P.D., Peterson, D.S., & Rodriguez, M.C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, *104*(1), 3-30.
- Tivnan, T., & Hemphill, L. (2005). Comparing four literacy reform models in high-poverty schools: Patterns of first grade achievement. *Elementary School Journal*, *105*, 419-443.
- Tomlinson, C.A. (1995). Deciding to differentiate instruction in middle school: One school's journey. *Gifted Child Quarterly*, *39*(2), 77-87.
- Tomlinson, C.A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C.A., & Allen, S.D. (2000). *Leadership for differentiating schools & classrooms*. ASCD: Alexandria, VA.
- Tomlinson, C.A., Moon, T. R., and Callahan, C. M. (1998). How well are we addressing academic diversity in the middle school? *Middle School Journal*, *29*(3), 3-11.
- United States Department of Education. (1993). *National excellence: A case for developing America's talent*. Washington, DC: Office of Education Research and Improvement.
- United States Department of Education (2006). *Title I part a*. Retrieved January 22, 2007 from <http://www.ed.gov/programs/titleiparta/index.html>
- VanTassel-Baska, J. (Ed.). (2003a). *Excellence in educating gifted and talented*

- learners* (3<sup>rd</sup> ed.), Denver: Love.
- VanTassel-Baska, J. (2003b). *Curriculum for at-risk learners*. Storrs, CT: Monograph for The National Research Center on Gifted and Talented.
- VanTassel-Baska, J., & Bracken, B. (2005, November). *Findings on Project Athena*. Paper presented at the National Association for Gifted Children, Louisville, KY.
- VanTassel-Baska, J., Bracken, B., Feng, A., & Brown, E. (In press). A longitudinal study of reading comprehension and reasoning ability of students in Title I elementary schools. *Learning and Individual Differences*.
- VanTassel-Baska, J., Johnson, D., & Avery, L.D. (2002). Using performance tasks in the identification of economically disadvantaged and minority gifted learners: Findings from Project STAR. *Gifted Child Quarterly*, 46, 110-123.
- VanTassel-Baska, J., Johnson, D., & Boyce, L. (Eds.). (1996). *Developing verbal talent*. Boston: Allyn & Bacon.
- VanTassel-Baska, J., Johnson, D., Hughes, C., Boyce, L. (1996). A study of language arts curriculum effectiveness with gifted learners. *Journal for the Education of the Gifted*, 19, 461-480.
- VanTassel-Baska, J., Little, C., Rogers, K., Feng, A.X., & Drummond, D. (2002, May). *Social studies curriculum effectiveness study*. Paper presented at the Sixth Biennial Henry B & Jocelyn Wallace National Research Symposium on Talent Development, Iowa City, IA.
- VanTassel-Baska, J., & Stambaugh, T. (2006). *Comprehensive curriculum for gifted learners* (3<sup>rd</sup> ed.). Boston, MA: Allyn and Bacon.

- VanTassel-Baska, J., & Stambaugh, T. (2007) *Overlooked gems: A national perspective on low-income promising learners*. Washington, DC: National Association for Gifted Children.
- VanTassel-Baska, J., Stambaugh, T., & French, H. (2004). *Jacob's Ladder Reading Comprehension Program*. Williamsburg, VA: Center for Gifted Education.
- VanTassel-Baska, J., Zuo, L., Avery, L.D., & Little, C.A. (2002). A curriculum study of gifted student learning in language arts. *Gifted Child Quarterly*, 46, 30-44.
- Villaume, S.K., & Brabham, E.G. (2002). Comprehension instruction: Beyond strategies. *The Reading Teacher*, 55(7), 672-676.
- Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. (1986). *Thought and language*. (A. Kozulin, Trans.). Cambridge, MA: The MIT Press.
- Webb, M. (1998). Culture: A view toward the unexplored frontier. In M.E. Dilworth (Ed.), *Being responsive to cultural differences: How teachers learn*. (pp.61-77). Thousand Oaks, CA: Corwin Press.
- Weinfurt, K. (1995). Multivariate analysis of variance. In L.G. Grimm & P.R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp.245-276). Washington, D.C.: American Psychological Association.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performances. *Education Policy Analysis Archives*, 10(12).
- Westberg, K.L., Archambault, F.X., Jr., Dobyms, S.M., & Salvin, T. (1993). *An*

*observational study of instructional and curricular practices used with gifted and talented students in regular classrooms* (RM93104). Storrs, CT: University of Connecticut, The National Research Center on the Gifted and Talented.

Westberg, K.L., & Daoust, M.E. (2003, Fall). The results of the replication of the classroom practices survey replication in two states. Storrs, CT: *National Research Center on the Gifted and Talented Newsletter*.

Wharton-McDonald, R., Pressley, M. & Hampston, J.M. (1998). Literacy Instruction in Nine First-Grade Classrooms: Teacher Characteristics and Student Achievement *The Elementary School Journal*, 99(2), 101-128.

Willerman, M., & MacHarg, M. (1991). The concept map as an advance organizer. *Journal of Research in Science Teaching*, 28(8), 705-712.

Zemelman, S., Daniels, H., & Hyde, A. (1998). *Best practice: New standards for teaching and learning in America's schools* (2<sup>nd</sup> ed.). Portsmouth, NH: Heinemann.