

A COMPARISON OF EARLY READING OUTCOMES AND PROGRAM
COSTS IN FOUR PRIMARY READING PROGRAMS
FOR IMPROVED DECISION-MAKING

by

BRIAN DAVID GANDER

A DISSERTATION

Presented to the Department of Educational Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

December 2007

“A Comparison of Early Reading Outcomes and Program Costs in Four Primary Reading Programs for Improved Decision-Making,” a dissertation prepared by Brian David Gander in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Leadership. This dissertation has been approved and accepted

Dr. Gerald Tindal, Chair of the Examining Committee

NOV 21, 2007

Date

Committee in Charge: Dr. Gerald Tindal, Chair
 Dr. David Conley
 Dr. Keith Hollenbeck
 Dr. Jean Stockard

Accepted by:

✓

Dean of the Graduate School

CURRICULUM VITAE

NAME OF AUTHOR: Brian David Gander

PLACE OF BIRTH: Longview, Washington

DATE OF BIRTH: January 12, 1959

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
San José State University
Fresno State University
Oregon State University

DEGREES AWARDED:

Doctor of Education, Educational Leadership, 2007, University of Oregon
Master of Education, Health Education, 1989, Oregon State University
Bachelor of Science, Physical Education, 1983, Oregon State University
Bachelor of Science, Health Education, 1982, Oregon State University

AREAS OF SPECIAL INTEREST:

Educational Programs Evaluation
Cost Containment in Public Education

PROFESSIONAL EXPERIENCE:

Superintendent, Long Creek School District #17
July 1, 2007 to date

Principal, Salem-Keizer School District #24
August 2000 to June 2007

Principal, Campbell Union School District
October 1999 to June 2000

Assistant Principal, Campbell Union School District
August 1996 to October 1999

Assistant Principal, Dos Palos Oro Loma Joint Unified School District
August 1995 to June 1996

Physical Education Teacher, Dos Palos Oro Loma Joint Unified School District
August 1993 to June 1995

Health and Physical Education Teacher, Corvallis School District 509J
August 1989 to June 1993

Graduate Teaching Assistant, Department of Public Health,
Oregon State University
August 1987 to June 1989

Physical Education Teacher, The Columbus School, Medellin, Colombia, S.A.
July 1985 to June 1987

GRANTS, AWARDS AND HONORS:

President, Western Region Oregon Elementary Schools Principals Association.
2005-2007.

“Supporting Parental Literacy Skill Development” Community Foundation of
Silicon Valley. Lynhaven Elementary, Campbell Union School District,
Campbell, CA, 2000.

“Developing Video Production as Feedback for Improving Student
Presentations,” California Department of Education #1510 Grant for
Technology Implementation, Rolling Hills Middle School, 1998.

PUBLICATIONS:

Gander, B. D. (1982). Professional preparation: A student’s perspective. *Oregon Alliance of Health, Physical Education, Recreation and Dance Journal*, May 1982, 19-20.

ACKNOWLEDGMENTS

I wish to express sincere appreciation to Drs. Tindal and Hollenbeck for their support and dedication to teaching during the preparation of this manuscript. I would also like to thank the members of my committee, Drs. Conley and Stockard for their insights and support during this project. Their perspectives were helpful in making this study a reality. In addition, special thanks to Dr. Margaret Smith, whose advice and guidance many years ago put me in a position to achieve this degree.

To Sarah, Margaret, Matt, and Louise –
Chasing a dream takes a team and you are my dream team!

TABLE OF CONTENTS

Chapter	Page
I. STATEMENT OF THE PROBLEM	1
Research Questions	2
Introduction	3
Differing Perspectives of Reading Research	4
Constituency Concerns about Wasted Public Resources.....	6
Local, State and National Policy Pressures	8
Increasing Complexity of Student Demographics	9
Theoretical Framework	14
Reading Implementations Approaches	15
Explicit Instruction	15
Intensity of Instruction	16
Comprehensive Instruction	17
Socially and Emotionally Supportive Strategies	18
Cost Analysis	19
Distributed Leadership for School Decision-Making.....	21
Purpose of Study	24
II. LITERATURE REVIEW	26
Characteristics of Reading Programs that Support Early Grade Readers	28
Elements of Explicit Instruction	29
Elements of Intensity of Instruction	30
Elements of a Comprehensive Reading Program	32
Elements of Socially and Emotionally Supportive Strategies	34
Examples of Reform Models that Utilize the Elements of Effective Reading Programs	37
<i>Success For All</i>	37
<i>Success For All</i> as a Reform Model.....	38
<i>Success For All</i> Reading Curriculum	39

Chapter	Page
<i>Success For All</i> Student Groupings	40
<i>Success For All</i> Reading Tutors	40
<i>Success For All</i> Student Assessment	41
<i>Reading First</i>	42
<i>Reading First</i> as a Reform Model.....	43
<i>Reading First</i> Reading Curriculum.....	43
<i>Reading First</i> Student Assessment	44
Reading Initiatives Summary.....	45
Approaches to the Analysis of Educational Program Costs.....	45
Cost-Effectiveness Analysis	45
Ingredients Methodology	46
Costs	47
Discounting Costs.....	48
Sensitivity Analysis.....	48
Expenditure Review	49
Summary	50
 III. METHODS	 53
Participants	54
School Demographics and Reading Program Descriptions	56
Hamilton Elementary	58
Franklin Elementary	59
Adams Elementary.....	63
Jefferson Elementary	65
Measures.....	69
Reading Outcomes.....	69
School Decision-Making Practices.....	70
Cost-Analysis Measurement.....	71
Data Collection.....	72
Demographic and Reading Outcome Data	73
Leadership and Program Variables Data	74

Chapter	Page
Cost-Analysis Data	75
Assumptions about Subjects' Interactions with Three Reading Programs.....	76
Data Analysis	77
Summary	81
 IV. RESULTS	 82
Descriptive Statistics	82
Research Question 1	84
First Grade Demographics	84
Second Grade Demographics.....	86
First Grade Reading Outcomes.....	87
Second Grade Reading Outcomes	92
Research Question 1 Summary	97
Research Question 2.....	98
Resource Costs Data	98
Total Costs.....	103
Research Question 2 Summary	108
Research Question 3	109
Program Managers.....	110
Principals	112
Teachers and the Decision-Making Process	113
Leadership and the Decision-Making Process.....	119
Research Question 3 Summary	121
Research Question 4	122
Program Managers	123
Principals	125
Research Question 4 Summary	126
Research Question 5	127
Research Question 5 Summary	131

Chapter	Page
V. DISCUSSION	133
Research Question 1	135
First Grade	136
Span of Instructional Materials and Planning	138
Opportunity to Participate	139
Number of Days per Week.....	139
Second Grade.....	140
Research Question 1 Summary	143
Research Question 2.....	145
Program Ingredients Costs.....	145
Ingredient Cost Variables.....	147
Cost-Effectiveness	149
Cost-Effectiveness Ratios	149
Research Question 2 Summary	151
Research Question 3	153
School Leadership Model	153
School Personnel Responses.....	154
Research Question 3 Summary	157
Research Question 4.....	158
Research Question 4 Summary	162
Research Question 5	165
Research Question 5 Summary	169
Implications of Findings	171
Limitations	172
Conclusions	173
APPENDICES	175
A. PRINCIPAL AND PROGRAM MANAGER INTERVIEW QUESTIONS FOR UNDERSTANDING FOUR PRIMARY GRADE READING PROGRAMS	175
B. TEACHER SURVEY QUESTIONS FOR UNDERSTANDING FOUR PRIMARY GRADE READING PROGRAMS	178
REFERENCES.....	182

LIST OF TABLES

Table	Page
1.0 Percentage of U. S. Public School Students (ages 5-17) by Ethnic Distribution.....	10
2.0 Percentage of United States Public School Students (ages 5-17) Whose Home Language is Not English 1989-2003	11
3.0 School Demographic Characteristics	55
4.0 Grade 1 Demographic Information for Sampled Students in Four Participating Schools.....	56
5.0 Grade 2 Demographic Information for Sampled Students in Four Participating Schools	57
6.0 Reading Programs Characteristics in Subject Schools	60
7.0 Ingredients Reviewed in Four Primary Grade Reading Programs	75
8.0 Grade 1 Demographic Information for Sampled Students in the Four Participating Schools	85
9.0 Grade 2 Demographic Information for Sampled Students in the Four Participating Schools	88
10.0 Fall to Spring Primary Literacy Assessment Mean Scores for Grade 1 Students in Four Primary Reading Programs	89
11.0 Fall to Spring Primary Literacy Assessment Mean Scores for Grade 1 Students in Four Primary Reading Programs.....	91
12.0 Fall to Spring Descriptive Statistics Mean Scores for Grade 2 Students Reading in Four Primary Reading Programs	94
13.0 Fall to Spring Primary Literacy Assessment Mean Scores for Grade 2 Students in Four Primary Reading Programs	95
14.0 Program Costs Variables by Resource Type in Four Primary Reading Programs.....	99

Table	Page
15.0 Comparative of Total Cost of Four Primary Reading Programs Resource Cost Variables	103
16.0 First Grade Means Days Attended for Cost per Day of Program Implementation.....	104
17.0 Second Grade Mean Days Attended for Cost per Day of Program Implementation.....	105
18.0 First Grade Cost-Effectiveness Ratios for Students Meeting District Reading Benchmark	106
19.0 Second Grade Cost-Effectiveness Ratios for Students Meeting District Benchmark	107
20.0 Program Manager and Principal Expressed Knowledge of Cost- Effectiveness Concepts	111
21.0 Teacher Awareness of and Interaction with Cost-Effectiveness Measures.....	114
22.0 Frequency of Cost Effectiveness Variables Discussions at School Sites as Recalled by Program Managers	116
23.0 Frequency of Cost Effectiveness Variables Discussions at School Sites as Recalled by Principals	117
24.0 School Community Descriptive Variables Associated with Student Complexity Variables	124
25.0 Teacher Familiarity and Working Experience with Current Reading Program Materials	129
26.0 Program Manager and Principal Perceptions of School Program Fidelity Variables	130

CHAPTER I

STATEMENT OF THE PROBLEM

Schools and school leadership face an ongoing set of accountability and policy challenges. Schools are expected to meet the accountability expectations of district, state, and national policies requiring all students to achieve grade level academic proficiency by legislatively selected dates (Hummel-Rossi & Ashdown, 2002; Jefferson, 2005). Concomitant with these pressures are the pressures for leadership styles that distribute authority for schoolwide decisions amongst the practitioners and community stakeholders (Goertz, & Duffy, 1999; Kruse, 2001; Picus, 1999). Lastly, these distributed decision-making practices and the improved academic outcomes are to be achieved in a climate of uncertain funding, changing student and teacher demographics, and the increased pressure of public financial accountability (Brent, Sipple, Killeen, & Wischnowski, 2004; Hummel-Rossi & Ashdown, 2002).

Faced with claims of research-based interventions, locally designed instructional formats, and status quo circumstances, school principals and practitioners must maneuver a maze of program claims to find the appropriate instructional approaches for their communities' needs. In light of competing claims and rising community expectations, school administrators are looking for effective practices to reduce the

students at risk of missing grade level standards in a variety of academic areas, and reduce risk of not meeting standards in a manner that is cost conscious while meeting the needs of the professionals they work alongside (Allington, 2006; Lyon & Chhabra, 2004).

One means of evaluating the effectiveness of instruction and student interventions is through the comparisons of alternative programs (Levin, 1983). This requires three interrelated processes: (a) measuring student outcomes for each of the alternatives with a similar set of research defined variables, (b) delineating comparative variables that characterize the nature of alternatives under review, and (c) the analysis of program costs for the use in developing cost-effectiveness ratios.

Using those three processes, the goal of this research is to provide insight into a pragmatic framework that practitioners can utilize to assure they understand both the visible and less apparent implications of reading program implementations. By developing an understanding of the role that primary reading skill outcome measures and cost analysis tools play in the decision-making process, this study hopes to inform instructional leaders about how to efficiently utilize limited resources. In making program cost analysis more specific to student outcomes, the goal is to assure that students receive the most effective instructional programs as measured by attainment of district, state, or commonly agreed upon early reading standards.

Research Questions

- 1) Which schools' primary reading program had significantly better scores on each of the five Rigby Primary Literacy Assessment measures: a) letter identification, b) letter

sound identification, c) word recognition, d) reading comprehension, e) reading fluency and f) book level attained?

2) Which of the four reading programs had better cost-effectiveness ratios (cost of students reaching district benchmarks) of four reading programs when comparative groups of first and second grade students' early reading outcomes are used to evaluate all four programs?

3) What do schools' teachers, program managers and principals use as their decision-making process when programs are implemented?

4) How does the complexity of the school's student population affect the nature of the challenges faced by the leadership teams when evaluating student outcomes?

5) Do site administrators and program managers feel that the reading programs in place are being utilized within the design parameters of the developer leading to reading outcomes that represent a consistent application of the reforms intentions?

Introduction

Over the past 15 years a variety of school accountability concerns have forced school personnel to reevaluate educational program priorities. The accountability influences range from a lack of agreement on what the research recommends for effective reading instructions, to constituency discontent about inefficient use of tax revenues for school reform practices, and the implementation of local and national policy initiatives. These three policy pressures to increase the pace of improvement in student early reading outcomes while containing cost variables shape the decision-making process of district and school leadership teams (Allington, 2006; Brent, Sipple, Killeen, & Wischnowski,

2004; Jefferson, 2005). The nature of these pressures and how districts and schools respond to them require a diverse set of skills in program evaluation and the use of data by decision-makers to guide instruction.

Differing Perspectives of Reading Research

Reading professionals with differing perspectives on approaches to reading instruction articulate their positions on how best to meet various proficiency goals (Allington 2004; Lyon & Chhabra, 2004). The two groups (one favoring direct instructional protocols and the other favoring a balanced instructional approach) present arguments using reading outcomes, the nature of reading research, reading interventions, and, to a limited extent, projected costs to support their claims.

Lyon and Chhabra (2004) represented a view that believed it is possible (predominately through one-to-one intervention models) to get 90 to 95% of all students reading at grade level. Current research by these authors suggested that “the identification of children at-risk for reading failure coupled with the provision of systematic, comprehensive, and evidenced based interventions can reduce the number of students reading below basic levels to less than 6%” (p. 16). Rayner, Foorman, Perfetti, Pesetsky, and Seidenberg, (2001) in a review of how recent research has helped inform reading instruction also felt educators can achieve higher rates of success than is currently the case. Both Lyon and Chhabra, and Rayner et al., rely on the potential of students with a wide range of abilities to make the phoneme to grapheme connection as the base of instruction. Once these internal language codes are well developed, students can begin to learn strategies during the reading process to improve their understanding of the

phoneme-grapheme code and in the end progress through developmental levels to become effective readers. Thus, using a teachable representation of the phoneme and grapheme variables (represented as explicit phonics instruction), all but a very few of our school age children can become productive readers (Rayner et al., 2001).

A second perspective challenges the claims of those who say we have the methods and the finances to support all students reading on grade level by third grade (Allington, 2006, 2004). The effectiveness of one-to-one tutoring is not being challenged. However, the cost variables related to one-to-one intervention models seeking to reduce to less than 10% those students at risk of reading failure is in question. Researchers challenged those who looked for research-developed intervention models that will bring 90 to 95% of the struggling readers to grade level proficiency in cost contained programs (Allington, 2006, 2004; Krashen 2004). For many schools and districts under current economic conditions it is an unrealistic challenge for school personnel to bring all students (with the exception of special education students) scoring below the 20th percentile on accurately normed reading assessments up to the 45th percentile in effective and cost conscious programs. Careful reading of their argument shows that in a school of 400 students who are normally distributed in reading abilities, the lowest 20% accounts for 80 students (Allington, 2004; Krashen 2004). Finding interventions for 80 students that are effective, economically realistic, and that maintain a sense of order within the school may be a difficult task for school personnel.

The purpose here is not to prove either of these positions superior to the other. However, the competing perspectives present the situational context in which school-

based practitioners find themselves. Leadership of a school site, or leading a curriculum department through this level of dissonance, requires skills in program evaluation ranging from cost-analysis to defining effective reading assessment protocols for outcome measurement, and understanding the dynamics of school and district leadership.

Concerns about wasted resources come when these elements do not mesh well and the public sees schools changing from one approach to the next with less than thorough study of all facets of the problems involved.

Constituency Concerns about Wasted Public Resources

An ongoing concern within communities to change current practice arises from the conflicts about the appropriate uses of public resources (Jefferson, 2005). This is seen in the policy discussions about whether constituents within the United States fund education at a level adequate for all students' success (Flanigan, Marion & Richardson, 1997; Hedges, Laine, & Greenwald, 1994; Ilon & Normore, 2006). From the late 1980s to the most current editions of education finance journals, articles consistently establish qualitative and quantitative evidence for both the pros and cons of school funding adequacy.

Research throughout the 1980s, 1990s, and into this decade have attempted to define a variety of policy initiatives in economic terms (Hedges, Laine & Greenwald, 1994; Hummel-Rossi & Ashdown, 2002). Theoretical models on education programs production functions (Hedges, Laine & Greenwald, 1994; Marion and Flanigan 2001; Monk, 1989), along with reviews of the cost-effectiveness of reducing primary grade class sizes, and research on whether judicial interventions to level funding inequities in

state budgeting statutes, have all been part of the literature on the use of public funds and their adequacy. Oregon in the 1990s and early 2000s investigated the use of a Quality Education Model (QEM) to establish a base adequacy funding measure to guide state school funding policy. These developments in a production function approach is one of a variety of responses that states have pursued in an attempt to quantify for its constituency a base amount upon which to plan revenue expenditures, and then link that amount to student outcomes (Conley & Picus, 2003). These attempts to link policy initiatives, student outcomes, and the reduction of wasteful spending serve to increase the pressure on school-level decision makers to find and implement programs that are effective and cost conscious.

Therefore, associated with the call for better instructional approaches is a call for using our public resources effectively. In 2002, Levin estimated that the total expenditures in the areas of elementary, secondary, higher education, and educational training accounted for \$750 billion (Levin & McEwan, 2002). One can see how minor improvements in efficiencies may make a significant impact in the saving of public funds. A 5% improvement would net this sector of the economy \$37.5 billion. National pressures to improve school achievement, while at the same time monitoring the size of education budgets, presents a series of challenges to practitioners and researchers. One conclusion from the literature is that most research on the implementation of intervention programs presents the outcomes for students but lacks clear analysis of the costs. If educators are to make more informed decisions about the outcomes of instructional change, more complete information is needed on how expenditures and outcomes are

linked together. Chambers (1999) and Hummel-Rossi and Ashdown (2002) proposed that educational researchers link student outcomes with cost analysis tools. To achieve these goals educators need to be versed in the use of cost analysis tools like cost-effectiveness analysis, cost-benefit analysis, cost-feasibility analysis, and cost-utility analysis.

Local, State and National Policy Pressures

Since President John F. Kennedy's announcement in 1961 that the United States was going to put a man on the moon before the end of the decade, public policy about local schools has been responding to outside agency pressures. This use of education as public policy and political issue nationalized the concerns and debate about the *how* and *what* of our students' education; specifically in the areas of math and science. Although early policy documents were guides for local communities, as state and federal actions have attached incentives and sanctions to policy initiatives, local entities have lost increasing levels of self-determination (Conley, 2003). In the current political arena the blending of: (a) accountability measures through assessments and performance standards, (b) guidance documents for elements from quality teachers to curriculum, and (c) federal funding protocols has placed increasing pressures on district and site leadership teams to improve student outcomes by what the public perceives as targeted dates.

National and local policy initiatives have required school leaders to place increasing instructional focus and resources on student reading outcomes (Allington 2004, 2006; Borman, 2005). The *No Child Left Behind Act of 2001* has set a benchmark that all students read at grade level by the year 2014 (NCLB, 2001). Subpart B of the NCLB 2001 legislation was enacted to form grant opportunities within each state to

establish support of the variables defined by the National Reading Panel in its report *Teaching Children to Read* (Barone, Hardman, & Taylor, 2006; NCLB, 2001; NRP, 2000). In the NCLB legislation the necessary elements of a research-based reading program are to be used as an attempt to assure that educationally sound reading programs reach those who rely on Title I program funding specifically and all students generally. This process is known as the *Reading First* initiative and is used by states and districts in schools whose children are not meeting the reading benchmarks as outlined by the states in their adequate yearly progress formulas and approved by the U. S. Department of Education. The *Reading First* initiative, linking state and district policy initiatives and their academic and budgetary accountability standards, have in turn created a scramble within schools to find academic interventions that will meet the needs of all students. Specifically, district are looking more closely at the educational services provided to those students at risk for reading and math failure, while at the same time remaining cognizant of limited school budgets.

Increasing Complexity of Student Demographics

Shifting demographic characteristics with the United States school-age population have placed increased pressure on educators to approach reading and language arts instruction with more diverse strategies (Allington, 2004, 2004; Krashen, 2006). The complexity is addressed in two main discussions: (a) in the nature of the complexity (gender, race, and socioeconomic status), and (b) delineating student complexity variables that contribute to the achievement gap.

In 1980, K-12 classrooms in U.S. public schools contained 40.8 million students, in which 27.8 % were non-White and of whom 12.3% spoke a language other than English in their home (NCES, 2005). In 2003 the ethnic and language distribution of the 48.5 million public school students looked different. The 2003 NCES data suggest a 14% increase in non-White K-8 students in public school in the United States, and a 6% increase in students who speak a language other than English in their home (NCES, 2005). Although a variety of policy and educational solutions have been offered to meet this demographic shift, few ready-made solutions have been transferable across a variety of situations (Borman, 2005).

Tables 1.0 and 2.0 represent recent K-12 student demographic trends in U.S. public schools. The changes in the percentages of students who are of minority status and the percentage of students who do not speak English as their home language have had an impact on the way American youth are educated. The methods of instruction and school

Table 1.0

Percentage of U. S. Public School Students (ages 5-17) by Ethnic Distribution

Race / Year	1981	1995	2000	2003
Black	16.0	16.9	16.6	16.1
Hispanic	8.7	14.1	16.6	18.6
Other	2.9	3.5	5.4	7.0
White	72.4	65.5	61.3	58.3

Source: NCES, 2005.

policy responses have produced significant effects on educators; but have shown only marginal changes on student outcomes (Allington, 2006; Borman, 2005).

Table 2.0

Percentage of United States Public School Students (ages 5-17) Whose Home Language is Not English 1989-2003

Year	% of U. S. 5-17 year olds	Students (in millions)	% who Speak English w/ Difficulty	Students (in millions)
1989	12.3	5.2	34.6	1.3
1999	16.7	8.8	29.5	2.6
2003	18.7	9.9	29.4	2.9

Note. *Difficulty* is scored as a 4 or below on a 5-point Likert scale, NCES, 2005.

When the complexity of student characteristics is taken into account in U.S. schools, a number of items go beyond ethnicity and primary language. Student mobility, behavior and attendance patterns, preschool experiences, access to technology, and economic diversity are characteristics that can make the teaching of students more challenging and in turn add perspective to the discussions of the achievement gap in public schools.

The achievement gap is characterized in the literature as the statistical difference between students of minority and the normative student group. A more accurate portrayal of achievement differences is presented by Walsh-Symonds (2003) who compared the

specific groups of Hispanics and African-American to Asians and Whites. In a view that is characteristic of educational communities, Walsh-Symonds disaggregated the data of large-scale and local assessments to present a more specific context of the achievement gap. The National Center for Educational Statistics (NCES, 2005) measures of reading showed Asian students performed markedly better than their Hispanic and African-American peers. On the National Assessment of Education Progress fourth grade reading tests 39% to 41% of the White and Asian/Pacific Islanders are proficient or advanced compared to 13% to 15% of Hispanic and African-American students. The more startling statistic is that in general the fourth grade gap continued to grow as students progressed through the U.S. educational systems (Walsh-Symonds, 2003).

The second social context of the achievement gap is the interaction of race and socioeconomic status of the participants. Schools with the widest achievement gaps commonly are those schools that serve populations that are overall less economically fortunate. These schools have the fewest resources yet policy makers anticipate them to make the largest gains since their students face the largest deficits (Darling-Hammond, 2004). These social contexts are present as an “opportunity to learn” construct in which it is felt that necessary elements must be in place to assist disadvantaged student in accessing learning (Barton, 2004, 2003; Walsh-Symonds, 2003).

Barton (2003, 2004) presented 14 characteristics that are most often associated with opportunity to learn variables and thus contribute to the achievement gap in U. S. schools. These characteristics are grouped into three sets of variables: (1) health of children prenatal and in early years, (2) nature of home life, and (3) nature of schooling.

Measures related to children's health: (a) low birth weight, (b) lead poisoning, (c) hunger, and (d) nutrition both in prenatal and preschool years. A second set of measures is related to parenting practices: (a) reading to young children, (b) television watching, (c) parent reading practices, (d) student mobility, (e) marital status, and (f) parent participation in school. The final set of variables is related to schooling: (a) rigor within the curriculum, (b) teacher experience, (c) teacher preparation, (d) class size, (e) technology-assisted instruction, and (f) school safety. This final set is seen as possible policy influenced variables to control opportunity to learn variables (Barton 2003, 2004). Although the first two sets of measures are difficult for schools to influence, their importance is of concern for educators who attempt to link social services with school service provision to affect academic outcomes. Systemic reform advocates and educational advocates in general sought means to improve the school related measures without much control over the first two categories.

Systemic reform initiatives attempted to generate standards for education based on these variables and at times integrated these standards into the home setting. Thus some reform initiatives support parent liaisons connecting the family and school for dealing with behavioral and attendance concerns, training resources for parents around early literacy skill practice, and a variety of community events that attempt to integrate families with their schools to soften the effects of cultural differences. (Darling-Hammond, 2004; Herman, Klein, & Abedi, 2000). These student complexity variables contribute to the challenges of diverse schools and are just one of the four factors effecting leadership team decision-making.

In conclusion, these four challenges individually and in combination present a source of anxiety for district and school level leadership. Therefore, it is important that leaders are able to differentiate educational programs that generate positive student outcomes, programs that are responsive to a variety of learners, and approaches that are scalable to a variety of school sites. To accomplish these goals, decision-makers must be able to compare like outcomes and determine cost variables in a manner that all costs are understood. With these three sets of program information, leaders can make the appropriate choice for their locality. Hopefully the recommendations and subsequent decisions will have their genesis in a greater degree of certainty and succeed in supporting students' growth, helping staff as service providers, and reduce some of the pressures in the current educational policy environment.

Theoretical Framework

The context for viewing the challenges of supporting readers at-risk of reading failure, the measurement of costs related to that support, and the nature of school leadership is represented in three distinct theoretical perspectives. First, there exists a rich field of study examining the characteristics of effective reading instruction, effective reading interventions, and the measurement of the variables that signify reading success in early grades (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998). Secondly, a significant amount of current research addresses the context of professional learning communities with a focus on distributed leadership structures. Distributed leadership, as viewed by Gronn (2000), utilizes the elements of activity theory to summarize the *how* and *what* of school leadership actions. Finally, a research strand seeking to identify

production functions that support educational program evaluations through cost analysis has also contributed insight (Hummel-Rossi & Ashdown, 2002). The use of cost analysis in the development of educational policies that serve students and the public treasury has provided a consistent body of literature on a variety of cost analysis techniques (Jefferson, 2005; Marion & Flanigan, 2001; Verstegen & King, 1998).

Reading Implementation Approaches

The instructional approaches for students at risk of reading failure have become the source of many articles in education's professional journals (Foorman et al., 1998; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). The most prominent strand of research to come out of the National Reading Panel Report of 2000 (NRP, 2000) is the focus on research-based theories around five main early grade reading components: (a) vocabulary, (b) phonemic awareness, (c) phonics, (d) reading fluency, and (e) reading comprehension. The proliferation of articles defining the critical elements of effective reading instruction and intervention programs can be distilled into four broad categories: (a) explicit instructional strategies, (b) comprehensive literacy tasks, (c) intensity of instructional formats, and (d) social emotional variables for student support (Foorman & Torgesen 2001; Pikulski, 1994; Shanahan, 1998).

Explicit instruction. Explicit instructional practices are defined by Foorman and Torgesen (2001) and Pikulski (1994) as a teacher-directed process that shapes the learning environment to achieve specific outcomes. These focused lessons are built around the five variables recognized by the Nation Reading Panel (2000), beginning with the basic relationships of language at the primary level and proceeding through

vocabulary, reading fluency, and comprehension tasks. For kindergarten and grade one that means phonemic awareness and phonics development as the base for linking grapheme (letter representations) and phoneme representations (letter sounds) in specific lessons for students to literal, contextual, and cultural gain representation. As students get older the explicit instructional practices move to lessons about the way students think while reading, the metacognitive processes. Forman et al. (1998), and Spiegel (1995) found that more often the students from disadvantaged homes need to be taught ways of thinking about reading in scaffolding activities of increasing complexity. These two definitions of explicit, increasing complexity of instruction, and focused instruction within a developmental context, frame the construct of explicit instruction.

Intensity of instruction. Intensity of instruction as an organizing structure is developed around the belief that all students do not enter school with equal experiences in language development. Some students must catch-up with their peers by passing through a well recognized series of steps in language development (Forman et al., 1998; Rayner et al., 2001; Spiegel, 1995). Although the steps are not completely understood, some students need greater time on task with reading skills to allow them to access what grade level expectations agree upon as grade appropriate skills. The variables in the research that fall into this category are group sizing, time of instructional focus (minutes of supplemental instruction), and the length of time (number of days of supplemental instruction) when looking at interventions (Pikulski, 1994; Shanahan, 1998). Each of these variables is important in understanding how students catch up to grade level expectations when they enter a program or classroom behind the normative group.

Comprehensive instruction. Comprehensive instruction as a construct looks at the variables that support reading lesson development: (a) assessment of student abilities, (b) assurance that all five NRP variables are included throughout the instruction cycle, and (c) use of a variety of instructional strategies (Hiebert & Taylor, 2000; Rayner et al. 2001; Taylor, Pearson, Clark, & Walpole, 1999).

Therefore, explicit lessons are assessed and the results used for guiding the next lesson (Pinnell et al., 1994); student grouping formats assure that they link classroom instruction with small group instruction (Pikulski, 1994), and lessons link the grapheme system of writing with the phonemic system of reading (Foorman et al., 1998; Rayner et al., 2001). Together these examples are pieces that make the isolated use of any one process more effective if bundled together to create a more balanced approach.

In comprehensive instructional settings consistent feedback is given to students in both summative and formative formats (Leslie & Allen, 1999). These assessment structures are used during one-to-one, small group, and large instructional formats that utilize a variety of instructional strategies based on student needs. Programs that are effective evaluate their programs and individual progress periodically as the year progresses to assure a match with instruction and student need (Monk, 1995; Slavin & Madden, 2001a). The comprehensive instruction element asks program developers to view programs from a wider lens, one that takes instructional outcomes aligned with instructional processes and program evaluation practices, to assure that all aspects of instruction are covered for the wide range of student abilities in each classroom.

Socially and emotionally supportive strategies. The socially and emotionally supportive strategies construct is the final critical element of the four that explains reading programs and interventions. This construct takes in the variables that do not fit tightly into the structures of intensity, explicit process, and comprehensive instructional practices. As a group these variables round out the teaching process and link the skills of reading to the student's effectiveness.

Effective programs recognize the primary-age child as having a specific set of needs in the instructional setting, needs that differ based on previous academic experience, cultural influences, language experience, and support within the home and family. Effective teachers recognize that the adult-to-student relationship should adjust to the students needs first and the programs needs second (Foorman & Torgesen, 2001; Pinnell et al., 1994; Spiegel, 1995). In this context teachers assure that instruction outside of class aligns with in-class formats, adult-to-student interactions are overwhelmingly positive, and early grades support is the emphasis for intervention.

The use of these four categories of critical elements as presented in the literature supports the alignment of program review variables used by researchers in program evaluation. Looking for the specific instructional activities, materials, time frames, teacher behaviors, and outcomes allows for the organizing of program characteristics data and student outcomes while measuring the five variables outlined by the National Reading Panel (2000). As an organizing structure, the four categories under the umbrella of critical elements of reading instruction and intervention also serve to organize the large body of literature about appropriate reading instruction practices: some with a research

base, some as practitioner reports, and others as proprietary programs. Using a publisher's assessment protocol measuring these five instructional elements, this study will link research-defined outcomes with program characteristics, allowing for a second phase of this research, connecting the cost of four reading programs for program comparisons.

Cost Analysis

A number of researchers have linked the elements of program evaluation with student outcomes for purposes of funding local programs, state systems (Chambers, 1999; Conley & Picus, 2003; Jefferson, 2005), and the comparison of alternatives (Levin & McEwan, 2001, 2002). Each approach used some type of comparative system to generate estimated costs in variety of ingredients. The ingredients method described by King-Rice (1997) and first modeled by Levin (1983) has been expanded by researchers over the past 25 years. Although there is not a unifying theoretic construct as seen in other areas of research, program effectiveness in the context of generating the highest level of targeted outcome for the least amount of inputs could be summarized as a systemic accountability model. A method for evaluating the costs for this type of model is called the *ingredient method of resource costs method*.

The ingredients methodology (as well as resource cost method) takes into account five major cost *ingredients* starting with personnel (typically the greatest expenses in school systems), facilities, materials and equipment, unique inputs, and inputs by the organization (King-Rice, 1997; Chambers, 1999). Using this framework for costs in various programs, a view of scalability attributes and an accounting estimation of

anticipated expenditure structures can be developed. As a methodology ingredient cost outcomes in combination with academic outcomes provide a categorization of the variables that go into comparable units, permitting the evaluation of programs for their effectiveness. The system accountability model can then use cost-effectiveness ratios to project cost for program planned on a broader basis.

Adding to the understanding of the budgeted outcomes and the academic outcomes requires a third set of variables be looked at in the process (Grissmer, 2002; Levin & McEwan, 2002). The constructs of opportunity cost and distributed costs as economic variables contributing to the economic outcome of program implementation are a qualitative component within the system accountability model. Understanding the concepts of indirect cost, changes in individual workload and impingements on other curricular areas are not well defined by the budgetary or academic outcome. Qualitative factors used in defining the costs quantification for changes in work load and adjustments to new approaches of instruction are met with increasing complexity and are not well represented in a theoretical base (Grissmer, 2002).

Therefore, the second element of cost analysis, as originally presented by Levin (1983), used a comparison of alternatives based on a common metric. Cost-effectiveness analysis relies on the comparison of like outcome variables for programs with differing internal structures, materials, and strategies for meeting outcomes. The common metric gives outside viewers an opportunity to take costs, outcomes, and situational context into account when making decisions about program implementation in the school community. Cost analysis as professed in the literature is not to be the sole criterion upon which

school program decisions are to rest; however it has the potential to inform the decision-making process of both staff and communities in the areas of potential academic gains, hidden challenges, and long range viability of a reform decision.

Distributed Leadership for School Decision-Making

Gronn's (2000) combination of institutional leadership theory (Schwandt, 1995; Weick, 1975) in conjunction with activity theory explained the leadership behavior in the program development and implementation process. Institutional leadership theorists view leadership as situated in sense-making, or the thinking and actions of the actors, framed by the institution's norms, rules, and definitions of the environment (Schwandt, 1995; Weick, 1975). The institution's norms and structures both enable and constrain opportunities of the actors. Institutional theories, therefore, focus on the rational behaviors of the actors within sectors of the organization, placing an emphasis on the embedded nature of the decision-making. They view actors making rational decisions based in shared norms (organizational forces) at work during the decision-making process (Weick, 1975). The emphasis on organizational structure, shared belief systems, and situational responses has contributed a great deal to the development of distributed leadership theories and fits well with distributed leadership grounding within activity theory (Gronn, 2000).

Distributed leadership assimilated many of the conceptual strands of institutional (organizational) theorists. However, where institutional theorists place considerable emphasis on organizational structures, their forces on actors, and the embedded context of decision-making, not enough attention is paid to the actual distributed nature of

leadership practices or the artifacts within the organization (Gronn, 2000; Spillane et al. 2004). Gronn and Spillane et al. move away from an organizational theorist conception of actors making rational decisions solely for the benefit of the organization, and inject individual agency into the sense-making behavior of actors.

Distributed leadership theorists step back from the narrow focus on specific individual roles and look at leadership in a broader context, the social context of leading. Spillane et al. (2002, 2004) recognized the value of norms, rules, structures and mediation forces on behavior presented by institutional theorists. However, Spillane et al. (2004) do not agree with the idea of organizational forms having greater consideration over the influence than individual actors in shaping the organization. The distinction of distributed leadership is a move away from seeing leadership as determined by the organization or the leader-follower roles of the organization, and moves outward to a larger perspective (Gronn, 2000; Harris, 2003). Spillane (2004) hypothesized that knowing what constitutes leadership is essential, but unless the artifacts, the local setting, the technologies, local values, and local knowledge are taken into consideration the reason why leaders do what they do will not be understood. Theorizing that leadership is stretched over the social context, distributive theory moves past the idea of studying one individual (whether teacher leader or principal) and looks to the broader social and cultural activity and the artifacts that frame the decision-making for learning outcomes.

Placing distributive leadership on a continuum of leadership theories situates it in clustered set of structures including transformational leadership, organizational learning theory as well as theories fitting the definitions of facilitative and participatory theory

nested in institutional or organizational theory. At the other end of the spectrum would be transactional leadership forms and near the middle would be contingency theories. Gronn (2000), Spillane et al. (2001, 2004) and Harris (2003, 2004) saw the differences within the clustered group as one of research perspective. Distributive leadership research broadened the qualitative view to include thick descriptions of the leadership act. In the distributed perspective, research should focus increased emphasis on the influence of artifacts, structures, and technologies within the situated context of leadership behavior, with less emphasis on the traits of leaders and the form of the leadership situation.

In conclusion, distributed leadership literature is closely matched with the ideas of facilitative leadership as envisioned by Goldman and Conley (1994), and participatory leadership presented by Kruse (2001), Marks et al. (1999), and others. Distributed leadership is also closely aligned with concepts within the institutional perspective of school organizations. Distributed leadership theorists extend the literature frameworks of current thinking by viewing the leadership practices from a larger unit of analysis. They propose a unit of analysis that is situated in the leadership act, encompassing the social context, the artifacts and tools, and the interconnected nature of schools. In this perspective they seek to understand not just the *what* of the decision-making, but the *how* in the division of labor, the influence of structure, and the *why* of the leadership act, regardless of who is doing the leading.

It is believed that describing the nature of the decision-making process, the various skill sets of the individuals within the process, and the culture of the organization

that support from a proposed leadership framework for program evaluation inform the research.

Each research strand provides pragmatic and theoretical approaches to their topics that will support practitioners' understanding of: (a) what is important to look at in a program evaluation, and (b) how to design an evaluation of the alternative programs they are considering. Understanding the nature of primary grade reading programs, the characteristics of effective daily reading programs, and effective elements of intervention programs will help school-based leadership teams to make effective judgments. Associated with knowing theoretical elements of primary reading programs is understanding the costs related to alternative programs and existing barriers to program implementation. Understanding cost variables assists these same decision-makers make the best choices for their limited resources within the context of their community.

Purpose of Study

The primary purpose for this study is to link two research strands currently active in the educational literature. The pressures of cost containment in shrinking budget times, in association with pressures from local and national policy initiatives, have created increasing pressure on district and site leadership teams. Bringing cost analysis practices into the evaluation of current programs can help districts understand what they are expending for student outcomes and whether those programs are supporting students in meeting district and state benchmarks.

The review of four programs (two structured mainly by outside entities, and two locally designed) will create the variables necessary to evaluate programs. Once these

variables are apparent they can be placed into a framework for principals and district level administrators' use.

A secondary purpose is to see how teachers, program managers, and site administrators collect information, and use the information they have to make decisions based on their local circumstance. This is viewed in light of the complex nature of some of the schools involved and how closely they work within the confines of their specific program designs.

The rationale for the melding of qualitative and quantitative perspectives into a coherent picture of the relationship of student outcomes and programs costs variables is to provide administrators with a pragmatic approach to evaluating the potential impact of school reading programs on supporting student achievement towards grade level reading standards in a climate of increasing student diversity, pressures from outside policy initiatives, and shrinking budgets. The literature in reading outcomes is well developed and the literature about evaluating educational costs is also well established. However, few studies have taken concurrently running reading programs, applied similar reading outcomes, and attempted to compare cost-effectiveness ratios within a specific geographical location.

CHAPTER II

LITERATURE REVIEW

The recent work of the National Reading Panel (2000) served to accelerate reading instruction research (Lyon & Chhabra, 2004). The research of reading instruction can be found in both pragmatic models and theoretical constructs (Foorman, Breier, & Fletcher, 2003; Potter & Wamre, 1990). Of particular interest is the call from the National Reading Panel Report of 2000 and the No Child Left Behind Act of 2001 for the universal use of research-based reading practices. These research-based practices are to replace instructional approaches of the past four decades that have limited, if any, research basis as a foundation for school level implementation.

Reading instruction over the past 40 years has moved through a variety of philosophical and methodological approaches, each trying to define the specific elements that support children learning to read (Foorman, Breier, & Fletcher, 2003; Morrow & Gambrell, 2000; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). In general terms, students in American schools over the past 40 years have experienced structured basal reading programs of the 1960s and 1970s, the whole language perspective of the 1980s and early 1990s, or more recently a balanced literacy approach beginning in the early 1990s and currently in practice.

Today, U.S. public school students are experiencing schools' and teachers' responses to the findings of the National Reading Panel Report (NRP). These findings support an early reading instruction format based on direct as well as explicit instructional practices over the more recent use of approaches referred to as *whole language* and in more recent literature as *balanced literacy instruction*. In this approach, the NRP report on teaching early reading skills has outlined five areas of instruction that should be the focus of teachers in the primary classroom: (a) phonemic awareness, (b) phonics, (c) vocabulary, (d) reading fluency, and (e) reading comprehension.

Specifically the NRP report stressed that the basis of instruction in early grades should be phonemic awareness and phonics instruction presented in a systemic manner. As students get older and have the phoneme and grapheme relationships firmly in hand, sight word and vocabulary development takes over the focus of instruction. The concentration on vocabulary and sight words moves students to applying the skills of phonetic decoding as problem solvers during the reading process. As vocabularies and sight word repertoires increase, the shift from applying decoding strategies as a conscious act changes to a nearly automatic act. As the skill of rapid recognition broadens within the reader, reading fluency improves. The shift in focus from having to think about segmenting words to rapid recognition of whole words moves the reader into the final set of strategies aimed at increasing skills in rate of reading (fluency) and comprehension. Fluency and comprehension skills include training for the accurate recognition of words and the metacognitive processes of reflection, clarification and rereading. Although the two skills of fluency and comprehension are closely linked, strategies for understanding a

text's meaning through the use of syntax and expressive language are the focus of latter primary grade and early intermediate grade reading instruction.

Differing methodological perspectives approach these instructional variables in differing manners, yet each tries to outline methods for reaching all learners. What is common in the differing methodological approaches is the recognition of the need for remediation of students at risk of reading failure. Each methodological perspective has proposed intervention strategies to address at-risk students' needs. The effectiveness and prolonged adoption of a methodological perspective (whether its roots are in explicit instruction formats or whole language formats) will be determined by the rigor of its research base and its ability to meet the needs of all learners. That is to say, the strength of the theoretical perspective is measured by how well it supports all readers, especially those who need augmentation: the high readers and those at risk for reading failure. Therefore, pedagogical perspectives with strong empirical support for effectively teaching early reading skills to general education at-risk readers will provide more informed teacher instruction over time.

Characteristics of Reading Programs that Support Early Grade Readers

Instruction for children and specifically children at risk for reading failure is the primary challenge for classroom teachers in first and second grades (Berninger et al., 2003; Foorman et al., 2003). The challenges begin with the fact that children acquire skills at different rates than their peers, but they also must acquire the same set of skills to progress through the stages of learning to read. Therefore, students at risk of reading failure are not at all equal, but experience a common incongruence between the

instruction they receive and their current level of performance. This mismatch requires instruction for at-risk students that contains a variety of elements clustered under one of four organizing concepts: (a) explicit instruction, (b) intensive instruction, (c) comprehensive instruction, and (d) socially and emotionally supportive strategies (Foorman & Torgesen, 2001; Pikulski, 1994; Spiegel 1995). Therefore, when looking at instructional programs these four elements, as well as the outcomes they produce, must be part of the evaluation process undertaken by instructional leaders.

Elements of Explicit Instruction

Effective reading programs are characterized as explicit when they include a variety of elements ranging from direct instruction methodologies to teaching metacognitive skills. Spiegel (1995) found that children in low-income families and with minority backgrounds do not discover complex reading strategies without direct instruction. Through direct instruction practices, learners are shown in concrete activities what a strategy is and how to apply it. Interventions techniques using explicit methods include the scaffolding of concepts in literature and explicit modeling. Foorman et al. (1998) referred to two types of scaffolding: (a) instruction that incrementally refines skills, and (b) teacher-student conversations that clarify for the student the steps in thinking through a task.

Explicit instruction formats teach readers to transfer known strategies into new situations. By making the goals of instruction clear for the learner, the teacher uses direct instructional practices in lessons. These lessons include teaching phonemic awareness, phonics, and vocabulary, as well as practicing reading fluency and comprehension with

connected text, in contrast to the fragmented study of isolated skills. Instructional activities include strategies that are metacognitive: The learner knows a problem exists, identifies the problem, and puts into effect the learned strategy to get around the problem (Foorman & Torgesen 2001; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994).

Lastly, a beginning reading program's foundational elements should emphasize explicit instruction in phonemic awareness and phonics as part of the curriculum (National Reading Panel, 2000). Understanding the relationship of sounds to letters is a critical aspect in putting words together and taking them apart. Therefore, an early reading instructional program that is explicit in nature provides a variety of components in both large group and small group formats to teach in concrete terms the more complex skills necessary to understand the phoneme and grapheme relationships represented in printed text.

Elements of Intensity of Instruction

An equally critical set of factors in early grade reading instruction is the *intensity of instruction*. For those who have achieved less than their peers, increased intensity (also referred to as increased opportunity or acceleration models) is required to support their catching up with their peers. Continuing to make normal progress will only maintain the gap between at-risk readers and those whose growth has followed anticipated skill acquisition patterns (Foorman & Torgesen, 2001; Spiegel, 1992).

Secondary to the explicit nature of the instruction, intervention intensity (time on task and length of program) is critical in determining a student's catch-up rate when compared to on-grade peers (Pikulski, 1994). In a review of 30 programs, Elbaum,

Vaughn, Hughes, and Watson-Moody (2000) found the length of the interventions typically varied from 8 weeks to 90 weeks. For those programs lasting longer, the effect size was +0.37, while for those programs with greater intensity (fewer weeks, equal number of instructional hours) the effect size was nearly double, +0.65. This analysis shows that the intensity of instruction is a factor in program decision-making; however instructional intensity's true role is linked with the type of instructional practice.

Pikulski (1994) and Pinnell et al. (1994) both asserted that programs similar to *Reading Recovery* and *Success For All* have shown intensive reading instruction to be effective in meeting student needs. These approaches blend regular classroom instruction with one-to-one tutoring formats by specialist teachers, trained paraprofessionals, or peers. In a review of five intervention programs Pikulski (1994) concluded that small groups and one-to-one tutoring will be necessary for students at risk for reading failure. The duration of the grouping formats is also of importance. The effectiveness of longer meeting times with fewer days (i.e. shorter interventions) appears to make them preferable to longer interventions, but it remains unclear what the exact nature of the adult-student interactions should be to develop improved student outcomes.

The results of studies attempting to find the relationship between time on task, and student-adult interactions are mixed (Shanahan, 1998). It is not well understood whether interventions should be all one-to-one, or at times small group, and whether interventions should be carried out over an extended time or in a concise format. Analysis of the literature suggests the importance of a variety of lengths for interventions, but that

none of the programs contained in the reviews showed significant effects for any one duration of the interventions.

Therefore, although instructional intensity is important, the interaction of intensity with the other critical elements of interventions is not well understood. What is clear is that those students who are behind their peers in reading skill acquisition need more instruction than is provided during the regular classroom instructional programs. At this time, reading program developers must determine if short intense bursts of support are better than longer less intense instructional efforts. Taking this analysis and blending it with the explicit nature of instruction referenced earlier has been found to close some of the gap that exists between students who are meeting grade level standards and those not meeting grade level standards.

Elements of a Comprehensive Reading Program

Reading programs should include a comprehensive set of literacy elements important to early reading success. This includes critical elements of early reading skill development, variety of instructional strategies, and assessment protocols (Hiebert & Taylor, 2000; Rayner et al., 2001; Taylor, Pearson, Clark, & Walpole, 1999). Effective early grade interventions make phonemic awareness, phonics, vocabulary, fluency and comprehension tasks a part of every lesson. Of equal importance is that intervention protocols must be congruent with reading instruction that takes place in the regular classroom. Lastly, students should be engaged in auditory, verbal, and kinesthetic activities when representing their early reading abilities.

Instruction as we have seen in the two previous sections (explicit strategies and increased intensity) should also utilize a variety of strategies. When students represent on paper the sound they hear, they make the connection between the visual and the auditory (Pikulski, 1994). Therefore, the inclusion of writing is also seen as an integral part of a comprehensive beginning reading program.

In conjunction with making the reading and writing connection, group formats, sustained silent reading, model reading and kinesthetic involvement strategies should be utilized. Students who can read in pairs respond to questions after listening to reading, and students who can cut up words to make new words utilize a variety of brain activity to facilitate learning. Reading programs that are significantly structured in what is to be taught, but open to a wide scope of instructional strategies, will provide more matching opportunities for students at risk than those that are highly prescriptive in design (Pinnell et al., 1994). Teachers who have a diverse repertoire of methods are more likely able to reach a broader cross-section of students; supported with a framework of routine and a set of necessary elements, they have the potential to meet the needs of all students in the developmental continuum.

A comprehensive reading program also includes regular assessment as a part of its scope. A child's attempts to make meaning from text should be monitored and reinforced. Specific reinforcement at the time of miscues will help students to undo bad habits and reinforce the use of appropriate skills (Foorman et al., 2001; Leslie & Allen, 1999). This miscue tracking and support also become the basis of teachers' future plans for reading instruction. Future lessons should come from the current mistakes that students make and

teachers should continue to adapt lessons bridging what students know and can do with progressively more challenging skills. This daily monitoring of student abilities to shape lesson planning should be matched with ongoing program level assessment to understand how all students are improving.

Formative and summative assessments of student abilities are integral parts of determining if an early grades reading program is meeting the needs of the students. Programs such as *Reading Recovery*, *Open Court*, *Success For All*, and basal reading programs utilize interval and end-of-the-year assessment strategies to reflect on the successes of the programs' practices, whether it is a unit test common to basal programs, a curriculum-based measure like the Dynamic Indicators of Basic Early Literacy, or the Stanford 9 Achievement Test. These assessment protocols take time and have additional costs associated with them that effect schools' decision-making (Monk, 1995). However, programs that are effective evaluate their progress periodically within the year to adapt to student needs (Pinnell et al., 1994; Skindrud & Gersten 2006; Slavin & Madden 2001a). Effectively monitoring instruction is a critical part of a comprehensive program as it supports the teacher's recalibration of strategies and five critical elements emphasized in the National Reading Panel's 2000 Report.

Elements of Socially and Emotionally Supportive Strategies

Socially and emotionally supportive strategies comprise a broad category of elements necessary for effectively organizing the instruction of early reading skills. The category includes critical elements defining precisely what is supportive of the child, who should provide the intervention, and how the intervention will integrate with the

classroom instruction. To some extent, this category includes variables that do not fit neatly into the previous three categories.

Effective teachers recognize that the social and cognitive support of intervention practices must take into account the primary-age child affect during the reading process. Positive feedback and continual reinforcement are continuous requirements when working with any student, especially students who are at risk (Foorman & Torgesen, 2001; Spiegel, 1995). At the most intensive instructional level one-to-one tutoring takes into account these supportive variables. The close relationship (one-to-one pairing with instruction focused on the student's abilities) of the teacher and child is seen as one feature that has made individual tutoring a consistent choice for reducing reading failures. The specialized skills of one-to-one teachers are seen as the second variable that increases the supportive nature of the interaction.

Children who struggle with learning may see themselves as different from their peers, and therefore they should be taught by the most qualified teachers (Pinnell et al., 1994; Spiegel, 1995). Specialized reading teachers use strategies that build connections between the instructional protocols and what is happening in the classroom. These teachers assure that the instructional program is congruent with classroom reading instruction. Through the use of repetitive routines, scaffold instructional practices, positive reinforcement schedules, and explicit teacher planned interactions, the child accomplishes tasks that could not be accomplished without assistance.

Lastly, one of the consistent themes in the research, regardless of the methodological perspective, is that interventions take place early in the child's school

life. This is a critical element of *Reading Recovery*, and a main focus of non-reading and reading activities in *Success For All*. Pinnell et al. (1994) and Spiegel (1992) noted that the earlier good habits are developed, the fewer old habits will need to be unlearned.

Waiting too long before instruction or having to undo bad habits highly influences the nature of the adult-student interaction. If too much time is spent on relearning during the intensive skill development sessions, and if interactions are more negative than positive, then the affective context of the student learning may lead to less than satisfactory results.

In summarizing the four conceptual areas of elements critical to effective reading instruction and supportive interventions, it is clear that the body of literature looked for specific elements (Foorman et al., 1998; Foorman & Torgesen, 2001; Pinnell, 1994). Effective programs start early, include explicit instruction and direct instruction using phonemic awareness strategies, phonics instruction, vocabulary, comprehension, and fluency skills as key components. Also, because interventions are in addition to regular instructional opportunities, highly qualified individuals are necessary to carry out the supplemental instruction (Spiegel, 1995). The literature also maintained that interventions offer help outside the classroom that connects with instruction in the classroom. There is also a need for the explicit teaching of metacognitive skills. These skills are typically taught by scaffolding content as a form of instruction and intervention that supports students' understanding of what to do and how to think while reading (Foorman & Torgesen, 2001, Slavin & Madden 2001a). In the end, reading programs should teach students to solve problems in a manner that helps them to think about reading, supports their needs in learning the complex skills of reading, assesses growth in those same skills,

and considers their social and emotional well-being. Currently a number of examples exist that education reformers highlight when reviewing effective instructional programs that support all readers.

Examples of Reform Models that Utilize the Elements of Effective Reading Programs

In recent years a number of approaches have used these four categories of critical elements in either the design of the actual reading programs or as an outline for professionals responsible for improving students' early reading outcomes. One of these reading programs is consistently seen as a model program for Title I schools; *Success For All* (Skindrud & Gersten, 2006). Also under consideration is a second approach to developing a reading program that attempts to influence both practitioner and schoolwide reform while relying on *Reading First*, one of a variety of current reading curricula. Both approaches have research backgrounds that help define their theoretical framework, and both programs have parameters by which those who interact with them are expected to function.

Success For All

Success For All (SFA) is a well researched and often cited early grades reading program (Borman & Hewes, 2002; Skindrud & Gersten, 2006). SFA began in the Baltimore, Maryland schools in late 1980s as a joint project with John Hopkins University (Slavin & Madden, 2001a). Over the past 15 to 18 years the SFA Foundation has continued that work, building SFA into a consistently replicable and portable schoolwide reading reform model developed for Title I schools in the U.S. (Borman & Hewes, 2002; Skindrud & Gersten, 2006). In its current form school professionals vote to

adopt the SFA model that includes a set of components seen as critical to student success in developing early reading skills. SFA contains a variety of principles (explicit instruction, grouping strategies, and basic reading skills) found in effective reading instruction research, but also encompasses elements of family support, teacher support, and school reform. SFA guiding philosophies are based in supports for teachers with the use of a reading facilitator or mentor, regional staff development, and implementation support. Cooperative learning is a key instructional strategy, and at its center SFA is a formulaic and prescriptive instructional curriculum. The curriculum includes early reading imperatives similar to those defined by the National Reading Panel in 2000, such as: phonetic skill development, predicting strategies, vocabulary development, fluency, comprehension strategies, and related reading skills seen as necessary in early grades reading programs.

Success For All as a reform model. SFA instructional philosophy is derived from the position that early intervention and focused support can prevent a myriad of potentially negative student outcomes in the life skill of reading. SFA developers Slavin and Madden (2001a; 2001b) believed that both special education referrals and primary grade retention can be significantly reduced, and in the case of retention, even eliminated with early intervention on a variety of fronts.

As a reform model SFA requires schools adopting its approach to reading (and potentially math) to pass an 80% *show of support* vote before beginning schoolwide implementation. Once schools have accepted the model, SFA focuses on three levels of support: (a) student and family support through the use of a school and home liaison, (b)

teacher support through curriculum materials and a teacher facilitator or mentor, and (c) schoolwide support through the use of regional SFA Foundation representatives charged with supporting staff development and fidelity of program implementation. Each of these levels of support may be funded through federally supplied Title I resources where they are available. However, SFA is not only used in Title I schools; some schools fund it from non-title revenue (Borman, 2005).

Success For All reading curriculum. SFA reading curriculum has two components: K-1 is *Reading Roots*, and grades 2-5 (up to grade 6) program, *Reading Wings*. The K-1 program uses SFA developed text in phonetically regular context to give students opportunities to work in groups or with a teacher. These contrived text formats utilize paired and guided reading strategies in which the adult controls the reading of text and the students read a phonetically regular text as part of the same story. This approach is first introduced in the second semester of kindergarten and concludes in the second semester of first grade. As early instruction progresses the student-read portion of the book begins to dominate over the teacher's controlled text. In addition to small group reading formats, letters and letter sounds activities link the phonemic and grapheme segments together in what the authors refer to as an active and engaging set of activities (Slavin & Madden, 2001b).

This process is carried through to the second grade as the emphasis of phonemic and grapheme skills gives way to strategies that develop student vocabulary, increase reading fluency, and develop stronger comprehension skills. When students reach the second grade they move into a reading curriculum called *Reading Wings*. *Wings* activities

are built around the skills of predicting, summarization, vocabulary building, and decoding practice integrated into story related writing. An emphasis on cooperative learning strategies places students into partner reading and vocabulary building activities. Direct instruction takes place in the area of comprehension with students learning how to rephrase, summarize and elaborate on the meaning of the materials they are reading.

Success For All student groupings. In SFA schools students' reading activities are carried out in schoolwide reading times where student-to-teacher ratios are reduced using all available staff. The program's reliance on cooperative learning strategies is emphasized through student regrouping practices. Students are grouped in heterogeneous classes for the majority of the academic day. However, students are regrouped for 90-minute reading and writing instruction sessions. With the use of Title I reading support personnel students are regrouped into homogenous groups by reading performance level. These groups are then led by all the adults in the school, including classroom teachers, reading specialists and teachers in areas of physical education, counseling, and library media services. This "common reading period" is viewed by SFA as a protected time that is not to be interrupted. This regrouping is a form of the Joplin Plan that has been shown to be effective at increasing on-task performance during instruction (Slavin, 1987).

Success For All reading tutors. One of SFA's philosophical bases is that one-to-one tutors who are highly trained in reading instruction offer the most successful support to students who are behind in reading outcomes (Slavin & Madden, 2001a; 2001b). Tutoring takes place outside of reading and math lesson presentations and lasts for about 20 minutes each day. Tutors are trained reading specialist who link the curriculum taught

during tutoring time to the basal program in use during the reading lesson earlier in the day. This approach assures that teacher scaffolding is done using the same story and concepts that are used during the student's reading group.

Success For All Student assessment. Student grouping is achieved by ongoing assessment. As a critical element of effective reading programs, assessment is used to regroup students on an eight-week cycle. The assessments results are used for three purposes: (a) to determine who is to receive tutoring services, (b) to evaluate the effectiveness of tutoring and classroom instruction, and (c) to reconstitute teacher-led groups (Slavin & Madden, 2001b).

Skindrud and Gersten (2006) emphasized that the strength of SFA's effectiveness is the inclusion of progress monitoring systems to guide instruction, small class size for reading instruction, and intensive early intervention for at-risk readers in conjunction with cooperative learning principles. Program evaluations of SFA schools by the SFA Foundation, and by independent researchers, show a number of differences between SFA students and comparative students in non-SFA programs. Of particular note is that SFA students in the bottom quartile of their class have scored better than their comparative peers in non-SFA schools (Borman & Hewes, 2002). This is true when compared to models like *Reading Recovery*, *Open Court*, and other less formulaic reading programs. These successes are typically attributed to grouping formats, tutoring services, and a philosophy of early intervention summarized by the program's name: *Success For All* (Slavin & Madden, 2001b).

The SFA Foundation has analyzed program fidelity among participating schools and finds support for the program through internal comparisons. Schools judged to have highly effective implementation practices see significantly better reading outcome scores when compared to control schools. Only those with poor implementation practices see little or no differences in the reading outcome measures in non-SFA school comparisons (Slavin & Madden, 2001a). The SFA Foundation relies on research by outside investigators, and their own internal research, to support claims that SFA is an effective instructional protocol that encompasses the elements seen as necessary for effective early reading instruction.

Reading First

A second reform approach with a less specific instructional model comes out of recent policy legislation as part of the No Child Left Behind Act of 2001 (NCLB 2001). The *Reading First* initiative is part of an ongoing set of national, state, and local education policies developed by a variety of interested parties and implemented by state and district educators. The latest policy approach is the current iteration of national education policies starting with the Elementary and Secondary Education Act of 1965 and progressing to practices defined in the NCLB 2001. Successive incarnations of the educational improvement policies have common goals with regard to early literacy instruction: (a) teach children to read in the primary grades, (b) improve reading skills by having teachers use practices that are supported in rigorous research, (c) expand the number of quality programs available to students and families, and (d) intervene early in

a student's education to reduce the number of student retentions and student referrals to special education (Barone, Hardman & Taylor, 2006).

Reading First as a reform model. The *Reading First* initiative as defined in sub part B of the NCLB 2001 has not deviated from past trends to improve student outcomes, and seeks to implement these same goals. Accountability is seen in the assessment requirements and in use of sanctions against those schools that repeatedly do not meet targets of adequate yearly progress. *Reading First* protocols use as their basis the National Reading Panel report to identify research-based approaches to instruction, and critical elements of reading programs capable of getting all students reading on grade level by 2014. *Reading First* also uses staff development and hiring criteria as means to create a highly qualified work force to work with Title I program eligible students. Lastly, *Reading First* legislation requires schools to identify readers at risk for reading difficulties and to develop interventions that will assure that all students are reading on grade level by third grade.

Reading First reading curriculum. Curriculum choice in the *Reading First* protocol is left up to the participating school or the district. State *Reading First* offices do not subscribe to a specific curriculum, but do provide a set of goals that are less reliant on specific curricula, and more reliant on a reform model. Each state office of *Reading First* sets criteria for reading program adoption, and many have used evaluations by local universities or consortia to guide that process. Schools are required to choose their reading program from a specific list of research-based curriculum. This choice allows schools to use a variety of products that fit the needs of their students and their program

assessment requirements. State and district administrators have looked at a variety of curricula that entail the elements outlined in the National Reading Panel (NRP, 2000) documents as critical to early reading success. As with *Success For All*, *Reading First* schools hire reading coaches to support teacher improvement and ongoing growth. However, there is not a prescriptive or formulaic reading curriculum required by the local grants, nor is a set of grouping strategies specified for reading instruction.

Reading First student assessment. *Reading First* schools' governing agencies have added the need for effective program and student assessment protocols. Progress monitoring is expected to be an ongoing element of each *Reading First* grant school. The Reading Coach (a reading specialist with supplemental training in reading and assessment) is required to summarize student growth in quantitative terms regarding agreed upon reading outcomes. This leads to supporting teachers in planning instruction for all students, with emphasis on those students who are not progressing in a manner that is expected. Along with progress monitoring activities is an added element of school and district level sanctions that can be implemented if schools do not show the improvement required by state offices of education. These sanctions link state reading assessment outcomes in grades 3, 4, and 5 with students' progress in reading.

The relatively recent creation of the *Reading First* initiative has produced very few research studies of its overall effectiveness as a support for struggling readers or a reform format intent on changing professional practices. At this time more evidence is required to draw firm conclusion about the policy as a means to improve student reading scores in some of the country's most challenging schools.

Reading Initiatives Summary

There exists a rich theoretical base on what critical elements effective reading programs should include and what professionals should be measuring to assure that programs are meeting the needs of students. However, the pressures of policy initiatives require decision-makers to look more concertedly at costs in providing these critical elements to achieve high quality student outcomes. In this context decision-makers should be able to use cost analysis tools in a manner that informs decision-making, but also communicates programs' strengths and weaknesses to the interested public. Looking into the theoretical framework of educational program costs analysis should provide a backdrop for understanding what is important to look for in cost analysis tools. Effective use of these tools may ultimately shed light on policy decisions by educational and governmental bodies.

Approaches to the Analysis of Educational Program Costs

Once the elements of effective instructional programs are known and administrators bring into the analysis the alternatives that contain these variables, then comes the time to look into the economic variables of the programs. King (1994) places four cost analysis tools in the toolbox of program evaluators: cost-effectiveness analysis, cost-benefit analysis, cost-feasibility analysis, and cost-utility analysis.

Cost-Effectiveness Analysis

Cost-effectiveness analysis as a method of program evaluation has been used to compare educational alternatives that achieve like academic outcomes (Borman & Hewes, 2002; Levin & McEwan, 2001; Rossi, Freeman, & Lipsey, 1999). The use of

cost-effectiveness protocols to analyze the costs of the alternatives in relation to their outcomes holds promise for administrators working with shrinking budgets. Cost-effectiveness procedures have the ability to get past the mere adding up of program expenditures to define the less apparent costs associated with changing work roles of practitioners, shifting of costs to constituencies, and the discounting of cost over time. The discrete nature of these costs variables may come to light only when individuals working in the program have an opportunity to interact with the researcher (Grissmer, 2002; Levin & McEwan, 2002). An added advantage of the cost-effectiveness analysis methods is that evaluators do not get entangled in attempting to gauge future benefits, or in developing challenging algorithms that attempt to convert nonmonetary outcomes into monetary values as would be required by cost-utility and cost-benefit analyses.

Ingredients methodology. Cost-effectiveness analysis most commonly relies on the *ingredients method* first described by Levin (1983), and subsequently elaborated on by Levin and McEwan, (2001) and King (1994; 1997). This approach to estimating costs is consistent with Chambers' (1999) approach referred to as the *resource cost model*. Both methodologies are an approach to cost estimation that is supported in the literature by King (1994), Hummel-Rossi and Ashdown (2002) and others (Monk, 1989; Tsang, 1997).

The ingredients methodology is built around the notion that the costs of reading programs or support interventions are made up of a variety of factors including facilities, materials (curricula, administrative and instructional supplies, etc.), human resource needs, client inputs, and unique program variables. A cost in this methodology has two

variables. It can be quantified as expenditure for goods or a service or quantified as an opportunity not pursued (Chambers, 1999; Levin & McEwan, 2001). Matching both expenditures and the costs of missed opportunities in each of these categories requires a more elaborate cost review beyond the tabulation of program budgets.

Costs. Costs by their nature are defined by two factors: (a) expenditure of resources and (b) the value of the opportunity lost in making the expenditure (Levin & McEwan, 2001). Costs can also be categorized by their nature within the specific program being reviewed. In this context costs can be either *fixed* or *variable*. Fixed costs are relatively static and remain somewhat unchanged in response to increasing client base or improvement of services. Variable costs are more dynamic in response to changes in variables of the program and can influence cost-effectiveness analysis if the reviewer does not take into account the scale of the project. Therefore, the educational costs of programs must be understood on two levels: First, what is the opportunity sacrificed under the current expenditure pattern, and second, what is the scalability of the current expenditure pattern, and how does that affect the analysis of the costs.

Many programs' costs can be seen in the budget expenditures documents and can be accounted for in a review of documents and interviews of program personnel. Interviews of the school principals, program managers, and those who work within the program provide insight into the expenditures and can give a more complete accounting of the programs costs (Hummel-Rossi & Ashdown, 2002; King-Rice, 1997).

Researchers of cost-effectiveness analysis find insufficient clarity in the definition of opportunities cost, and the discount costs of program alternatives (Levin & McEwan,

2001; Ilon & Normore, 2006). These costs, referred to as *indirect costs*, are the *discounted costs* (the potential impact of opportunities over time) of foregone opportunities and program demands, and the *distributed costs* to those who bear the burden of the implementation expenses. Each of these elements is an important measure of program cost analysis and is captured in: (a) a review of distributed costs, (b) a sensitivity analysis, and (c) an expenditure review (Grissmer, 2002; Levin & McEwan, 2001).

Discounting costs. Discounting costs is a means of accounting for past opportunities lost because the resources were used in the set up and running of an alternate program. In schools discounted costs have real effects based on the funding challenges faced in many communities. As an example, the allocation of money to set up a program to support reading cannot be used for the next best alternative, nor can it be recovered over time. Since students move through schools' systems, opportunities missed in the early years cannot be recaptured, except at differing costs later in the child's academic career. These costs must be considered in the cost-effectiveness analysis of program implementation. Discounting is an economic concept that is rarely considered by administrators; this lack of attention to indirect costs may over time erode the strength of one academic program while another receives a disproportionate investment of the limited resources (Levin, 2002; Levin & McEwan, 2001).

Sensitivity analysis. Sensitivity analysis is a technique to help program evaluators quantify uncertainty in a dynamic ingredient under consideration (Boardman, Greenberg, Vining, & Weimer, 2001). The analysis is an estimation of an ingredient's cost when a

range of costs may apply to a particular program ingredient, or when no reliable cost exists for the program ingredient. In the case of reading interventions this can be seen in estimating the cost of a specialized reading teacher, or conversely a volunteer who provides individualized support for a child. The findings of the sensitivity analysis can support the decision-making process when the assumptions driving the analysis are well understood by the analysts (Boardman et al., 2001).

The assumptions of the costs estimates form the basis for the cost effectiveness analysis. The sensitivity analysis estimates costs using a variety of assumptions, most analysis typically include high, medium, and low cost values. For variables with the greatest uncertainty, the accuracy of the medium value can be tested against the assumptions of the high and low values. If the cost-effectiveness ratio holds over the three analyses, then the medium estimate can be seen as a true estimation of the costs. This rather time-intensive process can be critical in making good decisions, but is often left out of analysis presented in current research. Levin and McEwan (2001) suggest that sensitivity analysis should be used when unique variables are being evaluated and that the assumption of uncertainty should be consistently tested even though the reporting of the sensitivity analysis finding is not really necessary. This is particularly important when the ingredients under consideration have a high degree of variability between the alternatives under consideration.

Expenditure reviews. Expenditure reviews of budgetary documents are undertaken to account for the planned and unplanned use of resources. In this approach the evaluator is looking to match line item budgetary planning with line item charges to

budget (Chambers, 1999). Those who include this approach in program evaluations feel that budgets are one of the tools evaluators can use to gauge the effectiveness of implementation and decipher real costs in attaining program goals.

Using budgetary expenditures is a necessary step in program evaluations. However, it should be one of multiple measures used to define resource use. The tenuous nature of educational budget documents can be traced to the lack of specificity of the documents to accurately: (a) state how the budget is to be expended (i.e., budgets are line items and not binding, budgets are predictions for a time period, major transactions are tied to a year, not discounted over time) or (b) state the exact nature of how the account is to be spent (i.e., accounts do not show exact charges to the budgets, record changes in budget priorities as the year goes forward, or show shared costs between budget sources). These cautions are presented by researchers to advocate for more effective training of school personnel in the use of cost analysis protocols. Levin (2002) uses the volatile nature of program spending to support training that assures that the public's resources are allocated in the most cost-effective manner.

Summary

Combining the theoretical frameworks of critical elements of early reading and intervention programs with the components of accurate cost analysis can support school personnel who make critical decisions about reading program implementations. Adding the variables of task enactment and task function to the analysis of costs can improve the understanding of hidden costs or clarify lost opportunities through the cost analysis process. Understanding the variables that change workloads and shift responsibilities

among parties and therefore add unforeseen costs is important for leadership teams. The ability to quantify these variables as well as the traditional costs of effective programs has potential to shed light on the background issues that make these early reading programs effective. Knowing what makes these programs effective may give rise to understanding how to scale programs to differing environments in a manner that is consistent with the program's foundations and be cost-effective at the same time.

A significant body of research exists defining the variables of reading programs that support the growth of both conventional and struggling readers. That research includes what to look for in teachers' instructional practice, and the measurement of student outcomes. Linking that research with the extensive research in costs analysis as related to educational programs is a gap that is not completely filled. There have been comparisons of programs (Borman & Hewes, 2002), but only limited work on evaluating programs that are running simultaneously, using similar outcome measures and inclusive of teacher, principal, and program manager feedback. Two approaches have been outlined from a review of the literature that could potentially achieve the research-based nature of reading instruction as outlined by the National Reading Panel 2000 report. Although both *Success For All* and *Reading First* approaches are used in schools today, and much is known about *Success For All*, little is known about the effectiveness of *Reading First* as a reform protocol and even less is known about the costs structures relative to instructional effectiveness. When compared with locally designed approaches the costs of all programs can be analyzed and reviewed. Undertaking and then understanding this process may potentially improve leadership team decision-making. In

light of shrinking resources, more demanding student characteristics, and shifting policy initiatives, this understanding may be critical for the wise use of communities' resources.

CHAPTER III

METHODS

The design and methods chosen for this study reflect the nature of program evaluation in schools as found in the literature (Levin & McEwan, 2001; Taylor Fitz-Gibbon & Lyons Morris, 1987). This mixed methods study used a quasi-experimental format of a pre-post design to capture differences between two locally designed reading programs and two structured programs in four elementary schools. Early grade reading outcomes were analyzed for approximately 110 to 135 first grade and 110 to 135 second grade students randomly selected from four area schools.

Survey methodology was used to capture the independent variables of program costs and the decision-making considerations of the schools' leadership. The first survey tool described the situational context of leadership decisions as they relate to leadership task functions, task enactment, the social distribution of task enactment, and situated distribution of task enactment (Gronn, 2000; Weick, 1975). A second survey tool evaluated the costs functions in each of the four programs. To triangulate the information gained from observation and program leader interviews, a document review was carried out. Both quantitative and qualitative data were used to generate a complete a picture of each reading program and include the development of accurate cost estimates. The ingredients method of cost estimates was used to generate cost-effectiveness ratios for each of the four programs.

Participants

This study is a comparative evaluation of four reading approaches in four elementary schools in a mid-size city in the Pacific Northwest. Each school had multiple sections of single graded first grade and single graded second grade, and two schools had blended first and second grades. The demographics of the schools represent the demographics of the city, and reflected the potential policy parameters principals and district administration would have to take into consideration when choosing between alternative reading programs. The schools in the study were a sample of convenience in an attempt to compare alternative reading programs in a continuous geographic area. The approximately 70 students within the school (35 at first grade, and 35 at second grade) were a simple random selection from the potential students at their grade level.

Subject anonymity was maintained through a double number system. In this format each student had a district SASI (school database software) database number and an Oregon Department of Education permanent tracking number (SSID) assigned to the data collections sheet. The researcher had access to the students' SASI information at each school.

Table 3.0 delineated the four schools' aggregate characteristics for school year 2006-07 as reported by school program managers, principals, teacher surveys, and the district for the Oregon Department of Education's report cards of schools website (Oregon Department of Education, 2007). The schools ranged in size from 327 to 527 students in grades kindergarten through fifth. Class sizes for each school were at or above the state average as reported on the school's state reports cards.

Table 3.0

School Demographic Characteristics

School Characteristics	Adams	Franklin	Hamilton	Jefferson
Total Population	317	527	342	504
Student Ethnicity	--	--	--	--
White	77	384	251	195
Hispanic	196	34	27	191
Black	4	11	1	1
Alaskan/Native Amer.	3	4	5	1
Asian / Pacific Ins.	7	23	9	70
No Response	19	15	17	23
Multi-Racial	11	56	32	23
ESL Percentage	43%	.01%	.005%	47%
SES (Free Lunch)	89%	32%	42%	76%
Teacher Experience ¹	N/A	13.4	10.8	3.1
Average class size (K-5)	19.2	24.33	21.2	20.8
% Met Grade Benchmark	--	--	--	--
2005-06 1st Grade	41%	78%	39%	65%
2005-06 2nd Grade	28%	92%	65%	87%
Student stability ²	21%	10%	16%	19%

¹Average years of experience for grade 1-2 teachers in the school responding to survey.

²Percent of new students in a school at the end of the year from the start of the year.

Tables 4.0 and 5.0 present the demographic characteristics of the sample first and second grade posttest populations after attrition was taken into account.

Tables 4.0

Grade 1 Demographic Information for Sampled Students in Four Participating Schools

	Schools				Total
	Adams	Franklin	Hamilton	Jefferson	
Sample size (<i>n</i>)	24	32	32	31	119
Gender					
female	11	17	17	13	58
male	13	15	15	18	61
Age					
<i>Mean</i>	5.54	5.49	5.53	5.49	--
<i>SD</i>	.31	.33	.27	.32	--
Free or reduced lunch					
no	5	21	20	6	52
yes	19	11	12	25	67
English as primary language					
no	4	1	1	23	29
yes	20	31	31	8	90
Ethnic code ¹					
2	2	2	--	9	13
3	--	2	--	--	2
4	7	2	1	10	20
5	13	18	24	11	66
6	1	5	4	1	11
9	1	3	3	--	7

¹Ethnic Codes: 1 = American Indian / Alaskan Native; 2 = Asian / Pacific Islander; 3 = African American; 4 = Hispanic; 5 = White, 6 = Multiracial; 9 = No Response.

School Demographics and Reading Program Descriptions

The study involved four schools, all referred to by pseudonyms, that were part of a school district of approximately 39,000 students. Two of the four schools used a 2002

Table 5.0

Grade 2 Demographic Information for Sampled Students in Four Participating Schools

	Schools				Total
	Adams	Franklin	Hamilton	Jefferson	
Sample size (<i>n</i>)	24	35	34	32	125
Gender					
female	10	19	20	13	62
male	14	16	14	19	63
Age					
Mean	6.69	6.52	6.49	6.52	--
SD	.39	.30	.30	.33	--
Free or reduced lunch					
no	6	24	12	10	52
yes	18	11	22	22	73
English as primary language					
no	3	1	3	20	27
yes	21	34	31	12	98
Ethnic Code ¹					
1	--	1	--	--	1
2	1	2	1	4	8
3	--	1	--	--	1
4	7	4	4	11	26
5	13	23	25	14	75
6	1	3	4	3	11
9	2	1	--	--	3

¹Ethnic Codes: 1 = American Indian / Alaskan Native; 2 = Asian / Pacific Islander; 3 = African American; 4 = Hispanic; 5 = White, 6 = Multiracial; 9 = No Response.

version of *Scott-Foresman Reading* produced by Pearson Education and purchased by the district in 2002. This curriculum was used within a locally selected instructional format partly the design of the publisher and partly the design of the school staff. A third school was in its ninth year of using the *Success For All* curriculum materials and the associated methodologies and reform practices as their reading instruction program. The final school

in the study had been required by state and local district guidelines to use a *Reading First* approved protocol for reading instruction. This school used the *Read Well* early reading curriculum program for their kindergarten and first grade students, and combined *Read Well* with the 2004 edition of *Scott-Foresman* curriculum at second grade for below grade level readers.

Each school had their particular historical perspectives, yet each was bounded by the design of the specific programs they were implementing, either by choice or by requirement of an outside agency. The descriptive elements of the reading programs are presented in Table 6.0 as a means of generating comparative elements for the review of alternatives programs required in a cost-effectiveness analysis.

Hamilton Elementary

Hamilton elementary was viewed as one of the locally designed program schools in this study. A school of 327 students, Hamilton had two first grade classrooms, two classrooms of second graders, and one classroom of first and second grade blended. Hamilton was situated in a middle to lower-middle class section of the city.

Hamilton school used the *Scott-Foresman* reading program and created their instructional approach based on its historical perspective in teaching reading, and the school's decision-making process under minimal influence from sources outside of the school. Hamilton, like all the schools in this district, was provided with a .5 full-time equivalent (FTE) reading teacher with the only stipulations that the teacher must reduce class size and serve students within grades K through 3. The locally designed model was set up as a reading *pull-out model* in which a reading specialist and an instructional

assistant took students from each of the K-3 classrooms over the first half of the day and supported the most struggling readers.

At Hamilton reading instruction was carried out in a 90-minute “literacy block” that began at the start of the day and was intended to be uninterrupted. The classroom teacher grouped students for work, and the reading specialist took students for about 30-minutes at a time from each classroom during that 90-minute instructional time. The reading specialist utilized the *Scott-Foresman* materials and directed the work of the instructional assistant who also worked with a small group of students. The process was a five-day-a-week approach to providing service to each of the K-3 elementary teachers. Hamilton school represented the typical non-title program school within the district. They received (as did each of the other schools) a .5 FTE reading teacher position, but did not have Title I program funds to purchase other elements that might support their reading program. Hamilton had a relatively stable student population in terms of growth with a school size that stays around 325 students. The costs under consideration when looking at Hamilton were those costs in facilities use, materials acquisition, personnel expenditures, unique inputs, and school required inputs. Being a fairly typically school, Hamilton was seen as having costs relatively typical amongst schools of its size and demographic, and thus representative of the district in a policy decision-making context.

Franklin Elementary

Franklin Elementary was a school of 527 students that had four first grades and three second grades all taught in a single grade format. It was situated in a predominately middle class section of the city. Franklin school used a locally developed program and

Table 6.0

Reading Programs Characteristics in Subject Schools

Programs Characteristics	Schools and Reading Programs			
	Adams SFA	Franklin S-F	Hamilton S-F	Jefferson RW-RF
<i>Reading outcomes measured</i>				
Letter ID (grade 1 & 2)	Yes	Yes	Yes	Yes
Sound ID (grade 1 & 2)	Yes	Yes	Yes	Yes
Word recog. (grades 1 & 2)	Yes	Yes	Yes	Yes
Book level (grades 1 and 2)	Yes	Yes	Yes	Yes
Fluency score (1-5 rating)	Yes	Yes	Yes	Yes
Comprehension (1-5 rating)	Yes	Yes	Yes	Yes
<i>Curriculum characteristics</i>				
Explicit phonics instruction	Yes	No	No	Yes
Use of leveled text	Yes	Yes	No	Yes
Decoding	Yes	Yes	Yes	Yes
Fluency strategies	Yes	Yes	Yes	Yes
Comprehension strategies	Yes	Yes	Yes	Yes
Periodic formal assessment	Yes	Yes	No	Yes
External auditor	Yes	No	No	Yes
<i>Instructional characteristics</i>				
Protected reading minutes	90	90	30	120
Avg. <i>n</i> students per group	10-12	6-7	6-7	5-6
Small groups frequency	5 d/wk	3-4 d/wk	5 d/wk	5 d/wk
Minutes of intensive reading	30-45	30	30	25-30
Reading specialist support	No	Yes	Yes	Yes
Instructional assistants support (min./week)	None	2 hrs/wk	1 hr/wk	None
Training of inst. assistants	Yes	No	No	Yes

S-F = *Scott-Foresman* (2002); SFA = *Success For All* (1999); RW-RF = *Read Well* (2005).

created their approach based on its historical choices in teaching literature reviews, and the school's site-based decision-making process. This school, like all the schools in this urban district, was provided with a .5 FTE reading specialist with the only stipulation that the teacher serve students within grades K through 3 by reducing class size.

The locally designed model was set up as a *reading flooding model* in which a reading specialist combined with two instructional assistants and *flood* into each first and second grade classroom for 30 minute intervals to support the classroom teacher in small group reading instruction. Approaches similar to this have been attempted in a variety of schools throughout the United States and Canada (Hedrick & Pearish, 1999; Homan, King, & Hogarty, 2001; Kinnucan-Welsch, Magill, & Dean, 1999).

Franklin reading instruction was carried out in a 90-minute *literacy block* that began at the start of the day and was intended to be uninterrupted. During the 90 minutes each morning, 60 minutes were teacher-directed large and small group instruction in which both reading and writing tasks took place. Thirty minutes of this time were what Franklin called *Reading Flooding Time*. In this 30-minute segment students were taught in small groups targeted at students' individual reading level using both the *Scott-Foresman* basal reading program and leveled books from various publishers.

The classroom teacher grouped students for work with four individuals: a reading specialist, the teacher, and one of two instructional assistants. This grouping was initially done in the fall and was based on three agreed upon parameters: (a) the previous year's spring Primary Literacy Assessment score, (b) the current book level as assessed through

a running record, and (c) reading ability span (range of book levels) of the class. Groups were reformed as need arose through the use of monthly running records.

The reading specialist took the lowest grouping of students, and the smallest of the groups at about three to four students per 30-minute segment. The small group instructional format was based on the reading specialist's background in teaching *Reading Recovery*. The classroom teacher instructed the second lowest group, usually a group of six to seven students who were at various degrees of *on* or *below* grade level. The middle high group (group three) had the students working on teacher-designed tasks *at* or *above* their grade level and carried out by an instructional assistant. Their flooding session work revolved around using a leveled text to develop vocabulary skills, fluency task practice, literal comprehension skills, evaluative comprehension skills, and early writing skills. The highest group also worked with an instructional assistant to develop more advanced skills in vocabulary, strategies for the development of literal and evaluative comprehension abilities, and writing tasks. These four groupings met for 30 minutes at a time and the students were returned to class simultaneously for resumption of activities in the large group format. This grouping and instructional approach was repeated throughout the morning in each of the seven classes that make up the first and second grades.

Franklin's costs were evaluated on the basis of how the ingredients (personnel, facilities, equipment and materials, unique program inputs, and school required inputs) in the reading program contributed to overall costs of getting all students to reading benchmarks. The flooding model that Franklin used dedicated personnel specifically to

reading tasks at the expense of other academic areas. This may be different than the other three schools' use of personnel as well as personnel use in other programs around the district, and may contribute to opportunities lost with the reallocation of funds. When compared to other schools in the study, the costs of the flooding model may be uniquely different than those seen in a school of its demographic within the policy development context of the district.

Adams Elementary

Adams Elementary was one of three schools in this district that used Slavin's and Madden's *Success For All* (SFA) reading instruction program. A school of just under 335 students, Adams had two parallel SFA reading programs in use, one for instruction in English as the native language and a second for instruction in Spanish as a native language. Adams was both a Title I school and a bilingual program school. Adams had two first grade only classrooms, one class of first and second grade blended, and two second grade only classrooms, and was situated in an economically disadvantaged section of the city. The demographic makeup of the school is presented in Table 3.0.

Adams had utilized SFA for both its English speaking and Spanish speaking students for eight years. Annually the staff reevaluated its desire to use SFA and in the most recent vote 95% of the staff approved its continued use. Adams, like all the schools in this district, was provided with a .5 FTE reading teacher with the only stipulation that the teacher must reduce K-3 class size. Adams used a program manager who served as a second reading specialist at the school. The program manager's role was to organize the SFA program for the school and help track student progress through formative

assessments designed by the curriculum developers. Adams also had a regional SFA representative available to them who visited on occasion to review testing outcomes, work with the program manager, and answer questions about program implementation and staff development needs.

As seen in the previous two schools, Adams' reading instruction was carried out in a 90-minute *literacy block* that began at the start of the day and was intended to be uninterrupted. The classroom teachers came together to group students for their reading instruction. Reading groups were made up of students from each of the school's classrooms based on student SFA reading levels. During group reading time students from a variety of grades joined each other based on the most recent instructional level. Teachers planned for their group level and students traveled to new rooms for direct instruction in cooperative group format. This grouping format lasted for 30 to 45 minutes of instructional time each day.

SFA utilized intensive tutoring services for those students assessed at below grade level. These tutoring services were scheduled outside of core content instructional times. Tutors used the SFA materials in conjunction with the reading group teacher to assure that skill practice was in alignment with the skills being taught in the daily reading group. The process was a five-day-a-week approach to providing service to each of the K-5 elementary grade teachers.

SFA had a number of program variables that led to increased personnel and administrative costs. A reading program specialist organized the program for the school and helped track students through formative and summative assessments designed by

SFA curriculum developers. SFA had been positioned as a program that effectively utilized Title I program funds to improve student abilities (Skindrud & Gersten, 2006). SFA had a number of variables that led to increased costs in facilities use, materials, personnel, and other costs unique to this structured program. It had consistent progress monitoring activities that served Title I reporting requirements and could be scaled within a school to students of differing demographic characteristics. Programs that redistribute students may have a variety of costs that fall into a distributed category. These are costs that are not directly related to expenditures, but that affect educational opportunities for either adults or students. SFA also had significant costs that are not directly related to providing reading instructional services to students. As a model SFA had costs associated with the home-to-school liaison, contract costs for oversight of program fidelity, staff development training, and initial and annual consumable curriculum materials. The costs associated with the SFA model provided a good alternative to the three other models in the study. SFA cost evaluations provided an opportunity to compare alternatives with other Title I schools that also are bilingual schools.

Jefferson Elementary

Jefferson Elementary was a school of 504 students with four first grades and four second grades all taught in a single grade format. It was situated in a middle class section of the city, but also served a large number of multifamily housing dwellings. Jefferson was a Title I and bilingual program school with second language learners speaking a variety of languages, as represented in Table 3.0.

Jefferson began using the *Read Well* curriculum in 2005 as its primary grades reading instruction program. Jefferson created their current approach to teaching reading based on its historical perspective in teaching reading, literature reviews, *Reading First* grant criteria, and the school's decision-making process. They were required by outside sources to review their approach to reading under the guidance of a district assigned *Reading First* grant coordinator. Jefferson had a total of 1.5 FTE reading coach positions. These positions were split between three teachers who worked under the *Reading First* grant. These specialists coordinated the *Reading First* activities of staff development, performed student assessment, and analyzed student outcomes for grades K through 2. Like all the schools in this district, Jefferson also had a .5 FTE reading specialist teacher; the only stipulation was that the district-provided reading specialist reduces class size while serving students within grades K through 3.

At Jefferson first and second grade reading instruction was carried out in a 120-minute *literacy block* that began at the start of the day and for most classes went uninterrupted. During the 120 minutes each morning first graders received the small group *Read Well* lessons. Students in second grade below grade level also received *Read Well* lessons. The on-grade level second grade groups received 60 minutes of the instruction in a large group format (whole class) and small group formats took over after the large groups finished. Groups were teacher directed in which both reading and writing tasks were the focus. The on-grade level students used the *Scott-Foresman* reading curriculum available to students at Franklin and Hamilton.

The first and second grade classroom teachers organized students in groups using scores from Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and unit assessments from the *Read Well* and *Scott-Foresman* programs. A small group approach was used with the *Read Well* program with the district-provided reading specialist in the classroom during much of the leveled group instruction. Second grade on-grade level students were grouped in a similar manner. The second grades grouping format was based on *Scott-Foresman* unit tests. At Jefferson the reading specialist also took the lowest grouping of students in a later part of the day to do pull-out work on discrete skills. These groupings went for 30 minutes and the teacher was at times joined by instructional assistants and their group of students.

Jefferson was in its fourth year of being a *Reading First* school. Due to extensive oversight of *Reading First* programs the cost structures were well documented in the ingredients categories, the only exception being daily consumables and site Title I and general fund contributions. This study attempted to clarify the personnel, facilities and materials costs, and specifically the unique inputs required by *Reading First* schools. The *Reading First* protocol required a reading coach, ongoing assessment monitoring, use of outside consultants, and use of instructional materials from a state approved list. Therefore, one might anticipate increased costs over a traditional reading program. The unique costs consisted of assessment protocols outside the regular reading assessments (DIBELS and SAT-10), the time used by the reading coach to analyze those outcomes for both training and instructional purposes, and the increased staff development time for implementing new reading curricula.

In using these four schools a variety of alternative approaches to early reading instruction will be taken into consideration. All four schools represented a level of student complexity seen in many U.S. public schools. They had students of various economic, ethnic, primary home language, and experience backgrounds. Using this complexity in combination with cost-analysis procedures enhanced the understanding of how cost-effectiveness analysis can work in a variety of learning communities. Collecting both quantitative and qualitative information generated from survey methodology and extent student outcomes shed light on the critical variables needed for a clear cost-effectiveness analysis.

The variables outlined in Tables 3.0 through 6.0 typified those characteristics measured within reading programs research (Foorman & Torgesen, 2001; Kame'enui, 2001; Pinnell et al., 1994). These variables were the elements that decision-makers can relate to their current context and make both comparisons and cost estimates based on their experience. The cost variables in Table 7.0 delineate those found in a variety of reading programs as represented by *Success For All*, *Open Court* and *Reading Recovery* (Borman & Hewes, 2002; Pinnell et al., 1994; Slavin & Madden, 2001a). Bringing together both sets of variables in a cost-effectiveness analysis supports the decision-making process of educators seeking to answer questions about recommend program changes.

Measures

Reading Outcomes

Reading outcome data collection should include measures that are understood by reading professionals to show student growth in skills associated with success in acquiring early reading skills (Kame'enui 2001; Pikulski, 1994; Rayner et al., 2001). All of the programs under study have protocols in place for assessing student growth over the course of the year, as well as the district's approach to assessing students' end-of-year abilities. However, cost-effectiveness research purported that common outcome measures are the only accurate means of comparing alternative programs for comparative cost-effectiveness ratios (Hummel-Rossi & Ashdown, 2002).

One of the goals of a program evaluation was to help practitioners understand the workings of their environment, and using measures that they do not have access to limits the effectiveness of the recommendations that come from the findings (Levin & McEwan, 2001). Although the goal of this study was not to recommend changes to those involved, it would be appropriate to compliment their involvement with findings that may have practical value for their schools. Therefore, the data used to compare the alternative programs' reading outcomes supplemented the district's current data collection program.

The outcome variables of letter identification, letter sound identification, word recognition, book level score, reading fluency, and reading comprehension were used. The assessment protocol was designed by Rigby in their *Primary Literacy Assessment PM Benchmark Kits 1 and 2*, (Nelly & Smith, 1999; 2002), and both are currently in use by the district. This assessment was given two times, in the fall and again in the spring.

These six measures of student growth were analyzed. Letter sound identification and letter identification assessed student skills in phonemic awareness and phoneme grapheme representations. These attributes were seen by reading researchers (Rayner et al., 2001; Foorman et al., 1998) as critical onset skills to reading development.

Word recognition skills were the students' abilities to decode new words and use the skill of automaticity. This task required that the grapheme and phoneme systems be firmly established, and students could put letter pairs and blends together to make sense of the orthography (Rayner et al., 2001). Once a set of words and the background knowledge that goes with them was established, students could move through text with greater fluency which improved the comprehension of text. Comprehension took place when fluent readers used their background knowledge and onset reading skills to generate meaning from text.

Measuring these variables in a pretest to posttest format allowed for an understanding of how the four reading programs supported the growth of these critical reading skills. The growth of students in these variables could then be compared both within program and between programs to determine if one approach achieved greater results than the others.

School Decision-Making Practices

The local practices of staff decision-making were measured through the sampling of teacher, principal, and program manager perceptions of how their particular school interacted in the decision-making process. Site level decision-making can be viewed in the context of what leadership tasks were undertaken in terms of who carries out the task,

who initiates the tasks, and the degree of involvement by all participants (Gronn, 2000; Kruse, 2001).

Using a teacher survey of 45 questions, teachers' perceptions of the decision-making process, their involvement in the selection of the reading program, the nature of their reading program, and their approach to teaching the reading program were assessed. These survey variables were used throughout the study to answer research Questions 2 through 5.

Principal and program managers' interviews consisted of 27 questions and were used to generate information about leadership perceptions in a variety of elements. The interviews took approximately an hour and covered: (a) perceived importance and leadership skills in measuring costs, (b) decision-making practices, (c) assurances of program fidelity, (d) the nature of each school's students, and (e) the essential components of their reading program.

Cost-Analysis Measurement

The analysis of program costs used a mixed methods format. The ingredients methodology described above was used to identify the variety of cost variables associated with each of the programs. Ingredients methodology, at times called the resource cost model (Chambers, 1999), is a cost estimate approach. The method requires that the researcher understand the instructional alternatives being considered, or in turn be able to develop through interviews, surveys and document reviews an accurate understanding of the variety of expenditures (i.e., the ingredients) at work in the implementation of the alternatives under study. Levin (1983), King (1997), and Levin and McEwan (2001)

identify five categories of ingredients to be taken into consideration by evaluators. Table 7.0 represents those variables and the subcategories that help define them as: (a) personnel, (b) facilities, (c) equipment and materials, (d) unique inputs to the program and (e) required client inputs. These five categories are reviewed using school planning documents, Title I program documents, and district budget forms as well as interviews of program managers and principals.

Data Collection

Yin (2003) identified six potential sources of documentation for researchers seeking to understand the nature of personal behavior and the programs people interact with. In this study data sources included the school personnel through interviews with the researcher, a review of expenditure documents, and survey results of staff who interacted with the students in the program. This triangulated approach provided a means to link nonmonetary with monetary expenditures in an attempt to uncover the associated costs caught up in program implementation. These nonmonetary variables were seen by Grissmer (2002) to contribute to the opportunity cost variable that typically goes unnoticed in program implementation. It is these costs that are of the utmost importance to administrators making decisions with limited funds.

Data was collected in four primary areas: (a) school and student demographic variables, (b) student reading outcomes in grades 1 and 2, (c) professional perceptions of the school community in areas of student characteristics, program implementation fidelity, and decision-making, and (d) program costs as reported in school and district documentation.

Demographic and Reading Outcome Data

Student demographic data was retrieved from each school's database and checked with Oregon Department of Education website information defined in the measures section of this chapter. Demographics information was reported on schools kindergarten through fifth grade (Table 3.0), and on the sample subjects (Tables 4.0 and 5.0).

Collecting student data on reading outcomes for first and second grade students was based on two separate assessment periods, and consists of approximately 35 students from first grade and approximately 35 students from second grade using a simple random selection process from the available classes at each school. Trained assistants who are familiar with the schools' primary grade reading instruction and trained in the administration of the *Rigby's PM Benchmark 1 and 2 Primary Literacy Assessment (PLA)* protocol administered the assessments.

The first and second grade baseline assessment took place in the fall, gathering information on students' abilities to complete early reading outcome tasks. In the *PLA* format, if the student was a nonreader or struggled with books above book level eight, four to six variables were assessed: (a) letter identification, (b) letter sounds identification, (c) word recognition list and, if capable, (d) book level achieved with the corresponding (e) reading fluency, and (f) reading comprehension ratings. Students who read book level nine or above received scores for the book level they achieved and reading fluency and reading comprehension ratings (Nelly & Smith, 1999).

The spring outcome assessment took place in a similar manner in late May with reassessment of the same variables, using different reading passages developed by Rigby

in their *Primary Literacy Assessment PM Benchmark Kit 1* (Nelly & Smith, 1999).

Reading outcome data were associated with the student demographic information consisting of gender, age, ethnicity, enrollment date, absentee rate, free lunch status (a socioeconomic status proxy), and self-reported primary home language.

Leadership and Program Variables Data

In this study two forms of survey methods were used to access information about the leadership styles and decision-making process at each of the school sites. The first form was recorded interviews of program leaders (teacher leaders, program managers, and site principals) that encompassed budgetary, program fidelity, student interactions, and decision-making variables important to program evaluations. Second, a survey of primary grade teaching staff asked 45 questions about: (a) reading methodologies (b) instructional characteristics of the reading program, (c) participation in decision-making, and (d) staff demographic variables. These tools were used to understand the task functions of the classroom teachers and teacher leader or program manger roles—particularly, how they related to decision-making. Understanding the perceptions of administrators and teachers working in the programs allowed for the ground level perspective to be portrayed, as well as provided insights into questions that may not be easily answered in face-to-face interviews (Berg, 2004). This clearer picture, situated in daily practice, had the potential to support the decision-making process of elementary principals when confronted with change.

Cost Analysis Data

Table 7.0 defines the ingredients that make up the cost triangulation process in this study of four alternative reading programs based on the variables in five cost categories typically referred to in the literature.

Table 7.0

Ingredients Reviewed in Four Primary Grade Reading Programs¹

Ingredient / School	Adams	Franklin	Hamilton	Jefferson
<i>Facilities</i>				
New spaces	None	None	None	None
Remodel / upgrades	None	None	None	None
Furniture	None	None	None	None
<i>Personnel</i>				
Instructional Asst.	9.0 hours	8.0 hours	8.0 hours	14.0 hours
Specialist	.5 FTE	.5 FTE	.5FTE	1.0 FTE
Program auditors	1.0 FTE	0	0	1.0 FTE
Volunteerism	0	4.0 hrs/wk	0	0
<i>Equipment & Materials</i>				
Supplies	Not Kept	Not Kept	Not Kept	Not Kept
Curriculum products	Yearly	Some	Some	Yearly
<i>Client Inputs</i>				
Staff Development	Monthly	Yearly	Yearly	Monthly
<i>Unique Inputs</i>				
Miscellaneous	Contract Fees	PTA Support	PTA Support	Consultants Fees and Assessment

¹Adapted from King-Rice (1997).

The ingredients methodology organized the information collected from school and district budget documents, local school improvement plans, federal and state report formats, and principal generated expenditure documents. These variables allowed the reviewer to develop a qualitative and quantitative view of three important concepts in program evaluation: (a) the sensitivity of the instructional program in changing reading outcomes, (b) the nature of the distributed costs in the program, and (c) a clear representation of all program costs, both direct and indirect (Grissmer, 2002; Levin & McEwan, 2001; Yin, 2003). This review of documents was combined with administrators' and practitioners' surveys to compose a complete picture of the programs.

Assumptions about Subjects' Interaction with Three Reading Programs

Assumptions have been made about how students were included into the subject group. In an attempt to give each reading approach an appropriate amount of time to influence student outcomes, all students who had fewer than 135 days of attendance and interaction with the reading program will be removed from the analysis. Present at school for 135 days accounts for 70% of the school year, but 85% of the schools days up to mid-May in the programs under study. Students who moved from one school within the study to another school within the study and whose first reading program was dissimilar to their new school's reading program were dropped from the data pool.

A second assumption of the study was that special education students, with the exception of the Adams school special education students, receiving special education services for reading will be withdrawn from the sample groups. Special education students would have been receiving the reading program in class as well as supplemental

instruction from the special education teacher. The ability to determine which instructional setting was producing the desired effect would only cloud the results, and depending on the percentage of students in the school, these student outcome scores may skew the results for that particular school. Adams school special education students were maintained in the sampling because of the philosophical belief of Slavin and Madden (2001a; 2001b) that the SFA program adequately supports special education students within the regular education setting.

Students in bilingual schools who received their primary literacy instruction in a language other than English are not included in the sampling pool. This constrained and also biased some of the sampling at the two smaller schools. However, the assessment protocol used in this study was an English only version and did not allow for the assessment of non-English language instructional programs.

Lastly, students with missing assessment scores who could not be reassessed within a week of their peers' assessment were set aside in the analysis. This was due to a limitation of the research funds. An attempt was made to include absent students' assessment outcomes, but there were limitations to funding of this study.

Data Analysis

This study used descriptive statistics to examine the differences in costs and outcomes when comparing four reading instruction models. The dependent variable was the achievement of the students in the four programs derived from the previously described reading outcomes. In a pre-post design format a repeated measures ANOVA for independent samples allowed for a comparison of the sample populations' abilities on

the reading measures. This comparison was to assure that the study is looking at four comparable student populations. In an attempt to determine the direction and strength of significant *p*-values an ANCOVA measure was used to find the main effects of the four models as they compared the four sample populations' growth to each other. A posttest on the same variables allowed for a statistical analysis to determine differences in student growth on the collected reading measures. This process was used in relation to the independent variables to describe the type of reading program a student received and the demographic information collected (Borman & Hewes, 2002).

A second set of independent variables was the cost functions of the programs. A cost-effectiveness ratio was derived based on the dependent variables (whether the student reached grade level reading benchmarks) and the costs to produce the outcome. The unit of analysis for this study was at the student level. Levin and McEwan, (2001) and Chambers (1999) asserted that analysis at levels above the outcome (aggregated by grade or by school), used as a comparison, does not accurately link the program costs to the instructional change. A cost-effectiveness ratio was calculated based on student days of attendance as the cost function of the reading program in which the students participated as presented in Equation 1. Therefore, the cost per student day of receiving the program when compared to the growth of the students in the program can be seen as

$$C/E = \frac{\text{Cost per sample student per day}}{\text{Number of students meeting Benchmark}} \quad (1)$$

representing the program's ability to support students' success on the outcome measures. The cost estimates for values of costs per student per day was developed using an

ingredients methodology described by King-Rice (1997) and Levin and McEwan (2001) and presented in Table 7.0. The relative effectiveness of each of the programs under study was compared within the context of the qualitative variables to help decision-makers draw conclusions about the potential effectiveness of a program as it relates to their current context. Using survey results of individuals involved in the program's measures of distributed cost typically associated with task functions and leadership functions described by Spillane, Halverson, and Diamond (2001), Gronn (2000), and Weick (1975) further clarified leadership parameters that should be considered when making program changes.

To present an accurate representation of the expenditure value as a daily cost, the samples and their costs must be extracted from the whole group prior to cost analysis. This includes the cost of teaching reading. This extraction represented the true costs-effectiveness ratio of the students who participated in the study. A series of steps helped to parse out the data into a usable form for accurate comparison. Values for the following variables were determined: (a) the percentage of the total grade level (first or second) that the sample represents in the school (represented as S_1 , S_2 for each school), (b) the percentage of the total expenditure the sample population used throughout the study (represented as SPC_1 , SPC_2), and (c) the average daily cost for the sample group (represented as PDC_1 , and PDC_2). For the purpose of assigning costs to grade levels the program total costs were divided using an average daily membership formula to determine which portion of the aggregate costs should be assigned to first grade and which to second grade.

The following equations were used as a means to prepare the sample population data for use in the development of cost-effectiveness ratios. Equation 2 defines the percentage of the sample students representing first grade (subscript 1).

$$\frac{\text{Sample } n}{\text{Total } N \text{ of grade level}} = S_1. \quad (2)$$

S_1 (or S_2 depending on grade level) was the percentage of the total students in the sample represented for that grade level. This value was then applied to the total costs variable for the grade level as defined for each school as represented in Equation 3 where SPC_1 is the samples portion of the reading program's total costs.

$$(S_1) \times (\text{School's total costs}) = SPC_1. \quad (3)$$

Using the value SPC_1 the program costs for the sample groups as its share of the total resources to teach reading is defined. SPC_1 then became the value by which the average daily attendance was divided (Equation 4) to derive the cost per day per sample student, SDC_1 .

$$\frac{SPC_1}{\text{Sample Mean Average daily attendance}} = SDC_1. \quad (4)$$

Equation 5 then represents the formula for computing the cost-effectiveness ratio for the sample population:

$$C/E = \frac{SDC_1}{\text{Number of sample students meeting benchmark.}} \quad (5)$$

The value resulting from Equation 5 gave the cost-effectiveness ratio, or the expenditure to outcome attained value for the program and the resources used to get the

students who reached benchmark to that point, given only the economic factors under consideration.

Summary

The evaluation of four reading programs was undertaken for two primary purposes: (a) to understand which of the four reading programs supported early reading skill acquisition in the most cost-effective manner, and (b) to determine how the school leadership acted in the decision-making process when selecting the program. These two purposes involved applying the principles of cost analysis relying on common reading outcome measures, surveying program participants for their perceptions of the programs, and reviewing a variety of cost documents.

To accomplish this, a mixed methods format was used to generate data about the four schools. The data informed the researcher about how the adults viewed their decision-making structures and what program evaluation skills were present in the leadership staff.

An analysis of variance was used to compare the four programs on their outcomes. This information was then used to look at the daily costs of the programs in relation to the outcomes achieved. This produced a cost-effectiveness ratio for each of the four programs that were used as a comparison. The qualitative data were reported in descriptive format to allow the practitioners' voice to highlight their impression of the use of cost-effective techniques and how they as school leaders interact in the environment of pressures from outside agencies and constituents, and deal with the complex nature of their school's community.

CHAPTER IV

RESULTS

This chapter presents the statistical and the qualitative interview results for each research question posed in Chapter I. The main research questions are: (1) Was there a difference in the reading outcomes of the four reading programs under review; (2) Of the four reading programs, which one had a better costs-efficiency ratio; (3) What did schools' teachers, program managers, and principals use as their decision-making process when programs were implemented; (4) How did the student population's complexity contribute to the reading outcomes students achieved, and (5) Did site managers feel the reading programs in place were being used as intended by the program developers?

Descriptive Statistics

The descriptive statistics presented in the following tables utilize the subjects described in Chapter 3. The tables present the variables associated with primary grade readers in public school settings. The outcome variables are scores on six *Rigby PM Benchmark Assessment forms 1 and 2* subtests given in a pretest and posttest format. The reading outcomes were reported in book level scores for individuals above book level nine in either first or second grade students. Students capable of book level nine and above received a *book level read* score, a fluency score, and a comprehension score. To achieve a *book level read* score a student must have read the book successfully and received a fluency and comprehension rating of 3 or higher. For those individuals

below book level nine in the *Rigby PM Benchmark* protocol three areas of reading were assessed: (a) letter identification (52 possible), (b) letter sound identification (20 possible), and (c) word recognition (20 possible). These scores were then recorded along with the book level, scores. These data are presented in Table 10.0 through Table 13.0 as a basis for answering Question 1 and set the criteria for evaluating the cost-effectiveness ratios in Question 2.

Along with the statistics of the subject students' outcomes, qualitative survey data are presented. A 45-question survey of teachers and interviews with building principals and program managers or teacher leaders were utilized. These tools generated coded data regarding teacher, program manager, and principal perceptions of their relationship to: (a) the decision-making process for the reading program, (b) their understanding of the costs associated with the reading program they used, and (c) the practitioners' perspective of their program's characteristics. The data are presented in Tables 21.0, 22.0 and 23.0, and were used in the analysis of Questions 2 through 5 of the study.

In four schools' comparative research it is important to understand whether the populations within the sample were equal in the demographic variables identified in the research to be associated with reading skills acquisition. In this study it was important to assure that all four groups had similar distributions for age, gender, ethnicity, primary home language, free and reduced lunch status, and attendance rates. Table 3.0 displays schoolwide demographic information. Table 4.0 and Table 5.0 present the demographic information for the sample populations upon whose reading outcomes the cost-effectiveness analysis was calculated.

Research Question 1

First Grade Demographics

The demographic information of the sample subjects is presented for two grade levels. The sample of first grade students at the four schools is comprised of 119 students, both male and female. Table 8.0 shows that differences do exist between the first graders in the four schools. The frequency distribution for free and reduced lunch eligibility rate (socioeconomic proxy variable) was the most observable. Adams school uses the reading program *Success For All*, and Jefferson uses the *Reading First* protocol; each had a free and reduced lunch rate of 79% and 81% respectively. The two schools using their own locally designed models had free and reduced lunch rates at or near 34%. These percentages align with the general school variables with the exception of Hamilton whose schoolwide free and reduced lunch rate was 41% and the sample populations for first grade was lower at 34%.

The second variable that distinguishes the four schools was whether English was the primary home language. Jefferson school had only 8 of the 31 students (24%) who spoke English as their primary language. No other school, not even the *Success For All* school that has a dual language reading program, came close to this one-in-four ratio.

Table 8.0 includes statistics on two variables that previous research has shown to effect reading outcomes and cost-analysis equations. First was the age of students, especially if the sample contains retained students who may have been exposed to the curriculum in previous years. The second was the average days of attendance as a factor

Table 8.0

Grade 1 Demographic Information for Sampled Students in the Four Participating Schools

School	Gender		Age		Free or Reduced Lunch		English as Primary Language		Average days of Attendance of 163 days		Ethnic Code ^a							
	n	F	M	Mean	SD	No	Yes	Yes	No	Mean	SD	1	2	3	4	5	6	9
Adams	24	11	13	5.54	.31	5	19	20	4	148.75	14.42	--	2	--	7	13	1	1
Franklin	32	17	15	5.49	.33	21	11	31	1	155.11	5.50	--	2	2	2	28	5	3
Hamilton	32	17	15	5.53	.27	20	12	31	1	149.94	15.21	--	--	--	1	24	4	3
Jefferson	31	13	18	5.49	.32	6	25	8	23	151.06	11.98	--	9	--	10	11	1	--
Total	119	58	61	--	--	52	67	90	29	--	--	0	13	2	20	66	11	7

^aEthnic Codes: 1 = American Indian / Alaskan Native; 2 = Asian / Pacific Islander; 3 = African American; 4 = Hispanic; 5 = White; 6 = Multiracial; 9 = No Response.

in figuring cost-effectiveness ratios. Comparing these two measures across the four schools showed no significant differences. The age of the first grade students was not statistically different ($p = .90$) for the four schools sampled. The average days in attendance at the four school was also found not to be statistically different ($p = .21$). These two measures confirm that the students are similar in age and had equal amounts of exposure to the reading programs at their schools to use a pretest to posttest format with reliability.

Second Grade Demographics

The sampling of students in the second grade was similar to that of the first grade. Table 9.0 presents the statistics for the second grade sample. As with the first grade, age ($p = .13$) and attendance ($p = .27$) variables were not statistically different. A difference in the socioeconomic proxy variable free and reduce lunch was worth noting. The rates at Jefferson and Adams schools were the same percentages as the first grade students. However, Hamilton's sample subjects' two in three (63%) ratio more closely approximates the two Title I and bilingual schools than it does its own schoolwide percentage of 41%.

As with the first grade sample, one in every three second grade students at Jefferson spoke a primary language other than English at home, a ratio that was a little lower than the first grade ratio, but substantially higher than that of the other three schools.

First Grade Reading Outcomes

First grade reading outcome measures are presented in Table 10.0 for the variables of letter identification, letter sound identification, and word recognition. These skills were measured for those students who cannot read a level nine book with fluency and comprehension. In comparing the four programs the *letter identification* (Letter ID) measure was found to be statistically significant in the initial analysis of variance (ANOVA). Higher scores and a low standard deviation at Franklin ($M = 50.43$, $SD = 3.80$) and lower mean scores at Adams ($M = 41.14$, $SD = 14.70$) showed a broader variance amongst the students. Using the pretest as a covariant with the posttest score on Letter ID and ANCOVA analysis was utilized to determine if the growth in both sets of students could identify one model as significantly more effective than the other. However, the cell count for the posttest samples was too small (11 and fewer in the posttest sample) to say with certainty that a difference in the growth attained existed between the two groups.

In the analysis of letter sound identification and word recognition no significant differences were found in the pretest scores. A repeated measures ANOVA was used to determine differences in mean scores. The process produced no significant p -values associated with any of the four schools' programs. Therefore, there is a lack of interaction between the schools' programs and the pretests. When comparing the means tables values for the measures in the school-to-school comparison there are hints of main effects, but again the small cell counts for the ANCOVA make reporting of the results of the interactions spurious.

Table 9.0

Grade 2 Demographic Information for Sampled Students in the Four Participating Schools

School	Gender		Age		Free or Reduced Lunch		English as Primary Language		Average days of Attendance of 163 days		Ethnic Code ^a							
	n	F	M	<i>M</i>	<i>SD</i>	No	Yes	Yes	No	<i>Mean</i>	<i>SD</i>	1	2	3	4	5	6	9
Adams	24	10	14	6.69	.39	6	18	21	3	152.56	9.80	--	1	--	7	13	1	2
Franklin	35	19	16	6.52	.30	24	11	34	1	155.49	5.27	1	2	1	4	23	3	1
Hamilton	34	20	14	6.49	.30	12	22	31	3	154.99	7.60	--	1	--	4	25	4	--
Jefferson	32	13	19	6.52	.33	10	22	12	20	152.58	7.74	--	4	--	11	14	3	--
Total	125	62	63	--	--	52	73	98	27	--	--	1	8	1	26	75	11	3

^aEthnic Codes: 1 = American Indian / Alaskan Native; 2 = Asian / Pacific Islander; 3 = African American; 4 = Hispanic; 5 = White; 6 = Multiracial; 9 = No Response

Table 10.0

Fall to Spring Primary Literacy Assessment Mean Scores for Grade 1 Students in Four Primary Reading Programs

Schools	Letter Identification						Letter Sound ID						Word Recognition					
	Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>
Adams	21	41.14	14.70	11	48.45	3.08	21	11.36	6.87	11	15.50	6.07	21	3.82	2.75	11	7.27	4.98
Franklin	23	50.43	3.80	7	51.14	.90	23	13.57	5.41	7	18.29	3.25	23	2.14	2.19	7	11.86	4.18
Hamilton	23	42.78	13.48	11	47.27	4.80	23	10.55	6.17	11	16.64	3.07	23	2.09	2.43	11	6.09	4.83
Jefferson	24	47.88	5.20	10	49.50	4.20	24	12.70	5.40	10	18.30	1.64	24	2.60	2.07	10	10.40	6.28

¹p < .05. ²p < .01. ³p < .001

Taking a broader look at the sample population in terms of mean book levels achieved from pretest to posttesting in each of the programs and the corresponding fluency and comprehension measures allows for analysis of larger cell counts. Table 11.0 presents student reading abilities measured by pretests and posttests and improves the understanding of the four models involved. The means table for book level, reading comprehension and reading fluency all show program differences that are significant in the first grade sample.

When a repeated measures ANOVA was used the means tables show an effect for all schools ($p = .001$). In particular Franklin's ($M = 17.00, SD = 7.85$) growth on average of 9.94 book levels is significantly higher ($p = .001$) than the growth made by Jefferson ($M = 15.42, SD = 9.36$). This main effect was not universal among the programs and became evident only when the repeated measures compared schools' means scores in a Benferroni/Dunn analysis. Using this technique, the main effect of the pretest and posttest growth is seen to be statistically significant ($p = .03$) when comparing book level growth between Franklin and Jefferson.

Continuing with reading fluency, a repeated measure ANOVA was used to determine significant differences in first grade reading fluency among the reading programs. A significant interaction ($p = .05$) was found between the schools using locally designed models and the more structured programs based on the pretest to posttest scores. Both of the locally designed programs at Hamilton ($M = 3.95, SD = .92$) and Franklin ($M = 4.04, SD = .20$) produced mean scores on the posttest assessment that outpaced those of Jefferson ($M = 2.45, SD = 1.80$) over the length of the study ($p = .003$).

Table 11.0

Fall to Spring Primary Literacy Assessment Mean Scores for Grade 1 Students in Four Primary Reading Programs

Schools	Book Level						Reading Fluency						Reading Comprehension					
	Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>
Adams	14	3.25	6.58	24	11.63	9.26	5	.85	1.63	14	3.77	.83	5	.92	1.75	14	4.23	.44
Franklin	27	7.06	6.55	32	17.00 ¹	7.85	10	1.32	1.84	25	4.04 ²	.20	10	1.60	2.22	25	4.24 ²	.52
Hamilton	10	6.31	8.45	30	14.66	10.79	10	1.81	2.20	21	3.95 ²	.92	10	1.86	2.22	21	4.52 ²	.68
Jefferson	24	6.52	8.27	31	15.42	9.36	8	.81	1.54	20	2.45	1.80	8	1.03	1.96	20	2.77	1.98

¹p < .05. ²p < .01. ³p < .001.

The third variable in the analysis of reading gains was the interaction of reading comprehension gains associated with any of the four programs. There was no statistical difference in the pretest scores across programs. The use of the repeated measure ANOVA shows a similar set of results to that found for reading fluency. No interaction effect was found but significant main effects similar to those found in reading fluency were found between pretests and posttest mean comprehension scores in the program-to-program comparisons. Hamilton comprehension growth from pretest ($M = 1.86$, $SD = 2.22$) to posttest ($M = 4.52$, $SD = .68$) and Franklin's growth from pretest ($M = 1.60$, $SD = 2.22$) to posttest ($M = 4.24$, $SD = .52$) were significantly different ($p = .004$) to the growth attained by the students at Jefferson from pretest ($M = 1.03$, $SD = 1.96$) to posttest ($M = 2.77$, $SD = 1.98$).

Second Grade Reading Outcomes

The analysis of the second grade student reading data followed the same analysis format as described for the first grade. Of note is the fact that the variables of letter identification, sound letter identification, and word recognition are skills that second grade teachers anticipate students have mastered prior to the end of the first grade. However, it was not uncommon to have entering second grade readers who have yet to master these skills. Therefore, the second grade sample did contain small numbers of students who at the pretest were reading below book level nine and therefore had scores for these three variables. An even smaller number also had scores in these variables in the spring posttest.

Table 9.0 presents the demographic data for these second graders. Table 12.0 presents data on students who read below book level nine and had scores for letter identification, sound letter identification, and word recognition. Table 13.0 contains the data on second grade students who read above book level nine and had scores for reading comprehension and fluency as well. The cell counts for letter identification, sound letter identification, and word recognition variables were too few to allow for parametric statistics to yield reliable results. Of note is Hamilton's gain from the pretest to posttest for students below book level nine. Three schools moved from cell totals of five to six students to three students at the posttest. Hamilton began the year with 11 students reading below book level nine at second grade and ended the year with one student in that cell grouping.

The analysis of book level growth and improvement in reading fluency and reading comprehension relied upon the reading outcomes of the four reading programs. Using the *Rigby PM Benchmark 1 and 2* materials, student reading levels, fluency, and comprehension skills were assessed in all four programs. Table 13.0 presents the pretest to posttest for second grade. In the analysis of *Book Level* there is a significant difference in the pretest book level scores in the four sample schools ($p = .05$). Although all four programs' students did well, Hamilton's book level pretest scores ($M = 13.48, SD = 7.82$) are significantly lower than Franklin's pretest scores ($M = 18.37, SD = 7.35$) which is statistically significant at ($p = .05$). to account for these differences the pretest was used as the covariant, a significant posttest significant main effect was found ($p = .008$). Franklin's book level growth was seen in posttest mean scores ($M = 23.69, SD = 6.14$)

Table 12.0

Fall to Spring Descriptive Statistics Mean Scores for Grade 2 Students Reading in Four Primary Reading Programs

Schools	Letter Identification						Letter Sound ID						Word Recognition					
	Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>
Adams	6	41.50	20.50	3	51.67	.58	6	14.00	5.35	3	18.33	2.08	6	8.00	6.75	3	13.33	1.53
Franklin	5	50.40	.89	0	xx	xx	5	18.40	1.67	0	xx	xx	5	12.80	2.28	0	xx	xx
Hamilton	11	51.55	.82	1	49.00	xx	11	18.36	1.21	1	18.00	xx	11	15.09	6.33	1	13.00	xx
Jefferson	4	40.75	14.77	3	45.00	6.24	4	13.67	7.17	3	16.00	3.61	4	11.50	8.10	3	13.00	5.29

¹p < .05. ²p < .01. ³p < .001.

Table 13.0

Fall to Spring Primary Literacy Assessment Mean Scores for Grade 2 Students in Four Primary Reading Programs

Schools	Book Level						Reading Fluency						Reading Comprehension					
	Pretest			Posttest			Pretest			Posttest			Pretest			Posttest		
	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>
Adams	26	15.33	8.57	24	19.08	6.02	21	3.89	.68	22	4.11	.58	21	4.28	.46	22	4.28	.46
Franklin	35	18.37 ²	7.35	35	23.69 ³	6.14	30	3.53	.63	35	4.03	.56	30	4.00	.26	35	4.13	.63
Hamilton	34	15.33	7.82	34	20.29	6.69	23	3.78	.80	34	3.78	.52	23	4.09	.29	34	4.04	.21
Jefferson	30	17.25	7.06	32	19.69	6.20	30	4.21	.63	32	4.25	.44	30	4.11	.31	32	4.21	.42

¹p < .05. ²p < .01. ³p < .001.

and was significantly higher than that of Adams ($M = 19.08$, $SD = 6.02$), Hamilton ($M = 20.29$, $SD = 6.69$), and Jefferson ($M = 19.08$, $SD = 6.20$).

When scores for reading comprehension were analyzed in the same manner a significant difference was found ($p = .005$) for programs. Using the pretest as the covariant, no significant differences existed ($p = .61$). The same analysis was applied to the measure of reading fluency. An ANOVA showed significant program effect ($p = .031$). Primarily this difference was between Franklin school ($M = 3.53$, $SD = .63$) and Jefferson school ($M = 4.21$, $SD = .63$). However, when the pretest was used as a covariant there was no significant main effect between Franklin and Jefferson, or for any of the programs in the analysis of reading fluency and reading comprehension ($p = .08$).

Research Question 1 Summary

The results for Research Question 1 showed differences in the areas of primary home language and free and reduced lunch percentages among the four schools. However, all schools had comparable mean ages and comparable attendance patterns. Mean scores in beginning books levels, letter identification, letter sound identification, and word recognition were also comparable for both grades.

The ANOVA analysis of the first grade scores show no mean difference in five of the six variables measured by the *Primary Literacy Assessment*. The pretest of letter identification scores for first grade students at Franklin school were significantly higher than for the three other schools. However, this does not hold true in the posttest for letter identification. The posttest differences were associated with the three variables of reading

fluency, reading comprehension and, book level achieved. Scores in fluency and reading comprehension were significantly different in the two locally designed programs. Also Franklin's posttest book level read score significantly outpaced the scores of the other three schools.

The analysis of the second grade scores showed that Franklin pretest book level scores were significantly higher than both the Adams and the Hamilton schools' scores. The difference is maintained in the posttest book level scores as well. An ANCOVA analysis was used to determine if the growth at Franklin was still significantly different despite the students' starting points in the fall. The analysis showed that the locally designed model was significantly stronger than the two more structured protocols. Using pretest and posttest scores, the four programs showed different rates of gain in book level growth in second grade.

Therefore, differences in the four reading programs' ability to generate student growth did exist. Franklin school made stronger growth as the year progressed as measured by mean book level scores. The interaction main effects did not confirm whether it is due to the instructional programs or the nature of the students in the school. These complexity factors will be looked at in Research Question 4. The high percentage of students whose families speak a language other than English could be affecting first grade and some second grade students, but this may not completely explain the difference in the book level scores.

Research Question 2

Research Question 2 investigated the resource cost variables as reported by the program managers, principals, and the schools' and district's documentation. These descriptive values were then used in conjunction with the outcomes presented in Research Question 1 to generate a cost per student reaching grade level reading benchmark. This cost was then used to show the programs cost-efficiency in supporting students' growth towards meeting the district's reading benchmark for grades 1 and 2.

Resource Costs Data

At the center of this question is the quantification of the resource cost variables associated with the four programs as presented in Table 14.0. Program managers and principals' interview responses, school improvement plans, and district budgeting documents were used to determine program cost variables. These costs as reported by the program participants made up a large percentage of the overall program costs. However, in this analysis the opportunity costs and distributed costs were not included in the program evaluation. Opportunity cost concepts were set aside in the four programs because all four of the programs lacked documentation, or had no real remembrance of the past costs of the reading program. Distributed costs were minimized by the participants in their statements, so accurate experiences could not be gathered to generate estimated costs. Table 14.0 defines the four variables: (a) personnel, (b) client inputs, (c) materials and equipment, and (d) unique inputs most often recorded by the schools, either verbally or in their budget and expenditure documentation.

In the interview process the resource cost variable classified as *facilities* had minimal differentiated cost for the year, meaning that in the year of the study facilities expenditures were not a cost that the four schools incurred. Although the program managers and the principals had made some facilities related purchases over the four years

Table 14.0

Program Costs Variables by Resource Type in Four Primary Reading Programs

Adams Program Cost Variables			
Personnel	Client's Inputs	Equipment/Materials	Unique Inputs
Certificated Staff (1.65 FTE):	Staff Development:	Reading curriculum	SFA contract = \$1,660.00
1.0 FTE Program Coach \$59,700.00	Early Release Days: 15 days X 3.0 hours training @ \$900.00 / trainings = \$13,500.00	consumables at \$2,500.00 for Roots.	Parent Involvement \$1,960.00
.5 FTE Federal Reading Teacher \$38,600.00	New Teacher Training w/ SFA Consultant 2 days @ \$3,600.00		
.144 PE Conversion <u>\$11,560.00</u> \$109,860.00			
Instructional Asst. (9.5 hours): 3.0 hours Title I 2.5 hours CIF 4.0 hours Building \$41,733.00			
\$151,593.00	\$17,100.00	\$2,500.00	<u>\$3,620.00</u>
Total Costs			\$174,813.00

Table 14.0

Program Costs Variables by Resource Type in Four Primary Reading Programs (cont'd.)

Jefferson Program Cost Variables			
Personnel	Client's Inputs	Equipment/Materials	Unique Inputs
Certificated Staff: (2.5 FTE): 1.5FTE – <i>Reading</i> <i>First Grant</i> Program Coaches: \$119,400.00 .5 FTE Federal Reading Teacher \$38,600.00 .5 FTE Spring Support Teachers: (1/4 th of Year) <u>\$19,300.00</u> \$177,300.00 Instructional Asst. (9.5 hours total): 4.0 hours Grade 1 (Building) 4.0 hours Grade 2 (RF Grant) \$35,114.00 6.0 hours RF Grant (1/2 year) <u>\$26,433.00</u> \$61,547.00	Staff Development RF Grade Level Meetings: \$15,600.00 RF Early Reading Team Meetings: \$3,840.00 RF Coaches Training: \$550.00 RF Grant Trainings: \$18,070.00 Title I Reading Support: \$6000.00 Instructional Asst. Trainings: \$7,000.00 New Teacher Training w/ <i>SFA</i> Consultant 2 days @ \$3,600.00	RF Grant Training Materials: \$800.00 RF Grant: Grade 1 RW materials: \$4,750.00 Grades 1 & 2 RM Materials: \$750.00 Teach Take home Reading: \$2,400.00 <i>Scott-Foresman</i> Grade 2 materials: \$1,750.00 Supplemental Phonics: \$2,500.00	RF Consultants: \$24,200.00 RF Assessment Protocols: \$3,501.00 Parent Involvement \$2,292.00
\$238,847.00	\$54,660.00	\$12,950.00	<u>\$29,993.00</u>
Total Costs:			\$336,450.00

Table 14.0

Program Costs Variables by Resource Type in Four Primary Reading Programs (cont'd.)

Franklin Program Cost Variables			
Personnel	Client's Inputs	Equipment/Materials	Unique Inputs
Certificated Staff (.5 FTE): .5 FTE Federal Reading Teacher \$38,600.00	Staff Development: Early Release Day: 1 day @ 3.0 hours of training	Leveled Books -- \$2,000.00 New SF Basal Set \$675.00	PTA Resources \$258.00/teacher = \$1,806.00 Volunteer Hours: 60 minutes /day at IA rate: 28 hrs / wk by 11.50/hr = \$9,660.00
Instructional Asst. (8.0 hours): 4 - hour Building 4 - hours CIF \$35,114.00			
\$73,714.00	\$900.00	\$2,675.00	<u>\$11,466.00</u>
Total Costs:			\$88,755.00
Hamilton Program Cost Variables			
Personnel	Client's Inputs	Equipment/Materials	Unique Inputs
Certificated Staff: .5 FTE Federal Reading Teacher \$38,600.00	No Current year Costs	New WG Book Sets @ 3,500.00	Volunteer Hrs: 2.0 hours after school for six weeks with 4 people: \$552.00
Instructional Assist. (8.0 hours): 4 - hours Building 4 - hours CIF \$35,114.00			PTA Resources for Reading – \$2,000.00.
\$73,714.00	\$0.00	\$3,500.00	<u>\$2,552.00</u>
Total Costs:			\$79,766.00

of the programs (average length of existence of the programs), they had not made large purchases in the year of the study. The variables included in the tables were those defined by the participants and in schools documentation as items that most influenced the costs of their programs. The costs of print, paper, other non-reading program books, and school related consumables were not included in the analysis as none of the four program sites collected information at that level. In comparison to the items referenced in Table 14.0, the principals and program managers felt that these were small budgetary expenditures.

The largest resource cost in each of the four programs was the personnel costs incurred. The personnel expenditure values in this study did not include the regular classroom teacher salaries. The reading programs personnel costs range from 92% of the programs costs in one of the locally designed models to 71% at Jefferson, the school with the greatest degree of outside program support. The second highest category varies in each of the four models such that no recognizable pattern can be seen.

Table 15.0 shows the spending differentials of the four programs. The *Reading First* school resource cost variables were anywhere from 1.92 to 4.22 times greater than the other programs in the study. The structured programs (Jefferson and Adams) were consistently more expenditure intensive than the two programs that relied solely on district inputs. Looking at the interview results, Table 14.0 provides a comparison of differences in the four programs' costs. The greater costs seen in structured programs were based in personnel, training and consultations, and staff development variables. This cost comparison showed that the staff training intensity and teacher support activities of the *Reading First* and *Success For All* protocols made up significant expenditures. The

corresponding expenditures for the two locally designed programs relied on site-based resources.

Table 15.0

Comparative of Total Cost of Four Primary Reading Programs Resource Cost Variables

School / Total Program Costs	Adams \$174,813.00	Franklin \$88,755.00	Hamilton \$79,766.00	Jefferson \$336,450.00
Adams \$174,813.00	--	+1.97	+2.19	.52
Franklin \$88,755.00	.51	--	+1.11	.26
Hamilton \$79,766.00	.46	.90	--	.24
Jefferson \$336,450.00	+1.92	+3.79	+4.22	--

Total costs. Using district and school documentation, total program costs were aggregated for the four resource strands. These totals are presented in Table 14.0 and represented as differentials in Table 15.0. Total costs are the basis for generating the daily cost of providing each reading program. The total cost value was presented as the sum of the four resource variables measured for each grade level when the total cost of the program is distributed using average daily membership (ADM). Total costs were then divided by the grade level ADM to determine the cost of the grade level's program. Using the sample populations (S_1 for first grade) as a percentage of grade levels' ADM provided the total cost of services for the sample students (SPC_1 for first grade). This cost was then

turned into a sample population average daily cost (SDC_1); this value was used to compare programs in a cost-effectiveness ratio.

As described in Chapter 3 the sample total program costs (SPC_1 and SPC_2) were needed to make comparisons of the alternative programs in this study. Using Equations 1 through 5 from Chapter 3, Table 16.0 represents the first grade's sample total program costs (SPC_1), daily program costs (SDC_1), and the annual daily program costs (SDC_{1A}). Table 17.0 represents these same values for the second grade.

The cost-effectiveness of each program was calculated based on the sample students in the study over two time periods. First a value is generated for the sample students' mean attendance rate over the 163 days, from the start of the year to the posttest for book level ($SPC_1 \div M$ attendance = SDC_1). This value represented the program cost for

Table 16.0

First Grade Means Days Attended for Cost per Day of Program Implementation

School / Grade level program cost			Sample total program cost	Attendance		Sample daily cost	Annual sample daily cost
	<i>n</i>	<i>S_I</i>		<i>Mean</i>	<i>SD</i>		
			<i>SPC₁</i>			<i>SDC₁</i>	<i>SDC_{1A}</i>
Adams \$81,960.00	24	.40	\$32,784.00	148.75	14.42	\$220.00	\$191.00
Franklin \$49,410.00	32	.30	\$14,823.00	155.11	5.50	\$96.00	\$87.00
Hamilton \$40,469	32	.46	\$18,615.00	149.94	15.21	\$124.00	\$109.00
Jefferson \$187,487.00	31	.28	\$52,496.00	151.06	11.98	\$348.00	\$307.00

S_I is the percent the sample represents of the total first grade students.

Table 17.0

Second Grade Mean Days Attended for Cost per Day of Program Implementation

School / Grade level program cost	Sample total program cost		Attendance Days		Sample daily cost	Annual sample daily cost	
	<i>n</i>	<i>S</i> ₂	<i>SPC</i> ₂	<i>Mean</i>	<i>SD</i>	<i>SDC</i> ₂	<i>SDC</i> _{2A}
Adams \$92,869.00	24	.35	\$32,504.00	152.58	7.74	\$213.00	\$190.00
Franklin \$39,302.00	35	.41	\$16,113.00	155.49	5.27	\$104.00	\$94.00
Hamilton \$39,329.00	34	.74	\$29,103.00	154.99	7.60	\$188.00	\$170.00
Jefferson \$148,975.00	32	.34	\$50,651.00	152.58	7.74	\$332.00	\$296.00

*S*₂ is the percent the sample represents of the total second grade students.

the grade level to support the students in reaching their posttest book level reading fluency, and comprehension scores. The second value was the annual daily program cost (*SDC*_{1A}) using 171 days which is the length of the school year ($SPC_1 \div 171 = SDC_{1A}$).

Utilizing Equation 6 to define the value generated for *SDC*₁ (or *SDC*₂ depending on grade level) a cost-effectiveness ratio was generated by dividing the number of sample students who met the district's benchmark into *SDC*₁. To calculate the cost effectiveness ratio the following formula was used with the data derived from previous tables:

$$\text{Cost Effectiveness} = \frac{SDC_1}{\text{Number of sample students meeting benchmark.}} \quad (6)$$

The second grade cost effectiveness ratios were figured in the same format using the value SDC_2 from Table 17.0. However the second grade sample had less attrition over the length of the study. This in turn provided greater differential in the ratios using the equation presented above.

At first grade the four schools' pretest percentages of students at or below reading level were statistically equal. The initial data showed that Jefferson and Franklin schools were equivalent in student reading performance categories; however, the cost ratios were quite different. Tables 18.0 and 19.0 identified the cost per student ratios (cost-effectiveness ratios) for first and second grade sample students reaching benchmark. The values for Adams and Jefferson schools were appreciable higher than those of the locally designed programs.

Table 18.0

First Grade Cost-Effectiveness Ratios for Students Meeting District Reading Benchmark

Sample total program cost	Sample total	n Meeting benchmark	Sample daily cost	Cost Effectiveness Ratio	Sample annual daily cost
SPC_1	n		SDC_1	C/E	SDC_{1A}
Adams \$32,784.00	24	9	\$220.00	\$24.00	\$191.00
Franklin \$14,823.00	32	18	\$96.00	\$5.00	\$87.00
Hamilton \$18,615.00	32	15	\$124.00	\$8.00	\$109.00
Jefferson \$52,496.00	31	19	\$348.00	\$18.00	\$307.00

Table 19.0

Second Grade Cost-Effectiveness Ratios for Students Meeting District Benchmark

Sample total program cost	Sample total	<i>n</i> Meeting benchmark	Sample daily cost	Cost Effectiveness Ratio	Sample annual daily cost
SPC_2	n		SDC_2	C/E	SDC_{2A}
Adams \$32,504.00	24	16	\$213.00	\$13.00	\$190.00
Franklin \$16,113.00	35	28	\$104.00	\$4.00	\$94.00
Hamilton \$29,103.00	34	19	\$188.00	\$8.00	\$170.00
Jefferson \$50,651.00	32	21	\$332.00	\$15.00	\$296.00

The level of resources used by the two structured schools raised the first grade per pupil meeting benchmark costs and produced a cost-effectiveness ratio over three to four times the costs. The cost of getting a child at Jefferson to benchmark was three times greater than times that of the nonstructured programs. In other words, the cost to support a first grade student to meet benchmark status at Adams was four and one-half times greater than the cost of getting a child to benchmark at Franklin, or three times greater than Hamilton's that of Franklin or twice as expenditure intensive as Hamilton's program. The cost of getting a child at Jefferson to benchmark was three times greater than at Franklin or twice as expenditure intensive as Hamilton's program.

At second grade a similar finding existed. For Adams to get 25% fewer children to district benchmark their program costs was four times more costly than Franklin and

marginally more costly than Hamilton. Jefferson showed similar challenges, spending four times more per child reading at benchmark than Franklin and 50% more than Hamilton. Therefore, looking strictly at resource cost variables, the locally designed models were more cost-effective than the structured programs; both schools got more students meeting benchmark status at significantly less program cost.

Research Question 2 Summary

When locally designed and outside foundation or state agency programs were reviewed with an eye on cost estimation, the findings were not easily categorized. Principals and program manager interviews along with the local documentation they supplied went into an ingredients model (also known as a resource cost model) analysis. The ingredients of the programs showed that schools did not spend much if any money on facilities, nor did they keep good records on small expenditures like printing and consumable supplies. The variables that tended to differentiate the programs were required elements of reading programs including: (a) increased intensity in teacher training, (b) additional personnel for program oversight, (c) various student assessments, and (d) additional personnel for specific intervention and consultation tasks. These variables made the programs work for students, but drove up the program cost-effectiveness ratios.

Structured programs started with students in statistically equal book level standings as the locally designed programs in this sample of students; yet advanced between 12% and 22% fewer students to grade level standards in first and second grade. The cost ratios show that the *Success For All* and *Reading First* initiative programs used three to five times more resources to make these gains than the locally designed models.

These costs typically were personnel and training costs that directly impact instructional delivery. The costs were just one factor in the analysis of reading programs. In this light a look at the decision-making process and the use of cost analysis outcomes by school explained more about the current leadership context.

Research Question 3

Using elements from the teacher survey, along with the interviews of program managers and principals, the third research question attempted to discern the knowledge and use of program evaluation techniques within each school's decision-making framework. The degree to which site managers, both principals and program managers (reading coaches and teacher leaders), who implemented reforms used cost-effectiveness variables in decision-making was approached in two qualitative data formats. The first was an interview of program managers and principals; interviews were 27 questions (Appendix A) that ranged from 90 minutes to one hour and 30 minutes in length. The second tool was a 45-question survey of primary grade teaching staff (Appendix B) who were directly involved in teaching the primary reading programs.

The interviews of the program managers and principals included 13 questions regarding the use of cost-effectiveness strategies. The questions attempted to elicit information about: (a) the professional's knowledge of specific cost-analysis concepts, and (b) the professional's application of cost-analysis strategies in the evaluation of primary grade reading programs. The results of the program managers and the building principal's interviews are displayed in Table 20.0. Using a clustering of eight questions the data in

Table 20.0 data outlined varying degrees of expressed knowledge about cost-analysis concepts.

Program Managers

The individuals closest to the students and teachers (the teacher leader in the locally designed models and the reading specialist, or reading coaches, in the SFA and *Reading First* models) spoke about the need to use student outcome data in the evaluation of how students were performing in the current reading format. However, when asked specifically about the program data not related to student outcomes, they expressed only some knowledge about the variables related to the resources their program was using. Each program manager knew very little about how the program was built or had diminished over time, and what the program took away from other curricular areas and instructional practices. The program managers also expressed little if any understanding of the role of cost-effectiveness data in the evaluation of their reading programs.

Table 20.0 also identified a second set of variables of program evaluation; the application of the cost-effectiveness variables in the decision-making process. Five questions clustered around the idea of using cost-effectiveness tools as part of the decision-making process, including the key strategy of looking at a variety of programs prior to initiating or continuing use of the current model.

In the application of cost-effectiveness variables (accounting for resource expenditure, determining positive or negative effects on other programs, evaluating change in employee work habits), program managers had not used cost-effectiveness variables in the decision-making process for their school. Only one program

Table 20.0

Program Manager and Principal Expressed Knowledge of Cost Effectiveness Concepts

Cost Concepts	Program Managers (N = 4)			Principals (N = 4)		
	Expressed Knowledge			Expressed Knowledge		
	Specific	Some	None	Specific	Some	None
Data drives program decisions	4	0	0	4	0	0
Resource cost variables	1	2	2	3	1	0
Distributed costs	0	1	3	1	3	0
Opportunity costs	0	1	3	2	2	0
Use cost effects in decision-making	0	1	3	1	3	0
School has evaluated alternatives using reading criteria	1	3	0	1	2	1

manager had some knowledge of what the possibility could be for their school and three stated they did not know how looking into costs could be used to support their work.

On the second variable, understanding the need to evaluate a variety of models before making program choice decisions, three of the four program managers stated they did some research, and one leader at Jefferson spoke in depth about the staff traveling to other schools to review at least four to five programs in an attempt to understand what would work best for them.

Principals

Continuing with the interview process, the same set of questions was given to the site principals to elicit their understanding of cost concepts. Like the program managers, they understood the necessity of using data to drive decisions about their reading programs. Three of the four knew the resource cost variables without being prompted by the interviewer and could give specific examples of the items in their programs. When questioned about the variables of distributed costs and opportunity costs, only one could give specific descriptions of distributed costs and two could give answers showing only some knowledge of opportunity costs.

The principals of the four schools presented mixed understandings of how to use cost-effectiveness variables in the evaluation of their reading programs. Three of the four principals had some knowledge and talked about specific use of resource cost variables in their decision-making process. However, when questioned about the distributed costs, only one principal was able to reflect upon changes in work load at the teaching and supervision levels, reallocation of school funds from other programs, and less staff time focused in other content areas. The questions that focused on opportunity costs elicited specific responses from one principal, while two had only some knowledge of opportunity lost in both fiscal terms and teaching time in other content areas.

Although principals' understanding did not match the program managers' understanding of cost-effectiveness variables, the way they presented their assessment of the situation differentiated their responses. All four principals shared various anecdotal responses about when they worked with their leadership teams in discussing program

costs. In responding to questions about the use of cost-effectiveness concepts in decision-making, only one principal said his staff looked at cost as an ongoing concern. In this instance he engaged the entire staff to clarify the budget priorities for their site as part of school planning process.

On the second variable, reviewing a variety of programs before making the program decisions, one of four principals spoke of previewing a variety of programs before making a program decision. Two stated that their staff had made some peer inquiries but had not visited other schools or completed a journal review on the development of primary reading programs. One principal stated that “the school did not have the resources to purchase an alternative to the district’s adoption so looking around was a waste of the staff’s time.”

In conclusion, the program managers had limited awareness of cost analysis concepts and did not know how cost analysis could support their work. Principals had a better understanding but many had not implemented their conceptual understanding. Therefore, staff had little practical use for cost-effectiveness tasks in evaluating their programs over time.

Teachers and the Decision-Making Process

The teacher survey return rate at three of the four schools ranged from 70% to 100%; in the fourth school less than 30% of the teachers returned the survey. The 45-question teacher survey contained nine questions that addressed the concepts of: (a) use of data for decision-making, (b) conversations within the staff about program costs and materials, and (c) teacher contact with the program materials before purchase. The results.

Table 21.0

Teacher Awareness of and Interaction with Cost-Effectiveness Measures

Schools (N = Number of teachers responding to survey)												
	Adams (N = 2) ¹			Franklin (N = 5)			Hamilton (N = 4)			Jefferson (N = 9)		
Cost Concepts	# of Assessments Used	% Yes	% No	# of Assessments Used	% Yes	% No	# of Assessments Used	% Yes	% No	# of Assessments Used	% Yes	% No
Use data in decision-making	3	100%	0	3	71%	29%	3	75%	25%	4	90%	10%
Knowledge of cost-effectiveness concepts	NA	100%	0	NA	36%	64%	NA	25%	75%	NA	30%	70%
Involved in initial curriculum review	NA	0	100%	NA	33%	67%	NA	25%	75%	NA	33%	67%

¹Adams school had a 30% return rate to the study's survey. All other school s return rates ranged from 70% to 100%.

of the survey are presented in Table 21.0. The survey responses of the teachers reveal in general terms that the conversations about data driven instruction was universal in the context of reading instruction. However, the discussion of reading resources use was less universal

Teachers' discussions about reading program costs and the understanding of program cost variables were not much different than their program managers'. The percentages of respondents are inverted to the data discussion cluster. Only 25% of the respondents recalled some conversations concerning program costs. In the two schools that have extensive budgets and modified programs, very few respondents answered "yes" to questions regarding cost considerations in decision-making. In regards to Adams school this data is viewed with hesitation, given the limited return rate of the teacher survey.

Table 21.0 also highlights that schools that focused less on reading discussed their program cost variables less often. The two locally designed schools used an equal number of assessments, but recalled fewer conversations about data-driven decision-making. However, they were equal to their structured program peers in their limited knowledge of cost-effectiveness concepts.

The third variable regarding the teachers' ability review materials prior to adoption was irregular among the schools involved. From one quarter to one third of each staff had some participation in the review and selection of the materials their school was currently using, with the exception of the two long-time teachers at Adams. The data

Table 22.0

Frequency of Cost Effectiveness Variables Discussions at School Sites as Recalled by Program Managers

Frequency of Cost Effectiveness Variables Discussions at School Sites			
Program Managers (N = 4)			
Schools	At Leadership Team and Staff Meetings	Time Teachers Discussed Reading Program Criteria	Teacher Directly Involved in Choosing Curriculum
Adams	At leadership occasionally as a whole staff not really.	We've spent none, I do not really know cost stuff.	Ours came from grassroots brought to me by our previous Title I teacher nine years ago. It was
Franklin	Leadership and CSIP planning quite a bit. I do not know C/E well. So I do not know.	We talked about program design, not criteria, like group size and support. Only used peer conversations.	Curriculum and the format was teacher driven. We used the adoption at first but not very often now.
Hamilton	We have talked about costs at times, but not much. I think we'd use C/E information.	We have not really changed what we do, just materials. We have only looked at supplemental materials.	At our Team Level we discuss small changes, but not full reviews. We did choose supplemental materials.
Jefferson	Reading Coaches and Principals and Site Council. Cost effect would be difficult	We stick with the RF criteria; it is a good set of research-based ideas to focus our conversations.	Quite a few, I'd say at least five were reviewed. And we went on site visits, too.

Table 23.0

Frequency of Cost Effectiveness Variables Discussions at School Sites as Recalled by Principals

Frequency of Cost Effectiveness Variables Discussions at School Sites			
Principals (N = 4)			
Schools	At Leadership Team and Staff Meetings	Time Teachers Discussed Reading Program Criteria	Teacher Directly Involved in Choosing Curriculum
Adams	All our staff sees our literacy program as a clear priority; so we do not discuss costs much at leadership or staff meetings. C/E information would be a little scary but informative.	We've been doing this for nine years and that is before me; so I really do not know what criteria were used.	I believe that the previous Title I teacher brought it to the staff... But I was not part of that. Now they choose curriculum pieces from SFA based on their groups' assessment feedback.
Franklin	Leadership and CSIP planning segments at the first of the year, we talk frequently about costs and what we are giving up to do our flooding model. Having C/E information would help me greatly in those conversations.	We did not really go much outside our four walls. We did some reading about programs but no visits. We have talked about program design, but not real criteria, like group size and support. We used peer conversations through local principals and at teacher leader meetings.	Our current reading curriculum was adopted and purchased by the district. We moved toward leveled books for daily instruction. Our design and implementation of the flooding model was completely driven by our K-2 teachers.

Table 23.0

Frequency of Cost Effectiveness Variables Discussions at School Sites as Recalled by Principals (cont'd.)

Schools	At Leadership Team and Staff Meetings	Time Teachers Discussed Reading Program Criteria	Teacher Directly Involved in Choosing Curriculum
Hamilton	I guess I sold the leadership team the idea and they agreed to it. We laid it out for the staff, but I'd say the discussion was limited. We usually do not discuss dollars, at least while I've been here. I think our funds are so tight; having no bilingual of Title I dollars, talking cost is pretty futile.	Well as I said, we do not have the funds to look around too much. We looked at getting our reading group ratios down and that was all about staffing not really about materials. So outside of that I'd say we did not look too far.	We've added some materials from the W. G. but we basically took what the district adopted and went with it. We did not have much of a choice.
Jefferson	I sit with the district grant director and the Reading Coaches and we figure a proposal to take to staff. But it really is not discussed much. Our Site Council covers the cost of staff development so they talk a bit. C/E would be difficult, it does not cover the entire picture, so I don't really need it.	We stick with the RF criteria, it is our required guide. Plus our reading coaches and I have a good amount of teaching experience between us so that works itself in too.	Quite a few, I'd say we did about three site visits and viewed in depth as many as five publishers. We are still looking at things as the years of the grant move forward.

around teacher selection of materials, and the perceived advantages of teachers teaching what they have faith in, is unclear and no conclusion should be drawn from it.

Leadership and the Decision-Making Process

The interviews of the four school principals and the four program managers or teacher leaders used six questions to understand their functional skills with cost-analysis strategies as part of their reading program evaluation strategies. The interviews used a double back strategy by rewording similar concepts to get to deeper held notions regarding the topics, and to judge the consistency of the interviewees' answers. The data in Tables 22.0 and 23.0 use the words of the program managers and principals to represent their leadership practices. Their words represented how frequently and to what depth they utilized cost-analysis strategies. The nature of the program evaluation strategies used by principals logically mirrors that of the program managers with whom they work, and also resembled the comments of their teachers. But there are differences between the program managers and the principals they worked alongside.

Of the six questions, a single question asked directly about the current or potential use of cost-effectiveness principles as an evaluation tools. The intent was to determine how program managers and principals looked at program outcomes within their leadership teams. Of the program managers three of the four needed explanation of cost-analysis concepts or did not understand how cost-effectiveness analysis could help inform their program evaluation. Most program managers recalled that the leadership teams at their school had some conversations, but only one program manager stated that cost conversations made it down to the staff level. None of the program managers talked

of reviewing costs at the end of the year as part of evaluating the effectiveness of their reading program. The most common place for cost discussions was at the leadership team level and three of the four program managers characterized these conversations as limited.

Principals' recollections varied at each site. Most often they recalled discussions in small groups at the leadership team level. Of the four, only one took costs discussion to the staff level. Only one in four used cost variables as part of the school's program evaluation practice. When questioned further, all four stated they had cost notions in their budget work, but they did not compare their programs results or costs with other sites. Principals had a more complete use of the vocabulary of cost analysis; however, most admitted little desire to undertake the practice as part of their annual reviews.

The principals' and the program managers' perceptions of discussions regarding the criteria used to choose the current reading model were consistent. Both groups' comments matched the survey results of the teachers. The reoccurring issues of (a) coming into the school with a program already in place, (b) being given a program by the district, or (c) following a well defined protocol provided by an outside agency were consistent in the interviews.

Principals and program managers also had varying experiences with curriculum choice and implementation. The principals in the locally designed models were given the curriculum and some staff development in 2002 when the new adoption was implemented across the district. The principal at Adams inherited the *Success For All* program when coming in as principal six years ago. The principal at Jefferson started with the district

adoption in 2002 and had been supplementing with a variety of curricula meeting *Reading First* criteria over the past four years.

Only one principal expressed an interest in seeing a cost analysis evaluation of the school's program, although the thought "scared me a little." One principal felt that the concepts of cost effectiveness had been used in program evaluations, but that principal did not use all the components. One principal was pretty impassioned that being cost conscious was important but costs were clearly out-weighed by the need for teachers to have time to work collaboratively, and develop relationships with their students. This principal ended by saying that "the teacher relationship with the student is critical and that cannot be measured in dollars and cents."

Research Question 3 Summary

Teachers generally had a universal understanding of data driving instructional practice, and peripheral knowledge about elements of cost effectiveness. They spent more time as a staff reviewing data for teaching and implementation decisions than for costs purposes. All four program managers were functioning at this same level of understanding for cost-effectiveness analysis. Each program manager used data to guide the instruction of teachers with whom they worked. Although some had a better understanding of program costs because of leadership team discussions, none of the four program managers used that understanding to evaluate their programs.

All of the principals had a means of looking at their programs, both in the context of student achievement and of costs. However, only one of the four used costs variables with the school's classroom level staff to support decision-making. Although each

principal had a certain way of measuring costs, none of the four had the means of comparing their programs with alternatives. This lack of comparison gave rise to continual application of the current approach with minor additions of supplemental materials.

Interviews of principals and programs managers along with local documents reviews were undertaken to gain information about local program elements and the use of cost variables in decision-making. Teachers who have come out of the classrooms to become program managers or who are still in the classroom and take on roles of teacher leaders profess little knowledge about cost-estimation and cost-analysis concepts. They understand specifically that student outcome data were to drive instruction, but did not apply cost-analysis concepts in the review of their programs. The data show similar tendencies for the principals at the four schools. Principals had a greater awareness of the cost-analysis principles to be taken into account, but lacked the application of cost-analysis principles in the site's decision-making processes.

Research Question 4

The fourth research question attempted to inform the interpretation of reading program outcomes by presenting the perceptions of program managers and principals regarding the complex nature of students' lives. These perceptions provided a context for the programs' challenges in nonacademic terms. The professionals discussed their impressions of the student's characteristics that challenge educators in the daily work of the classroom. Using four interview questions dedicated to instructional strategies, school wide programs for behavior and attendance, and general impressions of students'

families, each of the professionals presented barriers outside of their control when looking at their students.

Program managers. Table 24.0 presents the descriptive representation of student complexity for each of the communities. The professionals at all four schools consistently referred to three variables: (a) the perceived increase in the number of students who begin school with underdeveloped primary language skills, (b) an ongoing challenge to maintain home support for work that would make up the gap that teachers feel exist in on-set reading skills, and (c) the increased number of students with severe behavioral problems that make teaching more challenging.

The program managers and principals at Franklin and Hamilton commented about four issues they worked around regularly: (a) inconsistent attendance, (b) increased incidence of students with severe behavior, (c) under-developed basic language skills, and (d) lack of parental support with home reading tasks. The comments were by and large about increasingly severe behavior on the part of some students in the first and second grade. The perceived lack of language preparation combined with little or no support of teacher developed home reading tasks were seen as slowing the growth in reading skills development typically anticipated in past populations. The increasing amount of challenging behaviors was seen as breaking down reading group instructional dynamics and impinging on all students' ability to learn.

Adams' program manager felt that the generational poverty common among the Caucasian population in that attendance area was the source of many of Adams students'

Table 24.0

School Community Descriptive Variables Associated with Student Complexity Variables

Complexity Variable	Schools			
	Adams	Franklin	Hamilton	Jefferson
School K-5 Size	317	527	342	504
School Attendance Rate				
Grades K-5	91%	93%	94%	93%
Subjects	91%	95%	91%	92%
SES (F/R Lunch)				
Schoolwide	89%	32%	42%	76%
Subjects: Grade 1	79%	34%	34%	81%
Subjects: Grade 2	75%	34%	65%	69%
Mobility Rate	21%	10%	16%	19%
% Reporting non-White	53%	73%	73%	39%
Teacher Experience				
Grades 1 & 2	N/A ¹	13.4	10.8	3.1
Average class size (K-5)	19.2	24.33	21.2	20.8
% of LEP in School	43%	.01%	.005%	47%
% TAG in School	4%	5%	6%	3%

¹Two of six teachers returned survey, one reporting 28 years experience.

gaps in learning, and their low performance on standardized tests. For this program manager, poor attendance and student behavior rose above language concerns as areas of challenge in the classrooms.

At Jefferson the program manager felt that the main issue was language acquisition. The second most commonly referenced challenge was student mobility. Although in this study the data presented saw minimal attrition of subjects, the district reports an overall mobility rate for Jefferson of 19%. The program manager felt that this turnover of one in five students had long-term effects on the reading outcomes for their program.

As the complexity questions narrowed the conversation to current students posing the greatest challenge to reaching reading benchmark, the program managers became more specific. When they addressed their comments at the individual level (i.e., “What about your most struggling student?”), all four schools’ professionals discussed the developmental language deficits of their students. This concern, combined with increased severity of student behavior, became the two issues of focus for their comments. The range of language development and the disruption of the once well run classrooms were the overwhelming concerns of the program managers working with grade one and grade two teachers.

Principals. For the principals at Jefferson and Adams schools, complexity revolved around four main issues: (a) the lack of language development in some of the most struggling students, (b) family mobility-stability rate, (c) a limited experience base to create background knowledge associated with reading materials, and (d) limited cultural awareness of the school (associated with behavior and attendance concerns). When the discussion moved to each of these areas and principals were asked to rank their concerns, language underdevelopment was the challenge most often cited. Most said

attendance and behavior issues took up their time, but language development was a much harder gap to close.

At Franklin and Hamilton the issues focused around attendance and behavior. Both principals brought up the fact that severe issues of student behavior were challenging their teachers' ability to teach the whole class successfully. When asked to think of individuals who were most resistant to interventions, the Hamilton principal spoke of background knowledge differences in that school's students. The Franklin principal spoke of limited oral language development of the school's most struggling readers. Both felt that their teachers, who were conditioned to students coming in well prepared, were having to make adjustments to their instructional processes to meet the needs of these groups of children.

Research Question 4 Summary

In summary the complexity of the students in the study schools was a matter of degree. Franklin and Hamilton had from one third to one half the free and reduced lunch participants of Adams and Jefferson. Jefferson had an English as a second language population that is four to five times higher than at Hamilton and Franklin. The structured programs at Adams and Jefferson served mobile populations at nearly twice the rate of Franklin. However, school program managers and principals spoke with the same passion about the barriers to their students' ability to learn at a level assuring attainment of district benchmarks. Therefore the principals and program managers of the structured programs saw the same main issues of attendance and behavior but dealt more with

concerns around (a) the modeling of language in the home environment and the associated home practice of reading skills, and (b) student mobility.

The program managers and the principals at all four schools spoke about the gap in the language acquisition in both English speaking families and English as a second language learners. These complexity issues are a matter of degree as shown in Table 24.0, but are a reality for all the professionals at each of the four schools.

Research Question 5

The final research question sought to determine whether program fidelity factors affected the cost analysis of the four schools' reading programs. The question of program fidelity was considered because the degree of practitioner adaptation within the school may alter the intended effects of the instructional design by the developer—in this case, the reading program publishers.

Program fidelity variables were measured using all three professional groups. The teachers responded to six survey questions, while program managers' and principals' interviews contained five questions. Table 25.0 presents the results of the teacher survey. Teachers in the two structured programs had a high degree of faith in the curricula and strategies. However, the Jefferson school personnel also utilized supplemental materials at least one time per week in 70% of the cases. In both Adams and Jefferson schools the daily materials consisted of 80% to 100% of the *Success For All* and *Reading First* protocol adoptions. The teachers working in the two locally designed programs showed less faith in their basal program's ability to meet students' needs, greater levels of daily modification of publisher designed lessons, and consistently used supplemental materials

(> 20% of the time). The schools with oversight through the use of reading coaches or outside foundations reported higher favorable responses to questions of program faith, less use of outside materials, and fewer modifications to the program's lesson structure.

The teachers' self-reported tendencies when viewed in relation to the five questions asked of program managers and principals gave a more complete picture of the strength in program implementation fidelity. Table 26.0 highlights program managers' and principals' responses to questions related to program fidelity. Principals and program managers were not asked to quantify coaching situations or classroom walk-through situations; however, they were asked how they work with teachers. Their responses show they work in staff development implementation, observation, and leadership decision making. Five questions were used to determine: (a) the degree of individual modification permitted in instruction and planning, (b) personal perceptions of teacher engagement in the reform, (c) the linkage of planned staff development with instructional practice, and (d) the methods used by school leadership to assure that the program is being followed in the manner expected by the developer.

Program manager and principal responses indicated that because Adams and Jefferson had outside consultants and funding structures, those schools can hold a firmer line on discouraging teacher individualization. This was true in planning, material use, and linking staff development with instructional practice. These structured programs relied on their on-site coaches to assure structural fidelity, material fidelity, and support of teachers' use of resources. In the structured programs three of four program managers

and principals rated fidelity as strongly relying on outside monitors, reading coaches and principals to oversee daily implementation.

Table 25.0

Teacher Familiarity and Working Experience with Current Reading Program Materials

	Schools (N = teachers responding to survey)			
	Adams (N = 2) ¹	Franklin (N = 5)	Hamilton (N = 4)	Jefferson (N = 9)
Fidelity Concepts				
Time of program in school	9	4	4	4
Teacher average years of experience with program	6	2	4	2
Teacher faith in program	100%	80%	75%	90%
Daily lessons w/ core materials	80-90%	80-90%	70-80%	90-100%
Lessons the teacher modifies / week	1 x /wk	1-3 x's /wk	1-2 x's /wk	< 1 x /wk
Teachers using supplemental reading materials	100%	80%	75%	70%
Instructional assistant plan their own work	Yes	No	No	No

¹ Adams school had a 30% survey return rate; all other school's return rates ranged from 70% to 100%.

The locally designed models had both less formal approaches and fewer individuals involved in fidelity oversight. Principals and program managers of Franklin and Hamilton schools without Title I funds or outside reviewers differed in their

responses to the set of fidelity questions. The program managers (teacher leaders) and principals tended to provide fewer local controls over the teaching and learning

Table 26.0

Program Manager and Principal Perceptions of School Program Fidelity Variables

Schools

Professional	Reading Program Fidelity Variables				
	Teacher choice in curricula	Program fidelity rating	Training linkage to instruction	Fidelity assurance practices	Use of instructional assistants
<i>Adams</i>					
Program Manager	No	Good	Strong	Outside monitor, coaches monitor	Teach small groups
Principal	No	Strong	Strong	Outside monitor, coaches monitor, observations	Teach small groups
<i>Franklin</i>					
Program Manager	Yes	Good	Good	Team meetings, CSIP data	Teach small groups
Principal	Yes	Good	Fair	Team Meetings, CSIP Data, observation	Teach small groups
<i>Hamilton</i>					
Program Manager	Yes	Good	Good	K-2 Team meetings, emails	Teach small groups
Principal	Yes	Good	Good	Team leader, observations	Teach small groups
<i>Jefferson</i>					
Program Manager	No	Strong	Strong	Outside monitor coaches monitor	Support teachers, Teach small groups
Principal	No	Strong	Strong	Outside monitor coaches monitor	Support teachers, Teach small groups

environment. Without consultants and secondary measures teachers were given more choice in materials and strategies. Both schools stated they closely followed the district's reading *target skills* document; thus, program managers felt teachers were working on similar skills at the same time. However, teachers reported that during 10% to 30% of teachers' reading instructional time they were using supplemental materials or lesson designs. In Franklin and Hamilton three of four of the individuals gave *good* ratings to questions linking staff development to the reading program. However, their frequency in staff development is significantly less than in the two schools using structured programs.

The final variable of fidelity looked at how the programs used their instructional assistants. Use of instructional assistants was reviewed to see who was presenting the lessons as part of the instructional programs. Three of the four programs used instructional assistants side-by-side with teachers who model strategies and plan small group lessons. Adams school was the only school that allowed instructional assistants to carry out their own planning and lesson delivery under the direction of the program manager. The two programs with the larger funding levels (Jefferson and Adams) also had more instructional assistant hours per day and greater capacity to train and support their assistants through staff development opportunities.

Research Question 5 Summary

The results of teacher survey and principal and program manager interviews provide a practitioner's perspective of the four reading programs. The structured programs had greater oversight in terms of testing requirements, internal and external support of staff development, and greater staffing levels that led to stronger tendencies

toward complete program implementation. In the five questions asked of program managers and principals regarding teacher autonomy, teacher engagement in the reform, the linkages between training and instructional practice, the structured programs of *Success For All* and the *Reading First* protocol generated responses more consistent with programs exercising fidelity controls. The two locally designed programs had higher degrees of teacher individualization, fewer staff development opportunities, and fewer individuals giving support to teachers or monitoring teacher performance.

CHAPTER V

DISCUSSION

This study measured five primary questions presented in Chapter I: (1) Was there a difference in the reading outcomes of the four reading programs under review, (2) of the four reading programs which one had a better costs-efficiency ratio, (3) what did school's teachers, program managers, and principals use as their decision-making process when programs were implemented, (4) how did the student population's complexity contribute to the reading outcomes students achieved, and (5) did site managers and principals feel the reading programs in place were being used as intended by the program developers?

This chapter covers these five research questions in succession with specific focus on the main finding concerning the leaders in the four schools skills in using cost analysis techniques. The primary finding of this program evaluation is that cost-analysis currently plays a limited role at the site level when professionals are making program evaluation decisions. Principals and teacher leaders / program managers do not understand cost analysis principles well enough to integrate them into the technologies of their work. This is compounded by the finding that three of the four schools leadership teams and site principals tended to disregard cost conversations in favor of reading outcome information at the staff level. While using reading outcomes for data driven instructional

decision-making has begun in this sample of four schools, program cost conversations are still the purview of the leadership team. These findings will be expanded upon in the discussion of Research Question 3.

Secondary findings about the role of cost-analysis in informing the public about program alternatives is of equal interest, and similar to those presented by Levin and McEwan (2002) and Hummel-Rossi and Ashdown (2002). Research Question 3 explores gaps in practitioner skills and points to a recommendation that cost-effectiveness ratios ought to be one of five areas of consideration in the review of possible alternatives in school reform practices. Currently the schools in this study are: a) using common academic outcomes of primary reading skills, b) understanding school decision-making practices, c) disaggregating their school data on measures of community complexity. Two of the four schools also attend to fidelity variables through well developed program oversight practices. These four practices along with the use of cost-effectiveness variables could provide clear comparative information about the alternatives under review. Hummel-Rossi and Ashdown (2002) point out that publication of reviews of this type could inform the educational community about programs that support informed decision-making.

Cost-effectiveness ratios have the potential to inform the funding models like Quality Education Model and respond to community pressures for reform. Use of cost analysis in school level planning and program evaluation has the potential to inform decision-making and support reform efforts that generate academic outcome information for program evaluations. This study found that schools did not accurately track cost data

that could be used to define distributed costs, opportunities costs, or discount the cost of programs over the length of their existence. Districts did not require this type of data collection at schools and principals were not eager to take on the practice on their own accord. Concurrently, building principals although having more complete knowledge of cost analysis concepts, did not utilize that understanding in program reviews, and where not required by their district to carry out such reviews. Principals understood their cost, but did not analyze those cost in relation to student outcomes in primary reading programs. This finding will be discussed in more detail with relation to reading outcomes and cost-analysis as part of Research Question 2.

In Research Questions 4 and 5, the teacher survey tool and data from program manager and principal interviews were used to understand the variables related to student complexity and program fidelity at each of the four schools. Teacher responses characterize their impressions of each school's community complexity and how these traits affect their academic programs. Teachers, program managers and principals all shared their impressions regarding the degree to which the teachers adhered to the publishers and program design characteristics. Fidelity measures are important to gage if the outcomes that the students achieved are related to specific design parameters of the publisher or variables unrelated to the reading program.

Research Question 1

One of the consistent shortcomings in cost-analysis research has been the under utilization of comparative evidence due to non-equivalent outcome measures (Levin, 2002). This study looked at the combined group differences in the four reading programs

outcomes to discern which program supported students achieving district benchmark most effectively. Using *Rigby's Benchmark PM 1 and 2: Primary Literacy Assessment* (Nelly & Smith, 2002; 1999) the study generated a common set of reading measures as comparative variables to review alternative reading programs. To answer the question about which of the programs produced the greatest growth in students on the six reading variables a pre-test and post-test design was used over 163 days of instruction. A random sample of first and second grade students generated the reading outcome data for this study. The reading outcome data must be viewed with certain limitations because (a) differences existed in the number of students on free and reduced lunch when Franklin was compared to the other three schools and (b) Jefferson had significantly higher proportion of students whose home language was something other than English. Within these known limitations, the overall combined group differences on the pre-test reading scores for the six variables measured showed non-significant differences among the eight school groups.

First Grade

Of the four programs the first grade students in Franklin's flooding model presented the most growth for the time period. Based on the first grade scores in post-test book level achieved there existed a significant difference in the four schools. Franklin's average *book level read* score of 17.0 was above the districts standard of 16.0 and significantly different from Adams and Hamilton's post-test scores. There were two areas that may confound the differences seen in the Adams to Franklin comparison. First was the difference in socio-economic status as measured by free and reduced lunch

enrollments, secondly was the difference in early on-sets skills as measured by letter identification and sound identification in the pre-test outcomes.

Franklin's free and reduced lunch population was just less than half that of Adam's. Barton's (2003) 14 characteristics of complexity implied that the challenges outside of school that are faced by these students in their pre-school and kindergarten years may leave their skills underdeveloped compared to their peers. Students enrolled in free and reduced lunch programs are more likely to come from homes with little if any formal childcare practices, few educational materials (books, computer drill games) and as such early reading practices go taught, fewer educational experiences outside the home, and increased health complications (Barton, 2003).

However, the real significance may have come in Franklin's significant difference in letter identification at the pre-test period. This skill allowed students to access the phoneme and grapheme relationships and thus supported their putting together words in an attempt to make meaning from text (Rayner, et al., 2001; Foorman, et al., 1998). The letter identification advantage that Franklin students have allowed them to use early reading strategies of decoding, and context clues, in reading short picture books earlier (Hiebert & Taylor, 2000). Earlier access to decoding skills and strategies gave these first graders an advantage over time to gain higher book levels by the end of the trial period.

However, this did not explain the growth of Hamilton's students on these same variables. A second explanation was contained in the fact that small groups of students were used in both Hamilton's and Franklin's programs, but the characteristics of the small groups differ at the two locally designed schools. Using the information gained in

the teacher survey it was possible to use the four categories of essential elements of effective reading programs to ascertain why some programs achieved better outcomes than those with whom they are compared. The categories of explicit instruction, intensive instruction, comprehensive instructional approaches, and socially supportive environment were present in varying degrees in each program.

Of the four categories that distinguished Franklin from Adams and Hamilton, the intensity of the instructional approach provided the most complete explanation. Reading group size was most prominent in the comparison of these three programs. Group size affects intensity in three important ways: a) the span of instructional materials and the required planning, and b) the opportunity to participate, and c) number of opportunities per week. Each variable can be used to explain the difference in book level mean scores for first graders.

Span of instructional materials and planning. The span of instruction that a teacher had to attend to in both planning and presentation of reading material affects student outcomes (Foorman & Torgesen, 2001; Foorman, et al., 1998; Spiegel, 1995). Teacher planning and presentation for narrower skill groups could give the teacher an increased ability to scaffold instruction to a set of skills appropriate to all five to six students. Groups of five to six at Franklin had a narrower span of instruction than groups of eight to ten as seen at Adams, even when students are placed based on eight week assessments as is the case for *Success For All* programs. This span of instruction along with opportunity to participate, the second variable, may influenced the time-on-task for

reading that is necessary for groups to run smoothly and learning to take place (Shanahan, 1998; Pikulski, 1995).

Opportunity to participate. Opportunity to interact with the curriculum and the teacher was an advantage at Franklin when compared to both Hamilton and Adams. Franklin's most struggling readers work with a specialist teacher, *and* had a small group structure of three to four students. The approach when compared with instruction provided to the most struggling students at Hamilton with group sizes of five to six and Adam's with group size of eight to ten may have influenced student growth. Therefore, as presented by Shanahan (1998), the time on task and the student's ability to participate in all parts of the tasks explain more of the difference in the two schools reading outcomes than the nature of the student populations. There is little doubt that students from backgrounds in poverty, as seen at Adams, had challenges to accessing high quality pre-school, and fewer supports at home in their pre-school years. However, instructional programs based on acceleration require group structures to give adequate support to individual opportunities to practice. Maintaining large student groupings in early reading skill development may limit the acceleration model, and in fact contribute to the ongoing achievement gap.

Number of days per week. Franklin's groups met for 30 minutes four times per week and had five to six students in the three higher level groups, but only three to four in the lowest group working with the reading specialist. This contrasts with Adams who had groups five days per week for 30 to 45 minutes, yet worked in group sizes of eight to ten students. These contrasting group sizes may have affected the intensity of instruction

to the degree that the most struggling readers, regardless of whether they are receiving tutoring, may not have been able to access small group instruction as efficiently as those students who worked in groups that were two-thirds smaller.

Shanahan (1998) attempted to define whether it was group size or number of days and length of meetings that contributed to reading gains, however, he drew no clear conclusion. In this study it appeared schools that used shorter periods of time with smaller groups make greater gains in the book level read, reading fluency and reading comprehension scores of first grade students when compared with reading programs that used larger groups and longer time periods of student grouping. The intensity of instruction may be such that time-on-task and focus of the teacher's instruction can meet the criteria of intensity better than programs using larger groups supported by tutoring.

Second Grade

The findings of the second grade students were distinctly different from those of first grade in the areas of book level read, reading fluency, and reading comprehension. The areas of letter identification, letter sound identification and word recognition were seen by reading teachers as first grade on-set skills and as might expected small samples of 25 to 35 subjects surfaced only a few subjects in these assessment areas at second grade. The pre-test to post-test scores of these three variables provided small samples that could not be interpreted with any certainty. However, the more advanced skills of reading fluency and reading comprehension based on the level of book read gave insight into student abilities to use on-set skills of decoding and automaticity as well as advanced reading skills such as metacognitive strategies (context clues, picture clues), as well as

demonstrate expressive reading, understanding of syntax, and using literal and evaluative comprehension skills.

Franklin's pre-test to post-test book level growth was significantly better when compared to the other three models at second grade. This was the only significant finding in the outcome measures for second grade. Using the four categories of essential elements of reading programs and interventions of explicit instruction, increased intensity of instruction and socially and emotionally supportive approaches the differences in programs can be explained.

Using a small group format for two years may support students in three critical elements of effective reading programs a) explicit instructional methods, b) intensity of instruction, and c) socially and emotionally supportive practices (Foorman et al., 1998; and Hiebert & Taylor, 2000). The three variables of intensity (opportunity to participate, number of contact days, and span of materials) may have a cumulative effect over two years. Small groups for the lowest of students may allow teachers to scaffold instruction more adequately for improved acquisition of on-set skills that build into effective reading skills, including metacognitive strategies by the end of second grade. The drop from twelve students reading well below grade level in the fall to five students reading well below grade level in the spring can be explained using the categories of essential elements of reading programs and interventions. Franklin's program has been shown to differ in intensity variables (narrow span of instruction and opportunity to participate) at first grade. As these variables are added to the second year of reading instruction a greater number of students are brought up to grade level. In year two the variables of

socially and emotionally supportive approaches and the use of explicit instructional strategies may finally take hold on the most resistant readers. This in combination, increased intensity and explicit instruction, carried out by a reading specialist working with the most struggling readers supported student most effectively out of the four programs (Pinnell, et al., 1994).

Foorman et al., (1998) and Spiegel (1995) found that explicit instruction using scaffolding strategies that incrementally refined skills and led to student teacher conversations about what step to take and where to apply the strategy supported student growth. Grouping size directly affects these conversations and the span of direct instruction, both of which explain the differences in book level scores for the second grade students involved in the reading flooding approach at Franklin. Teachers in the other three schools reported larger group sizes and alternative uses of the reading specialist as data collector, coach, and staff development leaders. These activities maintained larger groups for the readers, including those at risk for reading failure, and led to less effective instruction for the most resistant learners.

The descriptive statistics when viewed through the lens of essential elements of effective reading instruction and intervention give some insight into what was at work in the four reading programs. However, the outcome measures do not speak to how they were obtained. The complex nature of a school and its students may make growth a slower process for some students and some programs overall. It is clear that the Franklin students out-performed the students in the other three schools on most of the outcome measures, however explaining why that is the case is not as clear.

Research Question 1 Summary

The variables that stand out from the teacher surveys, program managers and principal interviews are related to the characteristics of the programs that can be clustered into the four categories of essential elements as found in the review of reading instruction literature. This study of four programs using similar instructional material with differing school populations appears to point to two of the essential elements as possible explanations of the program differences. First was the intensity of instruction as defined by group size and instructional opportunities. The second variable was the nature of the instruction as defined by both the emotionally supportive interactions of the reading specialist and the explicit nature of the instructional protocol.

Students at the lower end of the performance scale ought to have smaller groupings than students reading at grade level or above. The findings give rise to the use of groups of three to four in combination with a reading specialist as an effective means to accelerate student growth so they can catch-up to their peers in fewer months of work. Whether it is the narrower focus that smaller groups bring to instruction, or students improved opportunity to practice with direct feedback and personalized responses, the smaller groups at the lower end of the instructional scale supports improved outcomes.

Franklin's growth specifically at second grade may have its initial success in first grade with student having better letter identification skills and students attaining higher end of the year average book levels. However, over the two years Franklin moved more students out of the high risk areas of below book level nine and into reading at or near grade level by the end of second grade than the other three programs. This ability to reach

the most resistant learners was the function of the using the reading specialist to carry out instruction with the most challenging readers and not relying on the specialist to interpret program data, mentor teachers new to the profession or organize and oversee staff development.

The complexity characteristics of Franklin have a much smaller impact of the overall program than in Jefferson and Adams schools. Franklin has significantly fewer students of second language, a lower mobility rate, a lower free and reduced lunch rate and less ethnic diversity than all three of the other schools. These variables have been shown to impact the outcomes of student learning in a variety of contexts (Barton, 2003). The ability to parse out the effects of these variables on the student outcomes led to findings that the initial differences in the reading pre-test scores gave only a slight advantage to Franklin over Adams and Hamilton and no advantage over Jefferson at both first and second grades as measured by reading outcomes. For all groups the age of students, and the mean days of attendance were not significantly different. Therefore, each group of students received equivalent numbers of days of instruction and was developmentally equal based on age. The growth at Franklin in comparison to the other schools is significant; however it can only be said with relative certainty that the instructional program is the reason for the difference. Other factors such as the nature of the teacher, the complexity of the other school communities, may explain the difference more completely. Research Question 2 will attempt to associate program costs with these outcomes and later research questions will attempt to explain the four schools outcomes including a look at the complexity variables associated with public schools in the U.S.

Research Question 2

Principals and program managers were interviewed and school documents reviewed to determine program cost variables by category and their associated amounts. The processes used to define costs for the sample groups gave the reviewer an opportunity to look at cost-effectiveness ratios and other factors in the evaluation of the four programs. Research Question 1 showed that three of the four schools start in statistically the same place for five of the six reading outcomes variables measured. Adams and Hamilton are significantly lower than Franklin at grade one in letter identification and at grade two on the pre-test book level read variable. However, Jefferson is statistically equal to Franklin at both first and second grades on all six of the reading measures, yet Franklin's students finish with statistically significant higher book level read scores compared to the three schools associated with the study.

There are a variety variables that could contribute to the differences in outcomes ranging from different instructional strategies, differences in the allocation of resources by school, differing teacher abilities, and student complexity variables. Research Question 2 looked at the costs of each program as reported by program managers, principals, and school / district documents to discern a cost-effectiveness ratio for each program based on the number of sample students meeting the district's reading benchmark.

Program Ingredients Costs

In recent years federal education Title I programs have pursued a policy initiative that attempts to link student outcomes with funding evaluations (NCLB, 2002; Jefferson,

2005). These accountability policy initiatives were institutionalized in the No Child Left Behind Act of 2001. Accountability measures were pushed into non-title funded schools through requirements for whole school reporting of state sanctioned outcome measures. This approach had led to increased use of student data in a school wide context to review program effectiveness (Borman, 2005; Darling-Hammond, 2004; Hummel-Rossi & Ashdown, 2002).

The discussion about resource use with program managers and principals was a based on specific set of questions in the school level interviews. The concepts of opportunity costs, indirect, direct, fixed, and distributed costs effects on overall school programs were explored from three interrelated processes: (a) measuring student outcomes for each of the alternatives with a similar set of research defined variables, (b) delineating comparative variables that characterize the nature of alternatives under review, and (c) the analysis of program costs for the use in developing cost-effectiveness ratios (Levin, 2002; King, 1994; Grissmer, 2002). In the school personnel interview process questions were directed at the nature of costs (opportunity, direct / indirect, and fixed costs), how the data about those costs were tracked by staff, and how the costs information was used.

The definition of ingredients variables (resources) that contribute to understanding the overall program cost was the main focus of Research Question 2. Using the ingredients model defined by Levin (1986; 2002), King-Rice (1997), Stiefel, Schwartz and Rubenstein (1999), and Chambers (1999), program costs were categorized into four groupings based on school documentation. The categories of: a) personnel, b)

materials and equipment, c) client inputs and d) unique inputs were defined at each school site. The fifth variable of *facilities*, as defined by King-Rice (1997) Stiefel et al., (1999), was not considered as a factor in the year of study at the four schools due to a lack of data.

Ingredient cost variables. Locally defined cost considerations outlined in the ingredients and resource cost models made up the costs for analysis of the four programs. However, in the locally designed programs budget values in the school site plans were not compared with program expenditures. This planning process was not considered by either set of professionals as necessary since budgets concerns were not part of the program evaluation process in general.

As is typical in most public schools operations personnel accounted for between 83% and 92% of the local school programs costs and between 78% and 87% of the more structured programs personnel costs (Levin & McEwan, 2002). When compared with King-Rice (1997) and Chambers (1999) findings these cost outputs fit the research.

The personnel category of ingredients was reviewed with the greatest clarity by school principals. Principals of well funded programs had some troubles parsing out which personnel served particular groups of student and what funding stream paid for their duties. However, local documents were adequate to get accurate representations of personnel costs. Principals in locally designed programs could easily give totals of both the hourly level and costs including benefits. This process was clear to them and they could converse about personnel with some historical perspective as well.

This was not the case for materials and equipment variables. Materials and equipment was the most challenging for principals to document, this was especially true in the locally designed models. The locally designed models had been in existence over the past four years. Purchases of curriculum materials and been made by the district, with the exception of the supplemental materials. The principals did not know if they counted supplemental materials as unique inputs or as materials and equipment purchases. In the interview they had no recollections of the cost of material and equipment over time, so accounting for this variable presents a problem in the review of alternatives process that the ingredients model is intended to facilitate.

The challenge in reviewing alternate programs using the ingredients and resource cost formats comes from defining costs variables in two areas: a) unique inputs and b) client inputs. Therefore, the ability to neatly categorize client inputs and unique inputs has the potential to inform school site resources costs. In the programs evaluated here the two structured reforms had significant client inputs and unique inputs that distinguished the models. The use of outside oversight, staff development through time and consultants, and a reading coach format increased overall program costs, and were easy to identify. Compared to the locally designed models the unique inputs and client required inputs created a differential that was not matched. If these variables are not clearly identified the ability to compare alternatives within the evaluation process is very limited. Any marginal difference would not inform the decision-making process beyond the pedagogy or materials variables currently understood. Organizing inputs from a variety of resource streams and using common definitions would support the effective evaluation of costs

and allow for variables like discounted costs and opportunity costs to be viewed in light of the program's current context.

Distributed and opportunity cost variables were not available in this study and limited the use of cost-effectiveness analysis. The nature of these categories of costs would be highlighted if school personnel could accurately represent their local inputs. Models like Oregon's *Quality Education Model* would be better served if the professionals working within schools could inform the model of on-going cost factors that actually represent implementation trend characteristics in differing school environments, environments based on size, teacher capacity, training resources, school complexity variables and other relevant cost variables.

Cost-Effectiveness

Cost-effectiveness ratios. The literature described the most accurate cost-effectiveness ratios required the evaluation be measured at the student level to compare programs (Brent et al., 2004; Hummel-Rossi & Ashdown, 2002; Levin & McEwan 2002)). This approach gave the most accurate valuation of program costs directly attached to district assessments and the daily program practices (Chambers, 1999; Levin, 2002; Stiefel et al., 1999).

The findings in this four school comparison showed a difference between daily costs for the structured models (*Reading First* and *Success For All*) and the locally designed models to be nearly three to five times as large when considering the number of students meeting benchmark. The difference ranging from three to five times as large makes the cost variables significant criteria in the decision-making process, regardless of

the nature of costs incurred, how they were reported, and the context of the students the program is serving. The program at Franklin school that cost a little over one-third of the Jefferson model in grades one and two shows the greatest cost-efficiency. A comparison of all four schools showed both of the locally designed programs produced cost-effectiveness ratios much more favorable than the two structured programs.

The value of the cost-effectiveness ratios may be in the anticipated costs that schools of high complexity and low complexity attempt to project. From the comparison of the two structured programs it can be said that students in complex schools require resources estimates that include teacher training carried out by highly trained professionals to make reading gains sufficient to reach district benchmarks. The two structured programs in this study suggest that models anticipating a cost per student in the range of \$200.00 to \$300.00 dollars per student annually. Based on this study's findings, funding at this level complex schools can anticipate getting two-thirds of their students to reach district benchmark. To get 90% of the students to benchmark as stated by Lyons and Chhabra (2004) programs would have to spend significantly more annually. The amount is in question on the scalability of the models utilized. One could anticipate *SFA* costs reaching \$300.00 / student annually, and \$400.00 / student annually in the *Reading First* protocol. The cost-effectiveness ratios would suggest about \$100.00 dollars more per child, but economies of scale may reduce that value.

Locally designed models had much lower cost per child values. This is due in part because of the large difference in available funds to these programs. The students at Franklin and Hamilton had no Title I funding or bi-lingual funding, but did have

resources provided all schools in the district. The students come from more stable surroundings, are less inclined to qualify for free and reduced lunch, speak English as their primary home language, and move less often than their peers at Adams and Jefferson. However, current annual cost of \$87.00 to \$94.00 at Franklin and \$109.00 to \$170.00 at Hamilton. These expenditures get 80% of Franklin's students to grade level benchmark at second grade and 55% of Hamilton's student to benchmark at grade two. These two programs are clearly not reaching stated goals of 100% of students reading at benchmark at the end of the year. One could anticipate Franklin costs reaching \$110.00 / student annually and Hamilton spending \$245.00 / student annually to get closer to the policy goal of 100% of their students reading on grade level. The cost-effectiveness ratios would suggest shifting funds from other local programs to met this policy goal, however economies of scale may reduce the need by a unspecified amounts.

Research Question 2 Summary

Working with school documents and school personnel major costs are accounted for in each school's expenditures. Two of the primary findings of this study are: a) in this particular four school context program managers and principals did not track incidental costs expenditures considered minor cost categories which effect cost-effectiveness ratios, and b) costs over years where not tracked (nor were they required to track costs) as a data point for consideration of opportunity or distributed cost analysis. Whether this is consistent in program administration in other contexts (local government, non-profit organizations, and business) it may be a point of contention for constituents determined to use cost in a systems accountability model to critique school reform practices.

A challenge related to the ingredients model and resource cost model arise when schools in mandated reform protocols are compared to less formally supervised instructional programs. The four schools in the study were not required to cross reference costs from the variety of funding streams from which they draw resources. The costs presented were the best recollections of the school personnel and the documents at hand. This process under-reports the locally designed models costs, and more accurately reports costs from programs that had more stringent reporting criteria, namely the structured models. Districts that give little guidance in program design and program evaluation will have less refined structures for comparing all the elements in a program evaluation (Levin & McEwan, 2002). This is specifically true when using a cost-effectiveness ratio as an element of program evaluation.

Both *Reading First* and *SFA* used what could be termed a transparent cost reporting structure as part of reporting requirements for Title I funds and state grant awards. This level of transparency gave the reviewer clear budgeting and expenditure data while leveraging leadership to be accurate in both budgeting and expenditure reports. This process presented a more complete representation of federally funded programs. The accurate documentation in turn gives rise to more reliable cost-effectiveness ratios.

Local programs in this district did not have to function under these constraints. They had fewer documents recording their budgets and even fewer site level documents accounting for their expenditures. Local school site plans at times included budget values for reading program services, however the level of expenditure was not required (or

expected) in evaluating the success of action plans. The cost-effectiveness values presented for the locally designed models present two insights. First the low cost value most likely represents the under-reporting of costs due to inconsistent records. Secondly, when locally designed models are faced with complex populations (Hamilton's second grade with 65% free and reduced lunch rate) their cost per student reaching benchmark rises considerably.

Research Question 3

Research Question 3 looks at school leadership factors that might influence program review skills for improved instructional practice. It attempted to answer the question: What did teachers, program managers, and principals use as their decision-making process when programs were implemented? The survey of teachers and interviews of principals and program managers provided information about the use of cost analysis tools and school wide decision-making models.

School leadership model. All four schools used leadership models that limited teacher input about issues at the classroom level. A leadership team represented the staff in all four schools and teacher responses showed that the discussion about cost variables never got to their level, and were infrequent. They did recall talking about student outcomes and how they were measured; all cohorts said that discussion about program costs was undertaken at the leadership team level.

When principals and program managers were questioned, the majority could not recall spending a great deal of time discussing the costs of their programs. For the program with the most resources, the principal at Jefferson recalled that she discussed the

options with the two reading coaches and they made the budget decisions along with district personnel regarding the *Reading First* grant. The principal at Adams stated that she did not even put *SFA* up for a vote this year, they rolled it over and purchased the curriculum for implementation. These structures change in the discussions about staff development training as designed by the school site council of the grade level teams. By-in-large the schools in the study group are typified a top down leadership styles in reading program spending decisions contexts.

School personnel responses: Teacher responses to questions about the decision-making process at their school generated a variety of perspectives. The use of data is well understood by the participants at all levels in this study. From teachers to principals they understand that consistent use of student outcomes measures is important. In comparison of the structured programs to locally designed programs, the teaching staff at Jefferson and Adams talked more in terms of research-based practices than the teachers at Franklin and Hamilton. All schools used an equal number of assessments to gather information, however some schools' assessments were more formal than others.

When the questions turned to understanding cost variables associated with their programs all four teacher cohorts had little if any knowledge of resource variables associated with their programs. These results highlighted a cohort of professionals (teachers and program managers) not well versed in cost analysis skills. Chambers and Parrish (1994) found this to be one of the complicating factors explaining the differences in cost variable reporting. Program evaluation for these teacher cohorts was attended to through curricular practices (assessment and pedagogy) with little if any attention to costs

of resources in providing the program, and little understanding of the effects on concurrent academic programs. Program managers could not make the connections between costs and outcomes. For example, at Jefferson when literacy instructional time moved from 90 minutes per day to 120 minutes per day the program manager was concerned that other curriculum was not getting taught. However, no attempt was being made to define how much opportunity in other content areas was being lost with the additional literacy minutes. Both the programs manager and principal at Jefferson felt *reading is our top priority, without reading nothing else happens*. Yet indirect costs of the loss of staff development in other content areas, the lack of purchasing of new materials for other content areas, and devaluation of other content were not explored. This *make or break* mentality may be in the best interest of short-run reading assessment outcomes and reporting protocols, but could in the long run restrict student access to content and experiences associated with building background knowledge (Grissmer, 2002).

Specific costs variables beyond opportunity costs such as indirect costs were not well understood by program managers as a cohort. There is a singular lack of common vocabulary in cost analysis concepts. All four reading coaches / teacher leaders were recruited by their schools leadership from classroom assignments. The two locally designed formats had teacher leaders who were concurrently teaching 28 student classrooms while providing program oversight. In all four schools the melding of student outcomes with cost outcomes was not part of the decision-making process, or only partially discussed with teaching staff. To fully look at costs the leadership team would

require principals to lead this discussion. Yet in three of the four schools no process was in place. Both teaching staff surveys and interviews defined only one school, Franklin, as having pieces of a review process in place at the planning level.

For staffs to effectively look at their programs' successes or failures it takes more than disaggregating data on instructional outcomes (Kruse, 2001). Principals could be required to evaluate reform practices in light of cost variables linked to established reading outcomes. Jefferson's *Reading First* protocol is a good example; in terms of curricular review and training they are developing a highly effective staff for early grades instruction. Gersten, Chard, and Baker (2000) referred to this as developing the core-teaching structures of the staff for sustained impact of the reform. The *Reading First* approach reduces the autonomy of the teacher, but includes them in the decision-making process through more consistent training. Yet without consistent discussions at the staff level and clear understanding of what to review the school leadership team are ineffective in this area. District leadership that passes down the decision-making processes to the schools need to be aware that these variables are to be approached in a comprehensive manner. In this particular district the two local programs were given staffing and limited guidance (reduce class size in grades kindergarten through third) but were not given a plan on how to look forward in a planning process related to fixed costs in personnel resources, adjusting to changes in enrollment or changes in the workloads of existing teaching staff.

Research Question 3 Summary

To change the decision-making format in schools to a more distributed model the leadership practices would need to become less centralized in the leadership team. Teachers at the service provision level would have to understand the costs of the program and how they are measured. Using distributed leadership terminology of Spillane, Halverson and Diamond (2001), teachers would have to understand the *artifacts* of cost structures and be able to place them into context of the *technologies* of decision-making, integrating costs into the program evaluation structure. Teachers and program managers currently understand the assessment artifacts and technologies, and have integrated them into the program evaluation structures, so moving to understanding costs may not be much of a stretch in skill development.

The changes that need to happen are at both the district and site level. Fullan (2001) stated that a district level leadership goal is to provide support for local initiatives to assure staffs are connected with a vested interest in the outcomes. For this to be effective in answering the questions about effective use of community resources site staff needs increased training in the evaluation of programs. Specifically program manager roles filled by teachers from the classroom need to be re-considered. Schools that staff these positions melding data collection with coaching need to take three steps: 1) provide evaluation structures that are easy to link direct costs (personnel, staff development, equipment and materials) with indirect costs (planning time, changes in teacher schedules, changes in content presentation), 2) training site level principals in the integration of cost analysis variables in planning the years forward, 3) clear use of like

outcomes measures in the comparison of like student populations. The training in these three elements whether undertaken at the district level or in conjunction with university administrative leadership course work has the potential to take into account what Gersten, Chard, and Baker (2000) called the *core teaching structures* while at the same time developing leadership structures that increase teacher understanding of distributed leadership concepts. Distributing leadership task functions will increase internal pressure about cost structures and new artifacts for accountability and transparency requiring site leaders to explain program outcomes in more specific context.

Research Question 4

Research Question 4 looked at the complexity variables Barton (2003) defined as variables related to students opportunity to learn. This question is about schools who serve students from homes of economic disadvantage, with parents who had fewer resources in the form of academic skills to supporting classroom practices, speak a second language at home, or present a perspective that does not support homework practices.

The interviews of professionals in the four schools disclosed that all four schools perceived barriers to their reading programs success that were outside their immediate control. These barriers, represented in four areas, were seen by teachers, program managers and principals as critical to overcome if students were to reach grade level standards. The four areas were a) the increasing number of student who arrived in kindergarten with limited language development, b) on-going challenges with regular school attendance and mobility, c) the increasing number of students with severe

behavioral problems that interfere with instructional practices, and d) limited home support for programs that support the daily reading instruction.

Increasing numbers of students with underdeveloped language skills was the aspect most often stated by professionals. Barton (2003), Darling-Hammond (2004), and Spiegel (1995) linked this with pre-school opportunity to learn constructs around the availability of Head Start or early start programs for poor and minority students. Students who arrive in kindergarten with underdeveloped language structures in phoneme and grapheme representation, phonemic awareness and vocabulary have greater gains to make before accessing reading instruction (Foorman et al., 2001; Spiegel, 1995). The progressive nature of language development requiring all children to pass through virtually identical acquisition phases must be accelerated in schools such as Adams and Jefferson. Language acceleration requires time, and often by the end of second grade students from non-English speaking or economically disadvantaged homes are still mastering the language skills students in other communities have internalized. Leadership teams at Jefferson and Adams faced free and reduced lunch rates in the low 80 percent range, and have home languages other than English rates 45 percent to 75 percent higher than their comparison schools. This gap is further exacerbated by mobility rates in the 20 percent range. With one in five students not getting a full year of the academic program, outcomes variables aggregated at the school level can misrepresent accomplishments by both students and staff. Using language variables like phonemic awareness measures and phoneme segmentation outcomes in a growth model would more accurately represent the initial language challenges faced by teachers and how well they overcame them.

A second common set of characteristics was attendance and behavior. These two characteristics were not separated by participants, but in fact lumped into one concern. No program really had an on-going strategy unique to elementary schools to deal with these two concerns. All four schools had full-time counselors who work on these two issues. Adams and Jefferson used Title I funds to support attendance programs but principal comments stated they mostly were not effective in changing the behavior of the most challenging students.

In this study the mean attendance rate for all groups was statistically equal. The sample groups had no attrition from the study because of too few days in attendance. However, as a general classroom concern, teachers and program managers felt that attendance disrupted their instruction enough to make a difference. Foorman et al. (2001) and Hiebert and Taylor (2000) found that scaffolding instruction was difficult for students who were in and out of school, the lack of continuity of instruction slowed metacognitive skills development.

All program managers (each with eight years or more of teaching experience) stated the number of severe behavior concerns was on the rise in primary classrooms. Extreme student defiance and increasing numbers of students on individualized education plans for special education are presenting themselves as challenges compared to previous teacher and principal experience. These students change both the large group and small group instructional dynamic for teachers. This element is poorly represented in the literature about young students. However, this clearly represented in discussions about minority students and boys over representation in special education programs (Darling-

Hammond, 2004). The program managers in their interviews had no real solutions, however, these students increase the complexity of a school community and could present unrealistic challenges for teachers attempting to get all student to grade level benchmarks.

Finally each school staff defined a category that was related to the interplay between home and school in the development of early literacy skills. This variable revolves around two sets of concerns found in all four schools. First is the capacity in the home to carry out the supportive tasks associated with the schools reading program, and secondly was the perception of the importance of homework tasks as viewed by the family.

Parent capacity to sit with a child for side-by-side reading is seen as a critical skill in school associated with *take-home reading* programs. Two schools had some form of *take-home reading* program in which students had a reading book that went home each night for sharing with the family. When the child returned the next day a volunteer or teacher sat with them in the morning and covered the same book. At Jefferson and Franklin the program managers viewed this as critical practice for their struggling students, but both program managers were concerned that parent literacy and language skills were limited and unable to support children growth through the use of this type of support. The gap in early reading skills was not closing as fast for these students as for their primarily English speaking peers (Pinnell et al., 1994).

The second element was counter-acting a belief system that the work of teaching was the schools responsibility and not a parent role, but that of the school. Program

managers in all four schools and two principals mentioned this among the group of families associated with generational poverty. Anecdotal accounts of conversations represented this barrier as non-supportive parents who expected the school to *fix* the problems at school, because they did not have the time at home to do what the school was suppose to complete.

Research Question 4 Summary

Barton's (2003; 2004) variables related to the student achievement gap are grouped into three organizing sets: 1) children prenatal and early years health, 2) nature of the home life, and 3) nature of schooling. Measures related to children's health: a) low birth weight, b) lead poisoning, c) hunger, and d) nutrition both in pre-natal and pre-school years. A second set of measures is related to parenting practices: a) reading to young children, b) television watching, c) parent reading practices, d) student mobility, e) marital status, and f) parent participation in school. The final set of variables is related to schooling: a) rigor within the curriculum, b) teacher experience, c) teacher preparation, d) class size, e) technology assisted instruction, and f) school safety. This final set is seen as possible policy influenced variables to control opportunity to learn variables (Barton 2003, 2004). Although the first two sets of measures are difficult for schools to influence, their importance is of concern for educators who attempt to account for all program costs in the schools service provision that affects academic outcomes.

Systemic reform initiatives attempt to generate standards for education based on these variables and at times integrated these standards into the home setting. Thus some reform initiatives support parent liaisons connecting the family and school for dealing

with behavioral and attendance concerns, training resources for parents around early literacy skill practice, and a variety of community events that attempt to integrate families with their schools to soften the effects of cultural differences (Darling-Hammond, 2004; Herman, Klein, and Abedi, 2000). These student complexity variables contribute to the challenges of diverse schools and are just one of the four factors effecting leadership team decision-making.

In looking at program effects and the cost associated with these categories a broad perspective is required to see the distributed costs of working with variables that are not easily controlled. The time and effort of managing the initial gaps in learning and then accelerating student progress are critical criterion in explaining the success of a particular reading approach. This is where intervention hours under the construct of distributed costs could serve the reviewer looking at alternative programs. The time and training required to work with families can get lost when reviews only consider instructional minutes as resources used. Parent conferences, informal training of parents, costs for setting up home to school communication, counselor and administrative time for attendance and discipline concerns are all variables that are challenging for schools to quantify in the ingredients or resource cost models.

For example the distributed costs of a program may not totally rely on a percentage of students within aggregated categories, but size of school and spread of student abilities that are associated with complexity variables. The assumption is that principals and teachers in schools with larger minority, poverty and second language populations have greater challenges. It was beyond this study's reach to answer that

question. However, what was observed in the interviews is that all schools see the same barriers and have localized strategies to manage them. It could be that a school of 527 predominately middle class students with a free and reduced lunch rate of 34% and 10% mobility is dealing with the same number of concerns around attendance, contrary parents, and second language learners as a school of 327 students whose families are predominately at poverty level and who may not speak English at home.

The quantification of complexity variables contributing to the existing achievement gap is not overly difficult in many public schools in the U.S. However, weighing their effects on student outcomes and program costs is not easy to control for. This study found initial similarities in outcome measures in the pre-test period, but significant differences at the end of the year. This relationship was consistent with the persistent nature of the achievement gap as represented by Rothstein (2001), Barton (2003), and Darling-Hammond, (2004). School staffs recognized these variables in four specific contexts that limited their effectiveness with students. However, they did not recognize the distributed costs associated with their reading programs.

The outcome variables associated with each program have to be viewed through the lens of the complexity variables. The program evaluation protocol has to be viewed as a category of information to be used in relation to the context the reform is going to be used (Grissmer, 2002, Levin & McEwan, 2002). Schools with significant complexity variables must look at schools in like circumstances, taking the elements of the reading programs, the costs associated with the program, and bringing these variables into a coherent decision-making process to make the best choice for their schools context.

Research Question 5

The program fidelity measures sought to understand if the site managers and principals felt the reading programs in place were being used as intended. Program fidelity literature makes the distinction between *adoption* and *adaptations*. Early work by Fullan and Pomfret (1977) defined fidelity as the measurable difference between the designed use and the actual use based on a consistent variable. Mills and Ragan (2000) distinguish implementation fidelity as the placement of an innovation into the instructional process, whereas adaptation is the use of an innovation in a manner similar to but not equal to the intentions of the designer. The degree to which teachers implemented the agreed upon programs of *Success For All* (SFA) and *Reading First* will help the reviewer understand how the programs attributes contribute to the students' reading outcomes. The degree of perceived program fidelity informs the review of alternatives when linked with the cost-effectiveness ratio the program produced.

Using the teacher survey and the interviews of school personnel the program participants provided self-ratings in teacher faith in the program, number of lesson modified weekly, and the use of supplemental materials in the course of teaching reading. These measures were followed up with questions about the oversight of the programs and who was responsible for staff development and instruction.

The two structured programs had the highest self rating by school personnel, the fewest changes per week in lesson, and fewer teachers using supplemental materials. The focus was upon the two structured programs because of their target populations. *Success For All* (SFA) and *Reading First* are meant to serve under performing populations. The

complexity variables discussed in Research Question 4 are directly supported in the program implementation with both personnel in the form of a community liaison and in practice in the form of more explicit instruction that uses scaffold processes in the lesson design.

SFA at Adams school received strong ratings from the program manager and the principal in their self-assessment of fidelity. However, two observations must be made: 1) when asked about the use of the community liaison for attendance and behavior concerns (as prescribed by the *SFA* foundation) the principal stated these two things were not the role of the liaison at Adams, 2) The *SFA* foundation had loosened oversight while the teacher augmented lessons 20-30 percent of the time. Both of these elements are affecting implementation due to: a) comments by both the program manager and teachers that their second greatest concern about their students' success was the behavior and attendance of their students, b) adaptations by the teacher erodes the publisher's intentions, and c) instructional assistants planned lesson and led groups on a daily basis.

Adams has been using *SFA* for nine years and getting about 66 percent of their student to reading benchmark in its current year. The foundation feels Adams is running well and although they do come in annually and help with data evaluation and support the training of new staff, their relationship with the school was characterized as following the lead of the program manager rather than prescriptive in an attempt to get 100 percent of students to district reading benchmarks.

Jefferson's *Reading First* protocol had better control on fidelity than Adams due to program oversight by the reading coaches. This oversight is seen in three areas: a) the

use of consultants in a staff development format to shape instructional practices, b) the use of two program managers in oversight roles to support the implementation of staff development strategies and c) use of reading coaches to interpret and disseminate progress monitoring data. Jefferson had three coaching positions that supported student growth with small group instruction and side by side teacher observations and lesson modeling. The intensity of support and oversight links new teachers (of which Jefferson had four of eight) with resources and guidance in teaching practices.

Using the four variables measured at the teacher level Jefferson shows strong teacher faith in the program and supported links between the trained instructional format, teacher adherence to lessons formats, minimal use of supplemental materials, and good guidance of instructional assistants. These variables are linked to explicit instruction criteria referenced by Foorman et al. (2002). Teacher training and use of materials directly addressed the areas of increased intensity and comprehensive approach to primary reading instruction. The coaches assure the consultants trainings are implemented and that teacher adaptations are within the scope of the schools overall program.

SFA at Adams, is implemented in a school about two-thirds the size of Jefferson, the *Reading First* school, and with a budget about one-half the size. Yet *SFA* achieves the same levels of essential comprehensive reading variables. However, Adams is less in line with the literature in the size of student groups, Adams group sizes of 10-12 are double the size of all the other schools. They balance the larger groups with longer times in small group (30-45 minutes) in the instructional setting where the other programs had

longer times in large group reading instruction but shorter times in small groups. So Adams intensity measures reflect that state of current grouping research. Shanahan (1998) found that the more intense instruction over shorter periods was more effective than longer instructional settings. That seems to be the case here. Care needs to be taken to look at the sample size at Adams when drawing the conclusion, dropping group size below 10 would increase their cost while increasing and potentially their effectiveness.

The two locally designed programs had fewer controls on teacher practice and structure to curricular materials. Both schools use the district provided curriculum with supplemental materials. In Franklin's case the supplemental materials are a result of having a *Reading Recovery* teacher prior to the district providing a reading specialist. This historical context left approaches to teaching reading in leveled book materials with some basal support in place. At Hamilton teachers supplement the district adoption through the use of an outside publisher. Both programs were faithful to this district guide to reading target skills, and Franklin adhered closely to the critical elements of the primary reading program, including on-going assessment strategies.

Hamilton's program fidelity measure fit to a loose program design where teachers' autonomy was well established. They adhered to intensity elements in grouping and some elements of explicit instruction, but both survey responses and interview responses showed very little use of data collection strategies and little if any coordination in their literacy model. They had few oversight practices, however, the principal gave a great deal of control to the teacher leader to run the decision-making process.

Franklin had stronger approach to a comprehensive program, but weaker controls on the variables of explicit instruction. They increased intensity with the use of small groups yet the training of the instructional assistants in those groups was not as directed and continually supported as in the two structured programs.

Franklin's comprehensive variables compared quite well with the research. They utilize on-going assessment that could be used across the grade levels, integrated writing with their reading tasks, used a reading specialist for their most challenging students and linked out of class supports with what was happening in the class. Franklin was the only school with volunteer support. The volunteers are used along with daily take home reading program with a volunteer listening to students reading each day.

Research Question 5 Summary

In conclusion adherence to program characteristics was rated high by teachers, program managers, and principals in the two structured programs, but their gains were not overly impressive when pre-test to post-test scores were evaluated in comparison with non-structured schools. The use of reading coaches to oversee the implementation of the program staff development strategies and new curriculum generated improved outcomes and generated faith from the teachers in the programs ability to raise their student reading levels. The benefit for both programs was ongoing oversight by well trained professionals, improved program strength through two venues a) increased modeling of content by consultants and reading coaches, and b) increased modeling of instructional strategies. Based on the school personnel responses in the two structured programs both schools are seen as implementing the curriculum as intended by the designer and defined

in the literature (Fullan & Pomfret, 1977; Mills & Ragan, 2000). Adams appears to have permission from the foundation to modify its format, and thus can be seen as implementing *Success For All* with good faith.

The locally designed programs had agreements within their staff to pursue a program structure, but left instructional choices to the teachers. Franklin had better controls on the elements of comprehensive programs like assessment, but Hamilton and Franklin supplemented the districts intended curriculum with outside materials on a regular basis.

The four programs had varying degrees of oversight, which in the case of the structured programs contributed to overall implementation fidelity. Both structured schools achieved 66 percent of their kids reaching the benchmarks using intensive oversight to assure fidelity. In the locally designed programs oversight was limited by the decision-making processes of the leadership teams. Franklin chose to use the reading specialist as a small group leader. This may affect the 20 percent of students who did not reach benchmark by permitting teacher autonomy to focus on elements that are not directed at their needs. This is even more likely at Hamilton where 45 percent of the students did not meet benchmark. In the review of alternatives process necessary for good comparisons of cost-effectiveness ratios evaluators could look closely at the two structured models and know how the fidelity issues relate to school outcomes. This is less likely with the two locally designed models.

Implications of the Findings

The findings show that the comparison of four reading models using site level documents, personnel interviews and surveys can generate pictures of school reading programs. These reviews can then be used to generate cost-effectiveness ratios for the comparisons of programs. Using data from the reading programs, the four models can then be compared for their ability to support students reaching an agreed upon standard. Once the schools have the ability to generate reliable and valid results with accurate costs variables, then they can make academic judgments about their programs that fit their context. The findings indicate that it is necessary to look at student complexity variables (such as, free and reduced and/or second language) and generate information about program fidelity characteristic before making a final judgment about what program to pursue.

Issues to be considered are the structures in place to effectively record cost variables. The schools in this study did not have accurate representations of costs in all ingredients categories, nor did they track costs over time. Both shortcomings limited the completeness of the cost-analysis by excluding important concepts of opportunity costs, distributed costs, and discounted costs. Although these costs certainly support the decision-making process, they may be overly technical for site level leadership to undertake. The important implication is that cost variables can always be disregarded by decision-making entities, but costs analysis ought to be present in the decision-making process.

Limitations

This study provides a structure that can be improved upon in further research. One of the initial limitations came from the choice of sample schools. As a sample of convenience they provided an opportunity for the researcher to attempt the use of set common reading outcomes over 120 to 140 students. However, one of the schools with split English and Spanish reading programs limited the sample counts and the resulting small numbers of participants limited the analysis of some reading outcomes at the post-test measurement. The sample schools were not exactly true representations of each other, yet they represented structured and locally designed programs, and represented the district in which the schools were located. Generating cost for locally designed and structured programs was necessary, but a sample set of four schools limits the ability to generalize the findings to a broader population. Future research should have two or three schools in each of the categories to provide more accurate cost-effectiveness ratios for a complete review of instructional program alternatives.

Secondly, this research did not look to disaggregate the academic data by demographic data. This research would have been enhanced by being able to compare the reading outcomes by various sub-populations. However, because of small cell size, this statistical analysis was not completed. Because of NCLB, most schools are very interested in sub-population differences and knowing how the reading programs and costs are associated with sub-groups would inform the scope of program evaluations.

Lastly, the school personnel in the study had limited experience with cost analysis and therefore could not really support findings about discounted, distributed, or

opportunity costs. This required an over reliance on documents which most certainly over reports the structured programs with guidelines leveraging reporting of expenditures and requiring annual budgets. The cost-effectiveness ratios represent this accurate cost accounting, which is distinctly different from the locally designed programs requirements.

Conclusions

The findings of this study demonstrate that cost analysis currently plays a limited role in informing the school site leadership teams about how their current reading programs are generating the outcomes they anticipate. Leadership teams do work with curricular level information about reading outcomes and staff development training in effective pedagogy. However, principals do not track resource use effectively to anticipate distributed costs, calculate opportunity costs, and inform a review of alternatives. Schools are required to file school improvement plans on annual or bi-annual basis, but those reviewed did not contain budgets or expenditure formats that allowed for local review of goal attainment. The ingredients model or the resource cost model have the potential to support the three factors identified by Hummel-Rossi and Ashdown (2002) in effective program evaluations: a) using like measures to compare alternatives, b) clearly defining the characteristics that differentiate the alternatives, and c) the analysis of costs for developing cost-effectiveness ratios.

With considerations of student complexity and implementation fidelity variables a complete analysis can provide community groups a clear picture of what they are doing, or what the reform they are considering would have them do. This degree of information

used in a distributed leadership format has the potential to facilitate the melding of current artifacts and technologies with new expectations to assure the sustainability of a reform.

The principal cohort of this study had limited working knowledge of program evaluation strategies, especially the cost-analysis processes. Future study should address potential shortcoming to see if it is a general tendency. Future research also should look at time as a variable. Of interest to many districts is whether one academic program can get students to mastery level faster than other programs for less cost. The increased pressures to use research based programs, the use of data to drive instruction, and the need to communicate school improvement effectively to the public requires principals respond to policy initiatives with strong leadership skills and effective evaluations tools.

APPENDIX A

PRINCIPAL AND PROGRAM MANAGER INTERVIEW QUESTIONS FOR
 UNDERSTANDING FOUR PRIMARY GRADE READING PROGRAMS

*Comparison of Costs and Reading Outcomes in
 Four Primary Grade Reading Programs for Improved Decision-Making*

Narrative for Interview:

Read by researcher: *“As you are aware, I am researching how four schools approach primary grade reading instruction from a perspective of the costs of school reform. I am looking at a set of primary grade reading outcome measures in relation to leadership teams’ decision-making processes, and how schools arrived at the curriculum choices / instructional model for implementation. These questions attempt to help me understand these issues and explore what the research literature points to as successful instructional elements. Please feel free to “pass” on any of the questions if you’d like, or add perspectives you feel will help me understand what your school has been through to get you to where you are at this time.”*

Questions for Administrators and Program Managers

School Pseudonym _____ Principal Pseudonym _____

1. How many funding streams do you use to fund your reading program? Which of these streams are temporary and which are permanent?
2. How much time do you spend defining costs with your leadership team? Do you also take time with the staff overall?
3. How do you measure cost-effectiveness of the reading program you are implementing? What are the measures you choose to measure?
4. What cost variables are associated with this program’s implementation?
5. What do you think are the greatest distributed costs associated with your reading program?

6. Resource costs categories that you have considered in your funding worksheets include what variables? (Personnel, instructional materials, facilities space, equipment, and unique program needs).
7. Do you factor in costs incurred in past years of the implementation of the reading program?
8. Have you used an “opportunities lost” measure to determine program cost variables?
9. Of what use is cost-effectiveness ratio in support of your school based leadership?
10. Does your program have an external oversight component as a program resource?
11. To what degree do current available personnel play into the program choices you make for the reading program instruction and implementation?
12. Are there costs variables you feel you are losing control of in your current reading program?
13. To what degree do you feel your staff is consistent in implementing the curriculum protocols as they were designed by the reform developer?
14. How do you allow for individual teacher differences in the implementation of the reading program?
15. Do staff development opportunities in reading instruction provide learning that has been implemented by teachers?
16. How are instructional assistants used in your approach to teaching primary reading?
17. How do you assure the fidelity in implementing the reading program you are currently using?
18. What criteria did you use to choose the approach that you now use for primary reading instruction? How many years have you been using this curriculum?
19. How many alternatives did you consider before choosing the reading reform format you now utilize?
20. Are there curricular areas that you feel have gotten less attention because of your school’s particular approach to teaching reading?
21. What other reading instruction support is available for students who are behind on their grade level measures? When are these resources available to students?

22. How are or were teachers involved in your curriculum decisions about your reading program?
23. What are the variables you feel make up the complexity of your students?
24. Which student characteristic (demographic or otherwise) would you weigh as contributing most to your students challenges in acquiring early reading skills?
25. What complexity variables of your below grade level students do you think contribute to their resistance to learning and requiring the greatest level of intervention?
26. Is the student population growth rate in your school such that your costs are growing when your budget is not?
27. If you had to change reading programs for next fall, where would you start today? What are five critical steps you feel you would have to take?

APPENDIX B

TEACHER SURVEY QUESTIONS FOR UNDERSTANDING FOUR PRIMARY
GRADE READING PROGRAMS*Comparison of Costs and Reading Outcomes in
Four Primary Grade Reading Programs for Improved Decision-Making**Teacher Survey*

Directions: Please fill in the blank spaces with numerals answers and check (✓) the appropriate boxes as they apply. Some answers can have more than one response. When you have completed please place in the attached return addressed envelope and place in the regular post. Thank you for your participation.

1. I have taught for ___ years.
2. I have taught at my current grade for ___ years.
3. I currently teach ___ grade.
4. I have taught primary grades for ___ years.
5. I am a member of our schools Leadership Team. Yes No
6. My highest degree earned:
7. B. S. / B. A. M. A. / M. Ed. / M. S. D. Ed. / Ed. D. / Ph. D.
8. I have a reading endorsement along with my teaching credential. Yes No
9. I have used the current reading program for ____ years.
10. How many students are in your class? ____
11. Of these students how many are reading below grade level at this time? ____

12. Our reading instruction time is protected from interference by a building agreement?
 Yes No

13. Does your grade level ability groups for reading instruction? Yes No

14. Is this a school wide practice? Yes No

15. Our school uses a reading specialist to teach our lowest level students.
 Yes No

16. Our reading specialist uses:
 Predominately small group format to teach our lowest students. Yes No
 Predominately individualized format for our lowest students. Yes No

17. Individualized tutoring is a daily activity for my most struggling readers?
 Yes No

18. What measures do you use to understand your students for reading instruction level?

Check all that apply

- | | |
|-------------------------------------------------------------------|-----------------------------------------------|
| <input type="checkbox"/> Basal reading assessment | <input type="checkbox"/> Sight word lists |
| <input type="checkbox"/> DIBELS | <input type="checkbox"/> Reading passage |
| <input type="checkbox"/> Running records | <input type="checkbox"/> San Diego Quick Test |
| <input type="checkbox"/> Primary Literacy Assessment | <input type="checkbox"/> Other |
| <input type="checkbox"/> Qualitative Individual Reading Inventory | _____ |

19. Are these same measures used to group students for instruction? Yes No

20. Our kids are assessed regularly for reading gains?

- | | | |
|-----------------------|------------------------------|-----------------------------|
| Phonics skills | <input type="checkbox"/> Yes | <input type="checkbox"/> No |
| Phonemic Awareness | <input type="checkbox"/> Yes | <input type="checkbox"/> No |
| Vocabulary | <input type="checkbox"/> Yes | <input type="checkbox"/> No |
| Reading Comprehension | <input type="checkbox"/> Yes | <input type="checkbox"/> No |
| Reading Fluency | <input type="checkbox"/> Yes | <input type="checkbox"/> No |

21. We had discussions about the specific reading outcomes we wanted when considering the different reading program?

Yes No

22. During my reading groups I teach small groups for ___ time segments.

10-15 15 – 20 20 to 25 25 to 30 more than 30 minutes.

23. I see each of my small groups

5 days/week 3-4 days/week 2-3 days/week 1-2 days/week

24. The number of students in my small groups is

3-4, 4-5, 5-6, 6-7, more than 7 students

25. Our instructional approach is built around leveled book materials?

Yes No

26. I regroup my reading groups:

once / month, more than once / month less than once / month

27. I participated in the decision-making process for the current reading program?

Yes No

28. I was able to look at a variety of reading programs before we choose one for our school?

Yes No

29. We had discussions about the specific costs of the different reading program?

Yes No

30. We balanced costs and the anticipated student outcomes of the program we reviewed when making our decision?

Yes No?

31. I have faith in our reading materials ability to raise all our general education students reading scores up to grade level standards? Yes No

32. Staff development has helped me to be more confident in the implementing of our reading? Yes No

33. I understand many of the costs of the reading program that we are currently using?

Yes No

34. I use the building agreed upon reading curriculum?

90 to 100 % of the time, 70 to 80% of the time,
 80 to 90% of the time 60 to 70 % of the time
 I do not use the building 50 to 60% of the time,
agreed upon curriculum.

35. I currently spend more / less time in:

Reading planning for instruction	<input type="checkbox"/> more time	<input type="checkbox"/> less time
Assessing reading instruction	<input type="checkbox"/> more time	<input type="checkbox"/> less time
Filling out program reports	<input type="checkbox"/> more time	<input type="checkbox"/> less time
Correcting reading instructional materials	<input type="checkbox"/> more time	<input type="checkbox"/> less time

36. I modify or completely change the current reading program lessons:

3 or more times /week 2 times /week 1 time/week

37. The IAs I use follow my instructions in reading instruction? Yes No

38. IAs plan their own work for reading instruction? Yes No

39. Our measurement for reading gains is linked to our instructional practices?
 Yes No

40. I have to use other resources to teach concepts the adopted curriculum does not present well. Yes No

41. The tutoring my students receive links closely with what is taught in the daily classroom lessons.
 Yes No

42. The Instructional Assistants I use for primary reading activities are trained in:

Running small groups	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Assessing student skills	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Questioning strategies for comprehension	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Vocabulary skill development	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Primary Literacy Assessment	<input type="checkbox"/> Yes	<input type="checkbox"/> No
DIBELS	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Phoneme segmentation activities	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Phonics instruction	<input type="checkbox"/> Yes	<input type="checkbox"/> No

43. Our reading program has a focus on vocabulary?
 3 or more times/week 2 times/week 1 time/week

44. Our reading program explicitly teaches phonics skills?
 3 or more times/week 2 times/week 1 time/week

45. Writing is an integral part of the daily reading lesson? Yes No

REFERENCES

- Allington, R. L. (2004). Setting the record straight. *Educational Leadership*, 61(6), 22-25.
- Allington, R. L. (2006). Reading lesson and federal policy making: An overview and introduction to the special issue [Special issue]. *The Elementary School Journal*, 107(1), 3-15.
- Barone, D., Hardman, D., & Taylor, J. (2006). *Reading First in the classroom*. Boston: Pearson.
- Barton, P. E. (2003). *Parsing the achievement gap: Baselines for tracking progress*, Princeton, NJ: Educational Testing Service.
- Barton, P. E. (2004). Why does the gap persist. *Educational Leadership*, 62(3), 8-13.
- Berg, B. L. (2004). *Qualitative research methods for the social sciences* (5th ed.). Boston: Pearson
- Berninger, V. W., Vermeulen, K., Abbott, R. D., McCutchen, D., Cotton, S., Cude, J., et al. (2003). Comparison of three approaches to supplementary reading instruction for low achieving second grade readers. *Language, Speech, and Hearing Services in Schools*, 34, 101-116.
- Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2001). *Cost-benefit analysis: Concepts and practice*. Upper Saddle, NJ: Prentice Hall.
- Borman, G. D. (2005). National efforts to bring reform to scale in high-poverty schools: Outcomes and implications. In L. Parker (Vol. Ed.), *Review of Research in Education*, vol. 29, (pp. 1-27). Washington DC: American Education Research Association.
- Borman, G. D., & Hewes, G. M. (2002). The long-term effects and cost-effectiveness of Success For All. *Educational Evaluation & Policy Analysis*, 24(4), 243-266.
- Brent, B. O., Sipple, J. W., Killeen, K. M., & Wischnowski, M. W. (2004). Stalking cost-effective practices in rural schools. *Journal of Education Finance*, 29(3), 237-56.

- Borman, G. D. (2005). National efforts to bring reform to scale in high-poverty schools: Outcomes and implications. In L. Parker (Vol. Ed.), *Review of Research in Education*, vol. 29, (pp. 1-27). Washington DC: American Education Research Association.
- Borman, G. D., & Hewes, G. M. (2002). The long-term effects and cost-effectiveness of Success For All. *Educational Evaluation & Policy Analysis*, 24(4), 243-266.
- Brent, B. O., Sipple, J. W., Killeen, K. M., & Wischnowski, M. W. (2004). Stalking cost-effective practices in rural schools. *Journal of Education Finance*, 29(3), 237-56.
- Chambers, J. G. (1999). Measuring resources in education: From accounting to the resource cost model approach. (Report No. NCES-WP-1999-16). Kensington, MD: American Institutes for Research (CRESS). (ERIC Document Reproduction services No. ED433613).
- Chambers, J. G., & Parrish, T. (1994). Modeling resource costs. In H. J. Walberg and W. S. Barnett (Eds.) *Advances in educational productivity. Cost analysis for education decisions: methods and examples*, vol. 4, (p. 7-21). Greenwich, CT: JAI Press
- Conley, D. T., (2003). *Who governs our schools? Changing roles and responsibilities*. NY: Teachers College Press.
- Conley, D. T., & Picus, L. O. (2003). Oregon's quality education model: Linking adequacy and outcomes. *Educational Policy*, 17(5), 586-612.
- Darling-Hammond, L. (2004). Standards, accountability, and school reform, *Teachers College Record*, 106(6), 1047-1085.
- Elbaum, B., Vaughn, S., Hughes, M., & Watson-Moody, S. (2000) How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92, 605-619.
- Flanigan, J. L., Marion, R. A., & Richardson, M. D. (1997). Causal and temporal analyses of increased funding on student achievement. *Journal of Research and Development in Education*, 30(4), 222-247.
- Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidenced-based approach. *Developmental Neuropsychology*, 24, 613-639.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90, 37-55.

- Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small group reading instruction promote success in all children. *Learning Disabilities Research and Practice, 16*, 203-212.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research, 47*(1), 335-397.
- Gersten, R., Chard, D., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33*(5), 445-457.
- Goertz, M. E., & Duffy, M. C. (1999). Resource allocation in reforming schools and school districts. In M. E. Goertz and A. Odden, (Eds.), *School-based financing. Twentieth yearbook of the American Education Finance Association* (pp. 215-244). Thousand Oaks: Corwin.
- Goldman, P., & Conley, D. T. (1994) Facilitative leadership how Principals lead without dominating, *Oregon School Study Council, 37*(9) 1-43.
- Grissmer, D. (2002). Cost-effectiveness and cost-benefit analysis: The effect of targeting interventions. In H. M. Levin and P. J. McEwan (Eds.), *Cost-Effectiveness and Educational Policy*, (pp. 97-110). Larchmont, NY: Eye on Education.
- Gronn, P. (2000). Distributed properties: A new architecture for leadership. *Educational Management and Administration, 28*, 317-334.
- Harris, A. (2002). Effective leadership in schools facing challenging contests. *School Leadership and Management, 22*, 15-26.
- Harris, A. (2003). Teacher leadership as distributed leadership: Heresy, fantasy or possibility. *School Leadership and Management, 23*, 313-324.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies on the effects of differential school inputs on student outcomes. *Educational Researcher, 23*(3), 5-14.
- Hedrick, W. B., & Pearish, A. B. (1999). Good reading instruction is more important than who provides the instruction or where it takes place. *The Reading Teacher, 52*, 716-726.
- Herman, J. L., Klein, D.C.C., & Abedi, J. (2000). Assessing students' opportunity to learn: Teacher and student perspectives, *Educational Measurement: Issues and Practice, 19*(4), 16-24.

- Hiebert, F. H., & Taylor, B. M. (2000). Beginning reading instruction: Research on early interventions. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, R. Barr (Eds.). *Handbook of reading research: Volume III* (pp. 455-482). Mahwah, NJ: Lawrence Erlbaum.
- Homan, S., King, J. R., & Hogarty, K. (2001). *A small group model for early intervention in literacy: Group size and program effect*. University of Southern Florida: Tampa FL. (ERIC Document Reproduction Service No. ED461095)
- Hummel-Rossi, B., & Ashdown, J. (2002). The state of cost-benefit and cost-effectiveness analyses in education. *Review of Educational Research*, 72(1), 1-30.
- Ilon, L., & Normore, A. H. (2006). Relative cost-effectiveness of school resources in improving achievement. *Journal of Education Finance* 31(3), 238-254.
- Jefferson, A. L. (2005). Student performance: Is more money the answer? *Journal of Education Finance*, 31(2), 111-124.
- Kame'enui, E. J. (Eds.). (2002). *An analysis of reading assessment instruments for K-3*. Report prepared for members of the Assessment Committee. Eugene, OR: Institute for the Development of Educational Achievement.
- King, J. A. (1994). Meeting the needs of at-risk students: A cost analysis of three models. *Educational Evaluation and Policy Analysis*, 16(1), 1-19.
- King-Rice, J. (1997). Cost analysis in education: Paradox and possibility. *Educational Evaluation and Policy Analysis*, 19(4), 309-317.
- Kinnucan-Welsch, K., Magill, D., & Dean, M. (1999). Strategic teaching and strategic learning in first-grade classrooms. *Reading Horizons*, 40, 5-21.
- Krashen, S. (2004). False claims about literacy development. *Educational Leadership*, 61(6), 18-21.
- Kruse, S. D. (2001). Creating communities of reform: Continuous improvement planning teams. *Journal of Educational Leadership*, 39, 359-383.
- Leslie, L. & Allen, L. (1999). Factors that predict success in an early literacy intervention project. *Reading Research Quarterly*, 39(4), 404-424.
- Levin, H. M., McEwan, P. J. (Eds.). (2002) Cost-effectiveness and education policy. *Cost-effectiveness and educational policy* (pp. 1-16). Larchmont, NY: Eye on Education.

- Levin, H. M. (2002). The cost effectiveness of whole school reforms. (Urban Diversity series No. 114) ERIC Clearinghouse on Urban Education, Institute for Urban and Minority Education.
- Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis: Methods and applications*, (2nd ed.). Thousand Oaks: Sage.
- Lyon, G., & Chhabra, V. (2004). The science of reading research. *Educational Leadership*, 61(6), 12-17.
- Marion, R., & Flanigan, J. (2001). Evolution and punctuation of theories of educational expenditure and student outcomes. *Journal of Education Finance*, 26, 239-258.
- Marks, H., & Seashore-Louis, K. (1999). Teacher empowerment and the capacity for organizational learning. *Educational Administration Quarterly*, 35, 707-750.
- Mills, S. C., & Ragan, T. J. (2000). A tool for analyzing implementation fidelity of an integrated learning system. *Educational Technology Research and Development*, 48(4), 21-41.
- Monk, D. H. (1989). The education production function: Its evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, 11(1), 31-45.
- Monk, D. H. (1995). The cost of pupil performance assessment: A summary report. *Journal of Education Finance*, 20, 363-371.
- Morrow, L. M., & Gambrell, L. B. (2000). Literature-based reading instruction. In M.L. Kamil, P. B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 563-586), Mahwah, NJ: Lawrence Erlbaum.
- National Reading Panel (NRP). (2000). *Teaching children to read: An evidenced-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- Nelly, E., & Smith, A. (1999). *Rigby benchmark kit 1 teacher's notes*. Barrington, Ill: Rigby.
- Nelly, E., & Smith, A. (2002). *Rigby benchmark kit 2 teacher's notes*. Barrington, Ill: Rigby.
- Oregon Department of Education (ODE) (2006). *Oregon school report cards*. Salem, OR: Author. Retrieved June 4, 2006 from <http://www.ode.state.or.us/data/reportcard/reports.aspx>

- Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Solomon (Ed.), *Distributed cognitions* (pp. 47-87). Cambridge, UK: Cambridge University Press.
- Picus, L. O. (1999). Site-based management: A nuts and bolts approach for school administrators. In M. E. Goertz and A. Odden, (Eds.), School-based financing. Twentieth yearbook of the American Education Finance Association (pp. 22-39). Thousand Oaks: Corwin.*
- Pikulski, J. L. (1994). Preventing reading failure: A review of five effective programs. *The Reading Teacher*, 48, 30-39.
- Pinnell G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly*, 29, 9-38.
- Potter, M. L., & Wamre, H. M. (1990). Curriculum-based measurement and developmental models: Opportunities for cross-validation. *Exceptional Children* 57(1), 16-25.
- Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2(2), 31-74.
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systemic approach* (6th ed.). Thousand Oakes: Sage.
- Rothstein, R. (2001). Social class, student achievement, and the black-white achievement gap. *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*. New York: Economic Policy Institute, Teacher College, Columbia University.
- Schwandt, D. R. (1995). Learning as an organization: A journey into chaos. In S. Chawla and J. Renesch (Eds.) *Learning organizations: Developing cultures for tomorrow's workplaces* (pp. 365-379). NY: Productivity.
- Scribner, J. P., Cockrell, K. S., Cockrell, D. H., & Valentine, J. W. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly*, 35, 130-160.
- Seashore-Louis, K., Marks, H., & Kruse, S. (1996). Teachers' professional community in restructuring schools, *American Educational Research Journal*, 33, 757-798.

- Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. In P.D. Pearson, & A. Iran-Nejad (Eds.), *Review of Research in Education, Vol. 23* (pp. 217-234), American Educational Research Association: Washington DC.
- Skindrud, K., and Gersten, R. (2006). An evaluation of two contrasting approaches for improving reading achievement in a large urban district. *The Elementary School Journal, 106*(5), 389-407.
- Slavin, R. E., & Madden, N. A. (2001a). *One million children: Success for all*. Thousand Oaks, CA: Corwin.
- Slavin, R. E., & Madden, N. A. (2001b). Success for all: An overview. In R. E. Slavin & N. A., Madden (eds.), *Success for all research and reform in elementary education*, (pp. 3-15). Mahwah, NJ: Lawrence Erlbaum.
- Smylie, M. A., Lazarus, V., & Brownlee-Conyers, J. (1996). Instructional outcomes of school-based participative decision making. *Educational Evaluation and Policy Analysis, 18*, 181-198.
- Spiegel, D. L. (1992). Blending whole language and systemic direct instruction. *The Reading Teacher, 46*, 38-44.
- Spiegel, D. L. (1995). A comparison of traditional remedial programs and Reading Recovery: Guidelines for success for all programs. *The Reading Teacher, 49*, 86-96.
- Spillane, J. P., Halverson, R., & Diamond J. B. (2001). Investigating school leadership practice: a distributed perspective. *Educational Researcher, 30*, 23-28.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (2004). Towards a theory of leadership practice: a distributed perspective. *Journal of Curriculum Studies, 36*, 3-34.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*, 387-431.
- Stiefel, L., Schwartz, A. E., & Rubenstein, R. H. (1999). Measuring school efficiency using school level data: Theory and practice. In M. E. Goertz and A. Odden, (Eds.), *School-based financing. Twentieth yearbook of the American Education Finance Association* (pp. 40-75). Thousand Oaks: Corwin.
- Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). Effective schools/accomplished teacher. *The Reading Teacher 53*(2), 156-159.
- Taylor Fitz-Gibbon, C., & Lyons Morris, L. (1987). *How to design a program evaluation* (2nd ed.). Newberry Park, NJ: Sage

- Tsang, M. C. (1997). Cost analysis for improved educational policymaking and evaluation. *Educational Evaluation and Policy Analysis*, 19(4), 318-324.
- U.S. Department of Education. (2002). *No Child Left Behind Act of 2001*, Washington DC: Author, Retrieved February 19, 2005, from <http://www.ed.gov>.
- U.S. Department of Education, Institute of Education Statistics, National Center for Education Statistics. (2002). *The Nation's Report Card: Reading Highlights 2002 National Assessment of Educational Progress (NAEP)*. Washington, DC: Author.
- Verstegen, D. A., & King, R. A. (1998). The relationship between school spending and student achievement: A review and analysis of 35 years of production function research. *Journal of Education Finance*, 23, 243-262.
- Wang, J., & Goldschmidt, P. (1999). Opportunity to learn, language proficiency and immigrant status effects of mathematics achievement. *Journal of Educational Researcher*, 93(2), 101-111.
- Weick, K. E. (1975). Educational organizations as loosely coupled systems. *Administrative Services Quarterly*, 21, 1-19.
- Yin, R. K. (2003). *Case study research* (3rd ed.). Thousand Oaks: Sage.