

The Effect of Explicit Teaching of Comprehension Strategies on Reading Comprehension in
Elementary School

James Grant Atkins

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

© 2013

James Grant Atkins

All rights reserved

ABSTRACT

The Effect of Explicit Teaching of Comprehension Strategies on Reading Comprehension in Elementary School

James Grant Atkins

This study investigated the effect of explicit teaching of comprehension strategies on the comprehension performance of elementary school students. Two schools with similar demographics, including a significant proportion of students at risk for reading failure, participated in the study. One school utilized an explicit comprehension strategy instruction program while the other school utilized traditional comprehension programs. A quasi-experimental mixed methods design was employed; standardized and researcher-developed pretests and posttests were administered and observations were conducted over a two-year period.

Analyses of covariance that treated the pretests as covariates resulted in significant differences between the schools on several measures, but the overall results did not favor either school. Significantly more instances of explicit instruction were demonstrated by teachers using the explicit program than teachers using the traditional programs. Differences in the number of instances of explicit instruction were most substantial during the student practice and feedback sections of the lessons.

Correlations between teacher explicitness scores and student achievement scores and a comparison of the achievement scores of high-explicit and low-explicit teachers (regardless of the school's program) did not demonstrate a significant relationship between explicit instruction and student performance. However, teachers using the explicit program for their second year demonstrated more explicit instruction than teachers using the program for their first year.

Students in the second-year teachers' classrooms demonstrated superior performance on a passage comprehension assessment compared to students in the first-year teachers' classrooms. Also, a comparison of the students who had been exposed to the explicit program for two consecutive years and the students who had been exposed to the traditional program for two consecutive years found that second-grade students exposed to the explicit program outperformed their counterparts on most assessments.

These results did not indicate an overall relationship between teacher explicitness and student performance. The results do indicate, however, that an explicit comprehension program can result in more explicit instruction in the classroom and that instances of explicit instruction increase when teachers have more experience teaching such a program. Moreover, there is evidence that among experienced teachers, there is a positive relationship between teacher explicitness and student performance.

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
I. INTRODUCTION	1
Components of an Effective Reading Program	2
Urban Education Exchange	4
The Present Study	8
General Hypothesis	9
Research Questions	13
II. LITERATURE REVIEW	15
Reading Comprehension Skills and Strategies	15
Approaches to Reading Comprehension Instruction	16
Explicit Instruction of Reading Comprehension Strategies	19
Observations of Reading Comprehension Instruction	26
Mixed Methods Studies of Reading Comprehension	30
III. METHOD	37
Participants	37
Treatment school teachers	38
Comparison school teachers	38
Materials	39
Treatment school curriculum	39
Comparison school curriculum	44
Measures	47
Assessments	47

<u>Chapter</u>	<u>Page</u>
Classroom observation protocol	49
Procedure	42
Professional development	53
Assessments	53
Classroom observations	55
Interviews	59
Scoring of Dependent Measures	60
IV. RESULTS	63
Research Question 1	63
Research Question 2	76
Research Question 3	80
Research Question 4	87
Research Question 5	91
Research Question 6	94
Research Question 7	97
Research Question 8	100
V. DISCUSSION	103
Research Question 1	103
Research Question 2	107
Research Question 3	109
Research Question 4	110
Research Question 5	112

<u>Chapter</u>	<u>Page</u>
Research Question 6	113
Research Question 7	114
Research Question 8	116
Limitations and further analyses	117
Conclusions	120
REFERENCES	123
<u>Appendix</u>	<u>Page</u>
A: Demographics, Measures, and Procedure Tables	129
B: Tables for Research Question 1	137
C: Tables for Research Question 2	187
D: Tables for Research Question 3	190
E: Tables for Research Question 4	208
F: Tables for Research Question 5	214
G: Tables for Research Question 6	220
H: Tables for Research Question 7	223
I: Supplementary Tables for Research Question 7	228
J: Tables for Research Question 8	234
K: Issues with informed consent procedures	237

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. School Demographics	129
2. Assessments Conducted in Year One and Year Two	131
3. Observations Conducted in Year One and Year Two	132
4. Significant differences between the treatment school and the comparison school (ANCOVA results)	137
5. Differences between the treatment school and the comparison school (ANCOVA results)	138
6. GRADE – All Grades – Year 1	139
7. GRADE – All Grades – Year 2	141
8. GRADE – Grades K-2 – Year 1	143
9. GRADE – Grades K-2 – Year 2	145
10. GRADE – Grades 3-5 – Year 1	147
11. GRADE – Grade 3-5 – Year 2	149
12. GRADE – Kindergarten – Year 1	151
13. GRADE – Kindergarten – Year 2	153
14. GRADE – First Grade – Year 1	155
15. GRADE – First Grade – Year 2	157
16. GRADE – Second Grade – Year 1	159
17. GRADE – Second Grade – Year 2	161
18. GRADE – Third Grade – Year 1	163
19. GRADE – Third Grade – Year 2	165
20. GRADE – Fourth Grade – Year 1	167

<u>Table</u>	<u>Page</u>
21. GRADE – Fourth Grade – Year 2	169
22. GRADE – Fifth Grade – Year 1	171
23. GRADE – Fifth Grade – Year 2	173
24. UEE Comprehension Assessment – All Grades – Years 1 and 2	175
25. UEE Comprehension Assessment – Grades 3-5 – Years 1 and 2	177
26. UEE Listening Comprehension Assessment – Second Grade – Years 1 and 2	179
27. UEE Reading Comprehension Assessment – Third Grade – Years 1 and 2	181
28. UEE Reading Comprehension Assessment – Fourth Grade – Years 1 and 2	183
29. UEE Reading Comprehension Assessment – Fifth Grade – Years 1 and 2	185
30. Number of instances of explicit instruction observed per lesson – Years 1 and 2 (All Grades)	187
31. Number of instances of explicit instruction observed per lesson – Years 1 and 2 (Grades K-2)	188
32. Number of instances of explicit instruction observed per lesson – Years 1 and 2 (Grades 3-5)	189
33. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 1: All Cases)	190
34. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 2: All Cases)	191
35. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 1: Treatment School)	192
36. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 2: Treatment School)	193
37. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 1: Comparison School)	194
38. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 2: Comparison School)	195

<u>Table</u>	<u>Page</u>
39. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 1: Grades K-2)	196
40. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 2: Grades K-2)	197
41. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 1: Grades 3-5)	198
42. Correlations between teachers’ explicitness scores and the adjusted posttest GRADE scores of their students (Year 2: Grades 3-5)	199
43. Comparison of High-Explicit vs. Low-Explicit Teachers (split at the median score*) (Year 1)	200
44. Comparison of High-Explicit vs. Low-Explicit Teachers (split at the median score*) (Year 2)	201
45. Comparison of teachers that averaged more than 35 instances of explicit instruction vs. teachers that averaged fewer than 35 instances of explicit instruction (Split at empirical gap in scores*) (Year 2)	203
46. Comparison of teachers that averaged more than 25 instances of explicit instruction vs. teachers that averaged fewer than 25 instances of explicit instruction (Split at empirical gap in scores*) (Year 1)	205
47. Comparison of High-Explicit vs. Low-Explicit Teachers in the Treatment School (Year 1)	206
48. Comparison of High-Explicit vs. Low-Explicit Teachers in the Treatment School (Year 2)	207
49. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Treatment School)	208
50. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Treatment School, Grades K-2)	209
51. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Treatment School, Grades 3-5)	210

<u>Table</u>	<u>Page</u>
52. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Comparison School)	211
53. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Comparison School, Grades K-2)	212
54. Comparison of Teachers using the Curriculum for their First Year vs. Teachers using the Curriculum for their Second Year (Observations – Comparison School, Grades 3-5)	213
55. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Treatment School)	214
56. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Treatment School, Grades K-2)	215
57. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Treatment School, Grades 3-5)	216
58. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Comparison School)	217
59. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Comparison School, Grades K-2)	218
60. Comparison of Students’ Achievement of Teachers using the Curriculum for their First Year vs. Students’ Achievement of Teachers using the Curriculum for their Second Year (Assessments – Comparison School, Grades 3-5)	219
61. Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers (All grades)	220
62. Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers (Students in Grades K-2 in Year 2)	221
63. Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers (Students in Grades 3-5 in Year 2)	222

<u>Table</u>	<u>Page</u>
64. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (All Grades)	223
65. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grades 1 and 2 in Year 2)	224
66. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grades 4 and 5 in Year 2)	225
67. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students in Each Grade	226
68. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grade 2 in Year 2)	227
69. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grade 1 in Year 2)	230
70. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grade 3 in Year 2)	231
71. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grade 4 in Year 2)	232
72. Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Students in Grade 5 in Year 2)	233
73. Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers (All Grades)	234
74. Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers (Grades K-2)	235
75. Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers (Grades 3-5)	236

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Classroom Observation Form	133
2. Frequencies of Mean Instances of Explicit Instruction (All Teachers – Year 2)	202
3. Frequencies of Mean Instances of Explicit Instruction (All Teachers – Year 2)	204

ACKNOWLEDGEMENTS

I owe a huge debt of gratitude to the many people that have helped and supported me throughout the time I have been working on this dissertation, and to all the teachers that have inspired me during my studies of education and psychology.

First of all, I would like to thank my advisor and sponsor, Dr. Joanna Williams, for her continuous support throughout my studies. When I first met Dr. Williams while working at Urban Education Exchange ten years ago, I was inspired by her knowledge and insight regarding reading comprehension. I feel honored to have been a student in her classes, a graduate assistant in her lab, and to have her as a mentor for my work. I would also like to thank Dr. John Black for his continued expertise and feedback from the time I first proposed this study. Dr. Michael Kieffer, Dr. Matthew Johnson, Dr. Mitchell Rabinowitz, and Dr. Linnea Ehri were extremely helpful as members of my committee and the discussions I had with them regarding my study helped me to think deeply about my research and directions for the future.

Additionally, I would like to thank everyone at ReadWorks (formerly Urban Education Exchange) for their inspiration and support during this process. Nancy McDonnell Scharff, Nadia Rawls, Drake Hagner, Jeffrey Fleishhacker, and David Ciulla encouraged me to study the effect of explicit comprehension instruction on reading achievement and provided invaluable support throughout the processes of designing the study, collecting and analyzing data, and reporting the results. I have also appreciated the friendship, support, and collaboration of my colleagues in Dr. Williams' reading lab, Jill Ordynans, Lisa Pao, Rong Cheng, Jenny Kao, and Stephanie Ramsey.

Finally, I would not have been able to complete my dissertation without the love and support of my family and friends, particularly Seth Fleischauer and Kiersten Greene, my friends in education. My mother and father have provided much needed encouragement, motivation, and support throughout my career and Amanda, Andrew, and Chris are always there for me when I need them. Most importantly, I could never express how grateful and thankful I am to my wife, Jane, who has provided unending love, belief, patience, support, motivation, and the ability to always make me happy.

J.G.A.

CHAPTER I

Introduction

Many children fail to comprehend what they read. This failure has been well documented and widely studied by educators and researchers from a variety of perspectives. According to the most recent report of the National Assessment of Educational Progress ([NAEP] 2011), only 34% of 4th graders nationwide are reading at or above a proficient level. This number is not significantly different from the last two NAEP reports, in 2007 and 2009. A vast number of studies, especially over the past 40 years, have investigated how best to improve reading. Instruction in comprehension, however, has not always been the focus of these studies. This is especially true in the primary grades (K-2), where lower-order reading processes, such as decoding and word recognition, are often the focus of not only reading research, but also reading instruction. The lack of attention to comprehension is surprising, given that most educators would agree that comprehension (and not decoding or word reading) is the goal of reading. Lower-order processes such as decoding and word reading were much better understood and, therefore, a more straightforward subject to both teach and study. During the past decade, researchers and educators have begun to make progress in their understanding of reading comprehension. Despite this progress, there is still much to be learned about effective reading comprehension instruction, especially with respect to at-risk students. An understanding of the nature of reading comprehension is necessary to investigate how reading comprehension instruction can be improved for these students. After settling on a definition of reading comprehension, this paper will examine the evolution and current state of reading comprehension instruction.

The RAND Reading Study Group (Snow, 2002) proposed a practical and comprehensive definition of reading comprehension that has since been widely referenced: Reading comprehension is “the process of simultaneously extracting and constructing meaning through interaction and involvement with written language” (p. 11). It consists of three elements: the reader, the text, and the activity or purpose for reading; these elements interrelate within a larger sociocultural context (Snow, 2002). Reading comprehension instructional programs must take into consideration these three elements, as well as the sociocultural context, in order to be effective. Programs that focus on one or two of the elements at the expense of the others are likely to be less effective. Unfortunately, there is a lack of research evaluating instructional programs from each of these perspectives. The RAND Foundation Report revealed that educators face many obstacles when selecting methods of instruction to improve students’ reading comprehension, a result of unsystematic research and development efforts and an inattention to taking evidence-based practices to scale (Snow, 2002). The evaluations that are available to educators typically use methods that only analyze the instructional programs from the perspective of one of these elements (reader, text, or activity) as opposed to a more holistic evaluation.

Components of an Effective Reading Program

The theoretical framework of reading comprehension investigated in this paper emphasizes the interrelationship of lower and higher order skills and the importance of instruction of both, at all grade levels (Williams, 1998; Nation & Norbury, 2005). In reading instruction, teachers typically differentiate between word recognition and decoding (lower order skills) and comprehension (higher order skills). The traditional approach to reading instruction is

to focus on the lower order skills until they have been mastered and only then move on to instruction in the higher order skills. Students that take longer to master the lower order skills will receive instruction in the higher order skills later than their peers. Low-achieving and at-risk students often receive instruction that is heavily focused on lower order skills long after their higher-achieving peers have moved on, leaving them limited opportunity to attain critically important higher order comprehension skills.

A complete reading comprehension instructional program must address not only acquisition but also the extended phases of fluency (performing skills accurately and in an integrated manner; reading with speed and expression), transfer (applying skills to novel situations), and maintenance (demonstrating skills over time). Often, without careful instruction, students find it difficult to generalize what they have learned on trained tasks to fluent performance on tasks beyond those used in training (Gersten, Fuchs, Williams, & Baker, 2001; Best, Rowe, Ozuru, & McNamara, 2005). The goal of any reading instruction for at-risk or low-achieving students should be transfer of the comprehension skills and strategies they have learned to enable them to fluently read and comprehend novel texts.

Also, progress in reading comprehension cannot be made without domain knowledge, and most improvements in reading comprehension are associated with advances in domain knowledge. That is, skills and content are intertwined. The relative deficiency of domain knowledge among at-risk students further compromises their improvement in reading comprehension. The National Reading Panel Report (NRP, 2000) identified vocabulary and comprehension strategies as two of the five core components of reading instruction and emphasized the importance of both readers' knowledge and problem-solving thinking processes. Similarly, Perfetti, Marron, and Foltz (1996) found potential sources of reading comprehension

failure resulting from both knowledge and process deficiencies. In their model, knowledge deficiencies include domain and word knowledge, and process deficiencies include weaknesses in inference making, comprehension monitoring processes, decoding and word naming speed, and working memory. During the process of reading, these components of knowledge and processes interact with one another to yield meaning from the text (Perfetti, Marron, & Foltz, 1996).

An effective program to improve the reading comprehension of at-risk and low-achieving students, therefore, must include instruction in fluency, vocabulary, and domain knowledge, in addition to explicit reading comprehension instruction. The focus of this study is on the explicit reading comprehension instruction portion of a complete reading program.

Urban Education Exchange

ReadWorks¹ (formerly Urban Education Exchange) is a not-for-profit organization committed to eliminating the achievement gap in reading in urban elementary schools. Urban Education Exchange provides urban public and private schools – many of which contain a substantial number of students that qualify for free-and-reduced lunch – with a research-based reading comprehension curriculum. UEE has had over 15 years of experience working with urban children in developing, implementing, and evaluating curricula as well as conducting applied research focused on helping inner-city schools to promote student academic success.

The comprehension program developed by UEE (the UEE Comprehension Program), centers on the teaching of comprehension “concepts” termed the Concepts of Comprehension©. The UEE Concepts of Comprehension© are taught as reading strategies with the goal that they

¹ At the time this study began, ReadWorks was known as Urban Education Exchange (UEE). For simplicity’s sake, I will refer to the organization as Urban Education Exchange (UEE) and the evaluated program as the UEE Comprehension Program, although that has since changed.

become automatic skills. At any grade level, some of the Concepts of Comprehension© may already be automatic and effortless, in which case the students are practicing them as skills. At the same time, they are being introduced to other Concepts of Comprehension© as new strategies. Thus, throughout the program, the Concepts of Comprehension© lessons draw upon skills the students have already mastered and introduce new strategies to improve the comprehension of a text, which themselves, as the lessons proceed, become automatic and effortless skills.

The UEE Concepts of Comprehension© program was developed through a review of the research literature (Pressley 2000; Hirsch 2003; Hart & Risley 2003) and feedback from UEE partner schools whose students demonstrated difficulties in reading comprehension. UEE developed the Concepts of Comprehension© program to address the poor quality of many students' inferential thinking skills. At-risk children often lack explicit modeling and instruction in the inferential thinking skills necessary to become proficient readers who can comprehend well (Block & Pressley 2003, NRP, 2000). The National Reading Panel (NRP, 2000), which analyzed the research literature on beginning reading, concluded that comprehension can be improved by instruction that helps readers utilize specific comprehension strategies. Research has found that explicitly teaching comprehension strategies “makes a difference in learner outcomes, especially for low-achieving students” (Snow, 2002, pg. 33). Based on this research, UEE developed a comprehension program that teaches a series of comprehension skills and strategies directly and explicitly and teaches students to use them in an integrated fashion.

The UEE instructional model contains elements of both direct instruction and strategic instruction. A meta-analysis of approaches to reading comprehension instruction showed a combination of those two models had the largest effect sizes (Swanson, 1999). Successful

instruction is both explicit and intensive. It proceeds systematically and gradually from the simple to the complex, and it provides substantial practice at each step. It promotes purposeful activation and self-regulation of performance. Teachers explain the purpose of each activity and encourage self-reflection and self-evaluation of understanding and performance (Swanson, 1999). The UEE instruction model is sequenced in a highly structured manner, moving gradually to greater complexity; it is scaffolded, and the scaffolding is removed gradually; and there is extensive practice at every step along the way.

At the same time, the UEE reading comprehension program emphasizes the holistic nature of the comprehension process and the importance of integrating text meaning with already-learned concepts and personal experiences, while also underscoring the demonstrated value of structured, direct instruction. Few programs have focused on identifying mechanisms to facilitate comprehension or building an instructional strategy that combines proven components in order to enhance students' reading comprehension (Armbruster et al., 2001).

Correnti and Rowan (2007) identified six factors that promote the implementation of programs designed to foster instructional change in schools. They found that the programs that have the greatest rates of implementation are those programs that: focus on changing specific instructional practices, have clearly defined goals, have written and other supports for teachers, seek significant change to instructional practices, provide an effective professional developer to help teachers, and ensure that the program is implemented effectively (Correnti and Rowan, 2007). The implementation of the UEE curriculum in the treatment school meets all of these criteria.

The UEE comprehension program is designed to address the underlying causes of comprehension difficulties in urban, at-risk students from kindergarten to fifth grade. The fact

that comprehension instruction in the UEE program begins in kindergarten is significant. Too many reading comprehension programs only begin comprehension instruction in the third grade or later (Pearson and Duke, 2002). This intervention begins comprehension instruction as early as possible, before some students are even able to decode, through listening comprehension instruction and activities. The listening comprehension instruction in the primary grades is as explicit and systematic as the reading comprehension instruction for students in later grades.

The UEE comprehension program incorporates instruction of *explicit* comprehension skills and strategies with instruction of domain and word knowledge. This comprehension program, developed with instructional strategies incorporating proven research components, addresses the need of identifying mechanisms to facilitate comprehension and building instructional strategies (Armbruster et al., 2001). Moreover, the intervention is supported by *teacher training* to ensure proper implementation and, within the context of this study, *reading assessments* have been designed to provide information about student strengths and weaknesses in order to more carefully target instruction. These components of the UEE comprehension program – curriculum, teacher training, and assessment – are cited by the RAND Reading Study group as having the highest priority for further research (Snow, 2002). Therefore, the implementation of reading comprehension strategies through UEE's comprehension program and its preliminary assessment will provide valuable feedback on the UEE comprehension program and its potential efficacy. Furthermore, the intervention will address current research needs as it seeks answers to questions in the areas of curriculum, teacher training, and assessment in reading comprehension.

The Present Study

The present study investigates the effect of the explicit teaching of reading comprehension skills and strategies, through the Urban Education Exchange (UEE) Concepts of Comprehension program©, on the reading and listening comprehension of students in grades K-5. The UEE comprehension program was compared to widely-used comprehension programs through observations and assessments at two demographically similar schools with similar student achievement profiles. The population of both schools included a substantial proportion of students at risk for reading failure. The initial comparison of the UEE program to the traditional programs was carried out as an evaluation commissioned by UEE in order to establish the efficacy of the UEE comprehension program (Williams & Atkins, 2011). This study further evaluates the data collected during that evaluation to examine the relationship between teacher explicitness and student achievement.

Children who have difficulty understanding what they read, although they may have adequate word recognition and fluency skills, typically do not respond well to instruction that is not well-structured. Research has demonstrated they respond positively to programs that feature explicit and structured reading comprehension instruction (Duffy et al., 1987; National Reading Panel, 2000; Pressley, 2006). However, at this time, there is no firm evidence that an explicit curriculum in reading comprehension will prove superior to a less explicit curriculum in a typical school setting. This is in contrast to the great deal of evidence that indicates that explicit instruction in decoding and word recognition leads to superior performance (NRP, 2000), not only in research settings but also in a typical school setting. This mixed methods study investigates the relationship between explicit, highly structured instruction in reading comprehension and student outcomes across two curricula through an in-depth analysis of both

tests and observational data. The knowledge gained concerning the role of explicit, structured instruction will inform future instructional design, to result in more potent comprehension strategy instruction in the classroom.

This study also addresses the need for studies that follow the progress of elementary school students for more than one year, especially starting in kindergarten. Studies that do this are especially valuable because they allow researchers and educators to see the effects of early instruction in comprehension strategies (in kindergarten) on later tests of reading comprehension (after first grade). Reading comprehension assessments of kindergarten students are likely to suffer from floor effects, making assessment of these students after an additional year of schooling particularly worthwhile (Slavin et al 2009).

General Hypothesis

The general hypothesis of this study is that the explicit teaching of comprehension skills and strategies, as measured through classroom observations, will result in improved listening and reading comprehension as measured by standardized and researcher-developed reading comprehension tests.

Research Question 1: The overall comparison of the UEE curriculum and the other curricula provides a starting point for this investigation. It is unusual for small-scale studies like this one to show significant differences between instructional programs on standardized assessments (Dole, Duffy, Roehler, & Pearson, 1991; Pressley, 1998; Slavin et al., 2009), and we did not expect that to occur in this case. We asked further questions to examine the effectiveness of explicit instruction.

Research Question 2: We began by asking whether the UEE curriculum is, in fact, more explicit than the other curricula. Gersten, Dimino, and Jayanthi (2007) hypothesized that the

quantity of research-based instructional practices demonstrated by teachers would be a good indicator of *quality* of teaching. These practices included explicit comprehension instruction, but also other, more general classroom practices. The present study focuses more specifically on the explicit instruction of comprehension strategies as assessed by the number of instances of explicit instruction, which was one part of the Gersten, Dimino, and Jayanthi (2007) observation form.

Research Question 3: Despite the evidence (Duffy et al., 1987; National Reading Panel, 2000; Pressley, 2006) that explicit comprehension instructional programs can result in the improved comprehension of students, there is little evidence that it is the explicit component of these programs that is leading to the gains in reading comprehension. That is, there is little evidence of a link between the explicit instruction that is provided by the teachers and the achievement of the students (Chamberlain et. al. 2009; Ness 2011). Even within a given explicit comprehension program, there is likely to be variation in the amount of explicit comprehension instruction that the teachers provide, and it would be expected, following the literature, that those teachers who provide more explicit instruction would have students who perform better on comprehension assessments. Few studies have examined this relationship between explicitness and achievement in a typical classroom setting. This study will examine that relationship, and the conclusions reached will help to understand the link between teacher explicitness and student achievement in greater detail.

Research Question 4: In addition to the consensus (despite the lack of empirical evidence) that explicit teaching leads directly to student gains, there is also a consensus (supported by more empirical evidence) that the positive effects of new curricula are more robust once the school has been using the curriculum for more than one year (Fullan, 2001; Gersten,

Carnine, Zoref, & Cronin, 1986; Hall & Hord, 2001; James-Burdumy et. al., 2012). In order to more fully understand the relationship between teacher explicitness and student achievement, therefore, we must also take the experience using the instructional programs for the teachers and students into account. While the effect of experience on outcomes has been widely studied, few studies have separated the effects of teacher experience from the effects of student experience. This study will investigate the relationship between experience and performance by examining experienced teachers and experienced students separately. First, the study will look at the teachers in both schools. Is level of experience related to teacher explicitness? As the explicit comprehension program utilized by the treatment teachers is specifically designed to increase explicitness, it is expected that the experienced treatment teachers will demonstrate more explicit instruction as compared to the less experienced treatment teachers. The traditional comprehension programs used in the comparison school, however, do not focus on explicit instruction and, therefore, it is not expected that the experienced comparison teachers will demonstrate more explicit instruction as compared to the less experienced comparison teachers.

Research Question 5: As has been established through previous research demonstrating gains in performance in the second year of using an instructional program, it is expected that the students of teachers who have more than one year using a comprehension program will demonstrate improved comprehension as compared to the students of less experienced teachers, in both the treatment and the comparison schools. Thus, Research Question 4 addresses whether experienced teachers are more explicit and Research Question 5 addresses whether experienced teachers have better student outcomes.

In answering these two research questions, this study will indirectly examine the relationship between teacher explicitness, teacher experience, and student achievement. Both the

treatment and comparison teachers will have more experience teaching the program in Year 2 of the study, but only the treatment teachers are expected to demonstrate more explicit instruction in Year 2 as a result of using the explicit program. On the other hand, the students of experienced teachers in both schools are expected to show increased performance in Year 2 of the study. The analysis of student achievement data in both the treatment and comparison school will enable a closer examination of the relationship between teacher explicitness, teacher experience, and student achievement.

Research Question 6: Next, we will examine the relative effect of the explicit versus the traditional program, when all the teachers are experienced, on student outcomes. The majority of the teachers in the comparison school had already been using the traditional comprehension program for several years before the start of the study whereas all the teachers using the explicit program were doing so for the first time in Year 1 of the study. By comparing the outcomes of students in classrooms of experienced teachers using the explicit program versus students in classrooms of experienced teachers using the traditional program, we are able to determine the effect of an explicit comprehension program when only experienced teachers are involved. This comparison allows us to consider the effects of an explicit comprehension program independent of the variable of teacher experience, and it is expected that the students of experienced teachers using the explicit program will outperform the students of experienced teachers using the traditional program.

Research Question 7: It could be, however, that the important factor in the gains found amongst schools using a curriculum for more than one year is that the students, rather than the teachers, have become more experienced with that curriculum. Thus, we examined those

students who were enrolled in the treatment or comparison school (and were therefore exposed to the explicit or to the traditional comprehension program) for two consecutive years.

Research Question 8: Finally, this study examined the effects of the explicit comprehension program versus the traditional program when both teachers and students were experienced in using their program. The comparison of the students enrolled in the treatment school who had teachers who had taught the explicit program in the previous year to the students enrolled in the comparison school who had teachers who had taught the traditional program in the previous year allows us to examine the effects of the programs more closely by equating the experience of both the teachers and the students as much as possible within the scope of our study.

Research Questions

1. Does instruction in reading comprehension that focuses on the explicit teaching of reading comprehension skills and strategies (Treatment: Program T) improve performance on standardized and researcher-developed reading comprehension tests more than a traditional reading comprehension program (Comparison: Program C)? *Answered in the Results section, pages 59-71, and in the tables in Appendix B.*
2. Does Program T result in more instances of explicit instruction during reading comprehension lessons than Program C, as measured by classroom observations? *Answered in the Results section, pages 72-76, and in the tables in Appendix C.*
3. Do the students in classrooms of teachers whose comprehension instruction is more explicit (regardless of the school's program) outperform the students in classrooms of teachers whose comprehension instruction is less explicit? *Answered in the Results section, pages 76-83, and in the tables in Appendix D.*

4. Do teachers who are teaching the same program at the same grade level two years in a row demonstrate more explicit instruction during their second year than teachers who are teaching the program for their first year? *Answered in the Results section, pages 83-87, and in the tables in Appendix E.*

5. Do the students of second-year teachers demonstrate improved performance on comprehension measures as compared to students in the first-year teachers' classrooms? *Answered in the Results section, pages 87-90, and in the tables in Appendix F.*

6. Do the students of second-year teachers in the treatment school demonstrate improved performance on comprehension measures as compared to students of second-year teachers in the comparison school? *Answered in the Results section, pages 90-92, and in the tables in Appendix G.*

7. Do the students who have been enrolled in the treatment school for two consecutive years outperform the students who have been enrolled in the comparison school for two consecutive years? *Answered in the Results section, pages 92-95, and in the tables in Appendices H and I.*

8. Do the students who have been enrolled for two consecutive years and have second year teachers in the treatment school outperform the students who have been enrolled for two consecutive years and have second year teachers in the comparison school? *Answered in the Results section, pages 96-98, and in the tables in Appendix J.*

CHAPTER II

Literature Review

This section will review research on reading comprehension strategy instruction. It will focus on the difference between reading skills and strategies, the explicit teaching of reading comprehension skills and strategies in elementary school, and what is known about instruction, assessments, and observations of reading comprehension.

Reading Comprehension Skills and Strategies

In recent years, particularly since the release of the National Reading Panel report (NICHD, 2000), research on reading comprehension instruction has focused on comprehension skills and strategies. Many programs have been developed that teach reading comprehension strategies individually or together as a set to improve the comprehension of a text. These strategies range from general strategies that good readers use when they have difficulty comprehending, such as rereading, to specific skills that are useful when reading a novel text, such as finding the main idea. It is important to define the terms *reading skill* and *reading strategy* in order to distinguish between them and explain how they are used in the literature and in this particular study.

In instructional programs, as well as in many research studies, the terms *reading skill* and *reading strategy* have been confounded and used interchangeably (Afflerbach, Pearson, & Paris, 2008). The use of these terms has led to confusion for researchers and teachers, and to a diminished understanding of both skills and strategies. Afflerbach, Pearson, and Paris (2008) propose that the difference between reading skills and reading strategies lies in the control the reader exercises over them. Reading strategies require effort and control on the part of the reader, whereas reading skills are effortless and automatic actions. It is not the concept, process,

or even the goal that determines whether a specific concept is a reading skill or a reading strategy. It is the level of mastery and automaticity that the reader has achieved with the concept that determines whether it is a skill or strategy (Williams & Atkins, 2009). In this paper, therefore, whether a technique or concept is referred to as a skill or strategy will depend on whether the reader exerts conscious control over it. We will primarily refer to *reading strategies*, with the assumption that we are focusing on the early stages of instruction, before the students have mastered and internalized the strategies into skills. The rate at which the instructed strategies become internalized skills will depend on the teacher, student, the reading task, and the particular skill or strategy. The goal is that, through sufficient modeling, guided practice, and, finally, independent practice, these strategies will become fluent skills for all the students.

Approaches to Reading Comprehension Strategy Instruction

There has recently been an increased focus on reading comprehension in elementary schools; this has manifested particularly as instruction in reading comprehension strategies. Instruction in reading comprehension strategies can involve general instruction on thinking strategically when encountering comprehension difficulties or explicit instruction on specific comprehension strategies. While both of these approaches to reading comprehension strategy instruction have proven effective to some extent, explicit instruction in comprehension strategies may be more effective in helping students transfer the instructional practices to novel texts. Teaching students strategies, through modeling and guided practice, that they can use when they encounter difficulties while reading any text allows them to eventually use these strategies independently (Liang & Dole, 2006).

The earliest forms of instruction in reading comprehension strategies may have been influenced by reading comprehension assessments (Pearson, 2009). These assessments tested

students' ability to perform skills like finding the main idea of a text, understanding the author's purpose, finding specific information in the text, and drawing inferences based on the text. It is unclear whether these assessments tested skills that were already being taught in schools, or whether instruction in these skills was a response to the assessments. Nevertheless, testing and instruction in reading comprehension, as in most fields, are closely linked. Instructional programs soon began to organize the strategies into a sequence that progressed from literal to inferential, but for the most part they only provided a framework for instruction and lacked any sort of explicit instruction in the strategies (Pearson, 2009).

In the first generation of research on comprehension strategy instruction, single strategies that could be used before, during, and after reading, such as question-formation, imaging, and summarizing, were taught in isolation. These studies typically consisted of teaching a treatment group the target comprehension strategy, while withholding instruction from a comparison group. In the majority of these studies, the strategies were taught in isolation and there was little attempt to help students integrate the various strategies while processing authentic text. Interventions of this type generally resulted in findings that favored the treatment group on researcher-developed assessments of use of the instructed strategy (Pressley, 1998). Students were able to perform the strategy in isolation, as they had been taught, but they rarely utilized the strategy on their own when reading connected text. Rarely did instruction of this type transfer to other strategies or result in improvement on more general assessments of reading comprehension (Dole, Duffy, Roehler, & Pearson, 1991). Interventions that teach students why and when to use individual strategies (instead of just how to use them) are generally more successful in attaining transfer and maintenance effects on comprehension assessments. Jitendra, Hoppes, and Xin (2000) integrated direct instruction on main idea strategy training and self-monitoring training;

this treatment proved successful on posttest measures as compared to a control group that received general reading instruction. The effects of treatment were maintained over six weeks, although transfer effects were less consistent.

Studies of skilled readers in the 1980's revealed that they typically employ a number of comprehension skills and strategies simultaneously while reading. These findings led to a new wave of comprehension strategy instruction in which multiple strategies were taught to students (Pressley, 1998). Researchers and educators found that it is important to teach each strategy and also teach ways to incorporate all relevant strategies simultaneously during reading. Models that do this, including Palincsar and Brown's (1984) Reciprocal Teaching, which incorporates four strategies – asking questions, summarizing, predicting, and clarifying confusing content – have been demonstrated to be effective (NRP, 2000). In Reciprocal Teaching, teachers demonstrate successful use of these four strategies in small groups before gradually releasing responsibility and allowing the reading groups to practice the strategies on their own. In each reading group, a student is designated as the “teacher” and is in charge of summarizing and posing questions about the reading. The actual teacher scaffolds this role for the student teacher, providing prompts and instruction when needed, and explicitly reminds the students that they should use these strategies when reading on their own. Palincsar and Brown (1984) found that this approach to strategy instruction can result in improvements in reading comprehension for students who are adequate decoders but struggle with comprehension, including on standardized assessments. A meta-analysis conducted by Rosenshine and Meister (1996) found that instructional interventions that teach students to generate questions are successful in improving reading comprehension, resulting in an effect size of .36 when evaluated using standardized assessments and an effect size of .86 when evaluated using researcher-developed assessments.

Studies of this type, when successful, have typically involved strong teacher modeling, consistent monitoring of strategy use, and substantial treatment duration. Overall, they demonstrate the importance of a direct, explicit approach for low-achieving and at-risk students. It is important to note this, because the approach toward comprehension instruction in general education has changed over the past few years. The research supporting instruction in reading comprehension has led to a proliferation of programs that include these strategies, but rarely provides the direct, explicit instruction that many students need. “The general rule is, teach children many strategies, teach them early, reteach them often, and connect assessment with reteaching” (Afflerbach, Pearson, & Paris, 2008). Unfortunately, while this rule may be successful in teaching a majority of students how to comprehend what they read, it is not enough for those at risk for reading failure.

Explicit Instruction of Reading Comprehension Strategies

The report of the National Reading Panel (NRP, 2000) indicated that explicit instruction can improve reading comprehension, because it teaches students to use specific cognitive strategies when they encounter barriers to understanding what they are reading. While readers acquire these strategies informally, to some extent, it is the explicit instruction of comprehension strategies that has been shown to be highly effective in cultivating understanding (NRP, 2000). Such explicit instruction is particularly important for students at risk for academic failure. While good readers can develop an understanding of comprehension strategies without explicit instruction, struggling and at risk readers often require teachers to explicitly teach them how, when, and why to use the strategies. In this explicit instructional model, teachers demonstrate strategies for students and guide practice until the students are able to carry out the strategies independently. Effective explicit instruction includes a direct explanation of the strategy,

modeling the strategy for students, guided practice in using the strategy, and, finally, independent application of the strategy (Armbruster et. al., 2001).

One of the first approaches to this type of strategy instruction was the Question-Answer-Relationship (Q-A-R) technique (Raphael & McKinney, 1983). This technique is based on the premise that question asking and answering can enhance the comprehension of a text and that there are a variety of questions that good readers ask and answer before, during, and after the reading of a text. Readers can be taught strategies to successfully answer these different types of questions. The idea of Q-A-R is to explicitly instruct students in these strategies and then give them practice applying the strategies to the specific question types. The questions are grouped according to where the reader must obtain the information that is needed to answer the question: from a single sentence in the text, from multiple sentences in the text, or from the reader's background knowledge. The questions are labeled as "right there," "think and search," or "on my own," and the students are taught to activate the strategies that are appropriate for each type of question.

A quasi-experimental study evaluating Q-A-R was conducted in 10 fourth grade classrooms (Raphael, Wonnacott, & Pearson, 1983). The researchers found that most teachers were sufficiently prepared to teach the program after an inservice training. They also found that the instruction provided to the students was effective in improving their reading comprehension. On a transfer task in which the students answered questions about a science passage, without being cued to use the Q-A-R strategies, the treatment group outperformed the comparison group, which had received their regular reading instruction. The researchers also found that this type of instruction especially benefited lower-ability students (Raphael, Wonnacott, & Pearson, 1983).

Another approach to explicit reading comprehension instruction was investigated with the development of Direct Explanation by Duffy, Roehler, and colleagues (1986). The focus of Direct Explanation is training teachers to teach students to treat reading comprehension as a problem solving process that requires strategic thinking to solve comprehension difficulties. First, teachers introduce the text for the class to read. Next, before reading, the teacher introduces the strategy for the class to learn, including declarative (e.g., what the strategy is), conditional (e.g., when to use the strategy), and procedural (e.g., how to use the strategy) knowledge regarding the strategy. Rather than simply introducing single or multiple comprehension strategies to students, teachers explain the reasoning behind the use of strategies and how to think strategically. Third, the teachers model using the strategy by “thinking aloud” about when, how, and why he or she would use the strategy while reading. Then, the class practices using the strategy with the teachers gradually removing scaffolding until the students are able to use the strategy on their own. Once the students are comfortable with the strategy, they read the text independently and practice using the strategy, while also reading the text for content. At the end of each lesson, the teacher reviews the strategy, including the declarative, conditional, and procedural knowledge that is necessary to use it appropriately. Teachers are trained to reframe the skills taught in basal textbooks as strategies; that is, the skills are not taught as simple steps applied without thought, but as processes applied flexibly and thoughtfully when comprehension difficulties are encountered.

In an experimental study, Duffy, Roehler, and colleagues (1986) found that, although fifth-grade students that received Direct Explanation instruction were more aware of strategies, more extensive instruction was needed to improve scores on standardized tests. In a subsequent study, the researchers evaluated the efficacy of Direct Explanation instruction for third grade

students over an academic year (Duffy et al., 1987). The students whose teachers received training in Direct Explanation of comprehension strategies outperformed a comparison group whose teachers received training in classroom management on a standardized reading comprehension assessment, and more educators and researchers began to focus on the potential of explicit instruction of comprehension strategies.

The National Reading Panel report and subsequent legislation, specifically Reading First legislation, resulted in an increased focus on the explicit instruction of reading comprehension strategies (Moss et al, 2008). Basal textbooks began to include comprehension strategies in order to conform to the new legislation and the focus on scientific, research-based instruction. This has led to a number of textbook companies including “research-based” comprehension strategies such as predicting, questioning, and summarizing in their comprehension instruction. Unfortunately, while there is research that supports the efficacy of teaching these strategies, this research does not necessarily make the basal textbooks programs “research-based,” as they claim. The inclusion of proven strategies into a program does not necessarily make the entire program effective; there are many other aspects of a program that will determine what effects it has on students’ comprehension (Liang & Dole, 2006, Slavin et al, 2009). While certain types of instruction in these strategies can improve reading comprehension, there is a lack of evidence that the particular ways these curricula and programs teach the comprehension strategies are effective. The majority of these basal textbooks and programs lack rigorous research analyzing their programs as a whole. Many of the textbook and program manufacturers are hesitant to subject their programs to a rigorous evaluation, resulting in few comprehensive evaluations of this type. Furthermore, while it is encouraging that there has been an increase in the presence of comprehension skills and strategies in reading textbooks and programs, the type of instruction

included in these programs is generally not consistent with the findings of the NRP or other recent studies (Dewitz, Jones, and Leahy, 2009).

Research on the impact of the Reading First legislation has demonstrated that teachers in Reading First schools are more likely to use research-based instruction, particularly in phonics and phonemic awareness instruction. The teachers in these schools are also more likely to spend more time on fluency, vocabulary, and comprehension instruction (Moss et al, 2008). The effects of this increased focus on the elements of reading deemed to be important for reading growth by the NRP, however, has not resulted in improved reading comprehension for students at Reading First schools (Gamse et al, 2008). One possible reason for this lack of reading improvement despite what experts might call “better” instruction could be the teachers’ reliance on basal textbooks. The majority of the teachers in Reading First schools relied heavily on Harcourt Trophies, Open Court Reading, Scott Foresman Reading, or Houghton Mifflin’s Nation’s Choice (Moss et al, 2008). As discussed earlier, although these basal textbook series include “research-based” instructional and reading strategies, the programs themselves have not been rigorously evaluated (Slavin et al, 2009).

Dewitz, Jones, and Leahy (2009) analyzed five of the best-selling and most popular basal textbook series to determine how they presented reading comprehension strategies instruction. They found that, while the textbook series included instruction in reading comprehension strategies, the instructional strategies that were included did not conform to the recommendations of the NRP or other reading research. There were generally more skills and strategies included than are recommended in the literature, and instruction in these skills and strategies was not sufficiently rigorous. For example, the instruction did not include enough examples of procedural knowledge, instructing students about why and when to use the comprehension

strategies, or metacognitive instruction, which would enable the students to control their use of the strategies. Instruction in the reading comprehension strategies did not follow a gradual release of responsibility, with sufficient scaffolding for the students to learn and then practice the comprehension strategies on their own. The authors concluded that these textbook series would require a great amount of supplemental instruction in the form of explicit modeling of the comprehension strategies, guided and then independent practice in using the strategies, and an assessment system to ensure that the students have learned the instructed strategies in order to meet the recommendations of the literature regarding explicit comprehension instruction (Dewey, Jones, & Leahy, 2009).

Slavin and colleagues (2009) conducted an extensive meta-analysis evaluating the effectiveness of reading programs for students in grades K-5. They reviewed studies that evaluated programs that teachers and administrators might consider to improve their students' reading achievement: reading curricula, instructional technology, instructional process programs, and approaches that combine curricula and instructional process programs. Literature searches resulted in a number of studies that evaluated elementary reading programs by comparing students to a control group using random assignment or matching with adjustments to pretest differences. The authors included studies that evaluated students using standardized or other comprehensive measures of reading, lasted a minimum of 12 weeks, and included a sample size of at least 15 students and 2 teachers. In the investigation of beginning reading programs (grades K-1), the researchers found a sample-size weighted mean effect size of +0.12 for seven reading curricula, +0.09 for 13 studies of instructional technology, +0.37 for 17 studies of instructional process programs, and +0.29 for 23 studies of combined curriculum and instructional process (all of the latter studies were evaluations of *Success for All*). For the beginning grades (K-1), they

found, not surprisingly, stronger effects on measures of decoding than for comprehension or overall reading achievement (Slavin et al., 2009).

Slavin and colleagues' evaluation of upper elementary programs (grades 2-5) resulted in a sample-size-weighted mean effect size of +0.06 for 15 studies of reading curricula, +0.06 for 31 studies of instructional technology, and +0.21 for 33 studies of instructional process programs. Instructional process programs included a subcategory of strategy instruction programs, which are programs that teach students what the authors called "cognitive and metacognitive skills such as summarization, graphic organizers, and prediction." A meta-analysis of five studies evaluating these programs resulted in a weighted mean effect size of +0.32. An analysis of the high and low-poverty schools in the sample revealed that, among studies of instructional process programs in both lower and upper elementary grades, the weighted mean effect size was +0.27 (from 45 studies) in high-poverty schools and +0.20 (from 31 studies) in low-poverty schools. This analysis suggest that instructional in comprehension strategies may be even more important when working with at-risk students. The results of the meta-analysis also indicated that those programs that provide "extensive professional development to teachers in specific classroom strategies" were more effective, especially in high-poverty schools, than programs that provide less professional development to teachers (Slavin et al., 2009).

These studies, along with many similar studies, have demonstrated that instructional programs that include explicit instruction in reading comprehension skills and strategies are effective in improving the reading comprehension performance of students, including at-risk students. The majority of these studies, however, are isolated experimental studies in which the performance of a treatment group is compared to the performance of a comparison group on

comprehension assessments. Many of the studies either utilized researchers as instructors and/or removed the students from their normal classroom environments in order to implement the intervention. Few of the studies made an attempt to situate the research within the specific classroom and school context in which instruction typically takes place. The next section will discuss studies that have attempted to explain and understand this classroom and school context in greater detail.

Observational Studies of Reading Comprehension Instruction

Observations of reading comprehension instruction are instrumental in helping researchers understand what really happens in the classroom, and how recommendations made for instruction can be and are implemented. These observations can also help researchers understand why certain practices and interventions that have proven successful in laboratory experiments may not necessarily work in the classroom. Methods of observation range from qualitative, narrative descriptions of instruction to quantitative coding of instructional acts. Some are based on the actions of the teacher, others on student engagement, some on the method and materials used for instruction, and others a combination of all these factors. The most common types of observational studies involve determining the amount of classroom time spent on reading comprehension instruction. By this measure alone, classrooms lag far behind the recommendations of research scientists.

The seminal observational study of reading comprehension was conducted by Durkin (1978/1979). Durkin found that almost no reading comprehension instruction occurs in reading or social studies classes in grades 3 through 6. In the classes Durkin observed, the teachers generally followed a pattern of mentioning a comprehension skill, giving the students time to practice the skill, and then assessing the students' mastery of the skill. In fact, she found that

very little *instruction* of any kind occurred during these classes. Although our understanding of reading comprehension and how best to teach it has changed considerably since Durkin's study, it is less clear how much actual instruction has changed since that time. In a similar observational study conducted 20 years after Durkin's study, Pressley and his colleagues (1998) found that little had changed in the amount of time that classroom teachers spent on reading comprehension instruction.

A more encouraging recent study was conducted by Ness (2011). In this study, she replicated the practices of Durkin and others to investigate the amount of reading comprehension instruction occurring in grades one through five. Ness (2011) observed twenty teachers in two schools (ten teachers in each school), for a total of 3,000 minutes over seven months. She observed two teachers per grade level per school three times each over the course of the school year. Overall, she found that teachers spent 25% (751 minutes) of their language arts time conducting reading comprehension instruction. Interestingly, the least amount of reading comprehension instruction occurred in third grade (67 minutes), while the greatest amount occurred in fourth grade (287 minutes). The author suggests that this discrepancy may be the result of third grade teachers focusing on test preparation, as third grade is the first year for high-stakes standardized testing in elementary school. The observations revealed instruction in the following reading comprehension strategies: asking questions (256 total minutes), predicting/prior knowledge (184 total minutes), summarization (101 total minutes), vocabulary instruction (85 total minutes), text structure (65 minutes), visual representation (34 total minutes), comprehension monitoring (19 total minutes), and question generation (7 total minutes). Although the data obtained in this study are an important component to reading comprehension research, the author points out that it would be useful to link the data to student

achievement data to determine whether the increased time spent in reading comprehension instruction results in student gains on reading comprehension assessments (Ness, 2011).

Other recent studies have used more pointed observational protocols to determine what teaching methods are used in reading instruction and how students respond to the instruction. Many of these observational protocols include time-sampling procedures in which the observer observes the classroom for a specified amount of time and then completes a checklist indicating components of the lesson, such as the focus of the activity, the teacher's and students' actions, and the materials being used (Gersten et al., 2007; McCutchen et al., 2002; Vellutino & Scanlon, 1996). These observational studies can help us understand what is really happening in the classroom and how teachers can better allocate their time for effective reading comprehension instruction. McCutchen and colleagues (2002) used a time-sampling procedure in which they completed a checklist of questions every 60 seconds. These questions addressed four categories: knowledge affordance, literacy activity, textual context, and group context. Each observation lasted 15 minutes, after which the total times for the codes within each category were calculated.

This time-sampling procedure was first used by Scanlon and Vellutino (1996) in their observations of kindergarten language arts instruction. In this study, the researchers observed the teacher for a short period (ten or twenty seconds) and then completed a checklist which consisted of a series of questions concerning the instructional group with which the teacher is interacting (individual child, pairs, small group, whole class); general focus of the activity (direct reading, indirect reading, writing, management, preparation); teacher's purpose (reading aloud, modeling, asking/answering questions, guiding practice, providing feedback); the materials being used; the focus of the instruction; and the students' expected response. The observers in this

study answered these questions approximately every 30 seconds. They then calculated the percentage of time that the teacher allocated to the various activities within each classroom.

In their observations of two cohorts of kindergarten classrooms, Scanlon and Vellutino (1996) found that nearly half of the instructional time was spent on classroom management, or noninstructional activities. They found that a distressingly small amount of time was spent on direct instruction of pre-reading skills (e.g., teaching letter names, letter sounds, phoneme awareness, decoding, etc.). Teachers often worked with individuals or small groups, leaving the rest of the class to work independently and without direct instruction for large chunks of time. The researchers split the students into two groups based on their performance on a letter identification task in kindergarten. They found that, for the low letter-identification group, a greater amount of time spent on instruction in phoneme awareness in kindergarten was associated with better performance on a first grade reading assessment. There was no such association with the high letter-identification group. The observations also revealed that children whose kindergarten teachers spent less time on direct instruction of word identification skills did not perform as well as those students whose kindergarten teachers spent more time on word identification. A negative relationship was found between time spent on recitation activities in kindergarten and performance on the first-grade reading assessment for the low letter-identification group, suggesting that this type of activity is less important for these students (Scanlon and Vellutino, 1996).

The time sampling procedure used by Gersten, Dimino, and Jayanthi (2007) was somewhat different from the techniques used by Scanlon and Vellutino (1996) and McCutchen and colleagues (2002). Rather than completing a checklist of questions every 30-60 seconds, Gersten, Dimino, and Jayanthi (2007) developed a Reading Comprehension and Vocabulary

(RCV) Observation Measure in which they used a fresh observation form every 15 minutes during a 90-minute observation, for a total of 6 intervals. This procedure allowed them to gather more detailed data during the 15-minute periods than would be possible in a 1-minute period. As part of this time-sampling procedure, instances of research-based instructional practices that occurred during the 15-minute interval were observed and catalogued. The quantity of research-based instructional practices was used as a proxy for quality of teaching. These practices included explicit comprehension instruction, but also other, more general classroom practices such as scaffolded instruction, allowing students sufficient time to answer questions, asking students to elaborate on their responses, encouraging broad participation from the students, maximizing instructional time, and effectively managing classroom behavior. The observation form used by the researchers also included global ratings of comprehension and vocabulary instruction and student engagement. There are three yes/no questions and five Likert scale questions that address the overall quality of comprehension and vocabulary instruction and the management practices of the teacher (Gersten, Dimino, & Jayanthi, 2007).

The observational studies reviewed help researchers understand the types of instruction that are occurring in classrooms. They could also allow researchers to link these practices to student performance. The next section discusses how mixed methods studies of reading comprehension can be effective in making that link more explicit and meaningful.

Mixed Methods Studies of Reading Comprehension

Studies that integrate qualitative data, such as observations, and quantitative data, such as assessment scores, are said to be taking a mixed methods approach to research. Researchers who favor mixed methods have argued that this paradigm integrates the strongest aspects of quantitative and qualitative research while minimizing the weaknesses (Johnson &

Onwuegbuzie, 2004). In the past 15 years or so, there has been an increase in research that combines a quantitative experimental or quasi-experimental design with a qualitative observation or survey design to study educational issues. Many of these studies seek to understand the differences in treatment groups by using descriptive data to explain these differences.

Mixed methods studies may come closest to addressing all four of Schwab's (1973) commonplaces: teacher, student, task, and setting. The combination of quantitative and qualitative methods allows the researcher to give equal weight to each of these commonplaces when designing the study, collecting data, analyzing the data, and drawing conclusions. Using mixed methods allows the researcher to use both numbers and words to describe the teachers, students, tasks, and settings in which the study takes place. The results of this type of research are especially valuable to practitioners because they allow educators to understand and analyze the complexities of 21st century classrooms better than quantitative or qualitative methods alone (Calfee and Sperling, 2010).

The use of both quantitative and qualitative methods enhances the validity of the research project by including multiple perspectives (Calfee and Sperling, 2010). Mixed methods studies are able to not only identify effective instructional approaches, but also identify why, with whom, and under what circumstances they are effective. The diversity of today's elementary school classrooms demand that researchers pay attention to the differences between students and how instruction effects both individual students and groups of students within the classroom. This requires an understanding and explanation of instructional approaches beyond what either quantitative or qualitative methods can offer alone. Particularly when working with struggling and at-risk students, it is important to conduct research that examines students' difficulties from different perspectives and using different methods. By approaching the problem from only one

perspective, or using only one method, the researchers will obtain only a partial understanding of the problem. In using both numbers and stories, mixed methods studies allow researchers to examine and hopefully understand all the complexities involved in students' reading achievement (Calfee and Sperling, 2010).

Educational effectiveness research has usually focused on large-scale quantitative studies, while school improvement and teacher development studies usually rely on more qualitative measures (Sammons, 2010). Research would benefit from combining these approaches to get the large-scale conclusions that quantitative research can provide with the context and understanding that qualitative research can provide.

As in other research paradigms, mixed methods research in the early elementary grades has lagged behind research in older grades. A mixed methods evaluation of The Reading Edge, a component of the Success For All school reform program, was conducted in two high-poverty middle schools (Chamberlain et. al., 2009). The researchers randomly assigned the sixth grade teachers in both schools to the treatment or control condition. The teachers in the treatment condition received the Reading Edge program, while the teachers in the control condition received books and materials appropriate for the reading program that those teachers had been using previously. The students were pretested in the first month of the school year using the Scholastic Reading Inventory and posttested the following spring using the Gates-MacGinitie Reading Test. Observations were conducted once in the fall and once in the spring, using an observation protocol that coded for levels of metacognitive strategy use, cooperative learning, goal setting and feedback, and classroom management/student engagement (Chamberlain et al., 2009).

The posttest resulted in a significant difference between the treatment and control groups on the Vocabulary subtest of the Gates-MacGinitie, and a difference that approached significance on the Total score of the Gates-MacGinitie (both favoring the treatment group). The observations also revealed differences between the treatment and comparison classrooms. Specifically, the researchers observed more than twice as much metacognitive strategy use, as well as more cooperative learning and goal setting and feedback, in the treatment classrooms. There were no differences between the treatment and comparison classrooms on the ratings of classroom management/student engagement. Despite being rated higher in the strategic areas that are a focus of the treatment program, the researchers felt that the implementation of these strategies was “mechanical” and “static” (Chamberlain et al., 2009). The researchers speculated that, as the teachers become more familiar with the program, they will likely improve their instruction. Unfortunately, the researchers did not examine the observational and assessment data within each condition in order to determine if those teachers that were better implementers had students with higher scores on the assessments. They acknowledge that this type of analysis would help them understand “how and under what conditions” the program is effective, and identify it as an area for future research (Chamberlain et al., 2009).

Gersten and colleagues (2010) conducted a mixed methods evaluation of the Teacher Study Group (TSG), a professional development program designed to help teachers translate research into practice. This study was conducted in 19 schools, with 81 first grade teachers and 468 first grade students. The schools were randomly assigned to either the TSG condition or a control condition, in which the teachers participated in their normal professional development programs as mandated by Reading First. The students were pretested using the DIBELS and the Word Attack and Letter-Word Identification subtests of the Woodcock-Johnson, and posttested

using the Oral Vocabulary, Reading Vocabulary, and Passage Comprehension subtests of the Woodcock-Johnson (Gersten et. al, 2010).

The researchers utilized the Reading Comprehension and Vocabulary (RCV) Observation Measure they had developed (Gersten, Dimino, & Jayanthi, 2007), and that was described earlier in this paper, to evaluate the quality of instruction in both the TSG and the control classrooms. They divided the results of the observations into two scales – comprehension and vocabulary – and summed the tallies for research-based instruction practices observed and scores for questions that evaluated comprehension and vocabulary instruction overall in order to obtain a total score for each scale. The researchers conducted one observation of each classroom during April and May in order to give the teachers time to implement what they had learned during the Teacher Study Groups (Gersten et. al, 2010).

Gersten and colleagues (2010) found that the TSG teachers scored significantly higher than the control teachers on both the comprehension and vocabulary scales of the RCV Observation Measure. The students of the TSG teachers outperformed the students of the control teachers on the Oral Vocabulary subtest. The researchers correlated the scores on the comprehension and vocabulary scales of the RCV Observation Measure and covariate-adjusted student posttest scores and found significant correlations between the comprehension and vocabulary observation scales and the Oral Reading Fluency, Letter-Word Identification, Word Attack, and Reading Vocabulary subtests. Scores on the vocabulary, but not the comprehension, scale of the RCV Observation Measure were also significantly correlated with scores on the Passage Comprehension subtest. The researchers conclude that the significant correlations found between teaching practices and student outcomes suggest that there is a link that future research should continue to investigate (Gersten et. al, 2010).

Another recent mixed methods study conducted by James-Burdumy and colleagues (2010) investigated the effects of four supplemental reading comprehension programs (Reading for Knowledge, Project CRISS, ReadAbout, and Read for Real) on the reading comprehension of fifth graders. This study was conducted over two years, and the researchers collected data from two cohorts of fifth graders. The first cohort included 89 schools, 268 teachers, and 6,349 fifth-grade students. The second cohort included 61 schools, 182 teachers, and 4,142 fifth-grade students. The schools were randomly assigned to one of five conditions: one for each of the four supplemental reading comprehension programs and a control condition that continued to teach reading comprehension as they had in the past (James-Burdumy et. al., 2010).

The students in the study were pretested using the Passage Comprehension subtest of the Group Reading Assessment and Diagnostic Evaluation (GRADE) and the Test of Silent Contextual Reading Fluency (TOSCRF). The students were posttested using the same subtest of the GRADE and a different assessment designed to measure comprehension of social studies and science texts. In addition to the assessments, the researchers conducted two types of observations within the classrooms: fidelity of implementation observations within the treatment classrooms and Expository Reading Comprehension (ERC) observations. The fidelity of implementation observations revealed that a majority of teachers used their assigned program and that a majority also implemented the practices that were deemed important by the respective program developers, although the percentage of practices implemented varied by program. The ERC observations were conducted in both the treatment and control classrooms and recorded the number of research-based instructional practices that the teachers demonstrated during comprehension and vocabulary instruction. The researchers divided these practices into three scales: traditional interaction, reading strategy guidance, and classroom management and student

engagement (James-Burdumy et. al., 2010). The techniques used and the categories of research-based instructional practices observed were very similar to those developed by Gersten, Dimino, and Jayanthi (2007).

Overall, the researchers found that the students in the classrooms of teachers who had taught the ReadAbout during year 1 of the study significantly outperformed the other students on the social studies comprehension test in year 2 of the study. The students who received Reading for Knowledge performed significantly worse than the other classrooms on the science comprehension test in year 1 of the study. There were no other differences between the groups. The researchers did not find any differences between any of the treatment groups or the control group on the reading strategy guidance or the classroom management and student engagement scales of the ERC in year 1 or year 2. The treatment groups as a whole scored significantly lower than the control group on the traditional interaction scale in year 1; the Project CRISS group scored significantly lower than the control group on the traditional interaction scale in year 2. In addition, the researchers correlated the results of the ERC observations and the student posttest scores and found that there was a significant and positive relationship between the posttest scores and the classroom management and student engagement and the reading strategy guidance scales. Although the classroom management and student engagement scale had more consistently positive and significant correlations with the posttest scores, the fact that there is also a correlation with the reading strategy guidance scale suggests that the connection between strategy instruction and student achievement should be investigated further (James-Burdumy et. al., 2010).

CHAPTER III

Method

Participants

Two schools in Queens, New York City, volunteered to participate in the study. One school (the treatment school) had chosen to use the UEE comprehension program in all grades, and in the other school (the comparison school), different reading programs were used. The treatment school served a total of 449 students during the 2008-2009 school year and 438 students during the 2009-2010 school year in grades K-5. The comparison school served a total of 598 students in grades Pre-K-6 during the 2008-2009 school year and 581 students in grades Pre-K-5 during 2009-2010; including 475 (2008-2009) and 527 (2009-2010) in grades K-5. Additional demographic data can be found in Appendix A, Table 1.

Both schools had received Title I funding for at least the school year prior to this study (2007-2008). They both received a grade of "A" according to the Children First Accountability Summary and had made adequate yearly progress in English Language Arts, Math, and Science (New York City Department of Education, 2010). The state-wide tests that were conducted in the Spring of 2008 (the semester before the beginning of the study) revealed that the students in the two schools had very similar achievement profiles. In the third grade, the mean score of the treatment school was 656 whereas the mean score of the comparison school was 654 (the state-wide mean score was 668.87 with a standard deviation of 39.50). In the fourth grade, the mean score of the treatment school was 654 whereas the mean score of the comparison school was 630 (the state-wide mean score was 666.19, with a standard deviation of 40.18). In the fifth grade, the mean score of the treatment school was 656 whereas the mean score of the comparison school was 652 (the state-wide mean score was 667.22, with a standard deviation of 31.02).

All students who were enrolled in general education classes and collaborative team teaching classes participated in this study, including the special education students who were enrolled in these classes; students in self-contained special education classes were not included.

Treatment school teachers.

In Year 1, the treatment school had a total of 24 teachers in 20 classrooms. One teacher in first grade and one teacher in third grade were replaced by new teachers during the spring semester. In addition, one classroom in fourth grade and one classroom in fifth grade were Collaborative Team Teaching (CTT) classrooms and had two teachers each. There were four classrooms in kindergarten (one self-contained special education class), three classrooms in first grade, three classrooms in second grade, four classrooms in third grade (one self-contained special education class), three classrooms in fourth grade, and three classrooms in fifth grade. There was one teacher in the fourth grade CTT classroom who was teaching for her first year.

In Year 2, the treatment school had a total of 20 teachers in 18 classrooms. One teacher in second grade and one teacher in fourth grade were replaced by a new teacher during the spring semester. There were three classrooms in kindergarten, three classrooms in first grade, three classrooms in second grade, three classrooms in third grade, three classrooms in fourth grade, and three classrooms in fifth grade. There were no teachers who were teaching for their first year.

Comparison school teachers.

In Year 1, the comparison school had a total of 22 teachers in 21 classrooms in grades K-5. One teacher in fourth grade was replaced by a new teacher during the spring semester. There were four classrooms in kindergarten (one self-contained special education class), four classrooms in first grade, three classrooms in second grade, four classrooms in third grade (one

self-contained special education class), three classrooms in fourth grade, and three classrooms in fifth grade (one self-contained special education class). There were no teachers who were teaching for their first year.

In Year 2, the comparison school had a total of 20 teachers in 19 classrooms. One classroom in kindergarten was a Collaborative Team Teaching (CTT) classroom and co-taught by two teachers. There were four classrooms in kindergarten, three classrooms in first grade, three classrooms in second grade, three classrooms in third grade, three classrooms in fourth grade, and three classrooms in fifth grade. There were no teachers who were teaching for their first year.

Materials

The treatment school partnered with Urban Education Exchange to implement the UEE Concepts of Comprehension© program in grades K-5 in 2008-2009 and 2009-2010. The comprehension program consisted of the implementation of the curricula; teacher attendance during the summer trainings and grade-level teacher training sessions; and the use of assessments to inform teaching practice. A demographically similar school served as the comparison school and implemented the Mondo Publishing comprehension program (Mondo Publishing, 1997) in grades K-3 and the Teachers College Reading and Writing program (Calkins, 2000) in grades 4-5 (both commercially-available and widely-used curricula).

Treatment school curriculum.

Before the study. Prior to the beginning of this study, the treatment school had used the Teachers College Reading and Writing workshop for five years, beginning in the 2003-2004 school year. The school received two days of teacher training at the beginning of the program, but professional development was limited after that point. In order to make the curriculum more

explicit, the treatment school had created an explicit scope and sequence that included a comprehension “Skill of the Week” and a writing “Genre of the Month.” These skills and genres were consistently implemented in all classrooms across the school. In addition to the Teachers College program, the school had tried to implement a guided reading program, but the principal felt that the teachers weren’t necessarily comfortable with it. The treatment school also tried to add more explicit instruction in reading comprehension strategies by utilizing the Spotlight on Reading (which focuses on reading comprehension skills) and Strategies for Reading (STAR) programs.

In kindergarten through second grade in the treatment school, the principal had supplemented the Teachers College Reading and Writing program with a phonics-based word recognition program (Words Your Way), read alouds with “real literature,” and shared reading with Big Books. Despite these adaptations to the Teachers College program, the principal and many teachers at the school felt that the curriculum was lacking in explicitness and focus, specifically in reading comprehension instruction, which led them to adopt the Urban Education Exchange Comprehension Program.

During the study. The Urban Education Exchange Concepts of Comprehension© program. The treatment school partnered with Urban Education Exchange (UEE) to provide a reading comprehension program beginning in the 2008-2009 school year. The UEE approach to English Language Arts focuses on explicit, structured instruction of reading comprehension skills, beginning in kindergarten. The treatment school implemented the UEE Concepts of Comprehension© program in grades K-5 in 2008-2009 and 2009-2010. This program was used as a supplement to, and did not replace, the existing English Language Arts curriculum in the treatment school.

The Concepts of Comprehension© provide a framework for explicit comprehension instruction. UEE identified twenty-one concepts that allow students to develop the inference-making abilities that underlie successful comprehension of both narrative and expository text. These concepts serve as skills and strategies that the students can utilize to tackle a text or address a reading comprehension problem. UEE has developed a framework that lays out (1) a scope and sequence for the introduction of each concept by grade, (2) learning objectives for each concept by grade, (3) lesson plan units that include scaffolded, explicit instruction and activities that provide independent practice, and (4) over 500 nonfiction passages for grades 2-6, with questions aligned to each Concept of Comprehension©. The lesson plan units use the “I”, “We”, “You” format, which consists of explicit modeling (frequently through read and think alouds of narrative and expository text), guided practice in whole-class or small group activities, and independent practice of the Concepts of Comprehension©. When learned and practiced through UEE’s model lesson units, the Concepts of Comprehension© is designed to enhance students’ ability in inference making and comprehension monitoring. The goal of explicitly teaching these concepts to students is for them to develop the habit of using these concepts as they read, first through thoughtful implementation as strategies and eventually through automatic execution as skills. Teachers also help students develop a meta-awareness in which they use specific cognitive strategies and reason strategically when they encounter difficulties in understanding text (NRP, 2000).

The UEE comprehension lessons in grades K-4 are based on reading aloud a book that lends itself to instruction of a particular Concept of Comprehension©. In the UEE model, reading aloud is an important way teachers help their students develop the attributes of successful readers by adding to the student’s general knowledge, word awareness, and listening

skills. Research has demonstrated that read alouds, if used correctly, can be an effective technique for improving students' comprehension of text (Santoro, Chard, Howard, & Baker 2008). Teachers first activate the students' prior knowledge and then explicitly model using the Concept of Comprehension© while reading aloud. Students then apply the comprehension strategies that have been explicitly modeled to answer questions and discuss literature. In the older grades (5-6), lesson plan units focus on novel studies, which combine instruction of several Concepts of Comprehension© into a four to six week unit. Each unit consists of 5 to 7 lesson plans, each of which follow the same explicit, scaffolded approach that the K-4 lessons use. The units also include worksheets that reinforce vocabulary and the Concepts of Comprehension© taught and a test that assesses mastery of those Concepts of Comprehension©. Teachers can differentiate instruction within the lesson plan units based on whether their students need more or less guidance while reading the novel and completing the activities.

Texts and passages are selected that exemplify and provide practice in using particular Concepts of Comprehension©. Approximately 500 non-fiction passages are available for teachers in grades 2-6, with questions aligned to each passage. The passages come from trade books, Weekly Reader (a children's news magazine), and from another part of the UEE curriculum, General Knowledge, which consists of a wide variety of short expository science and social studies texts. The passages are organized by grade and by the Concept of Comprehension© they reinforce; there are five passages per concept per grade. Each passage is followed by five comprehension questions; three questions on the concept the students are currently studying and two questions on concepts the students have previously learned. Students use a step-by-step reading procedure called Textual Analysis that includes attention to questions that focus on the important information in the text and that highlights particular Concepts of

Comprehension©. Students must “prove their answers,” i.e., identify the parts of the text that support their answers to the questions, and defend their reasoning in class discussions.

Overview of the Concepts of Comprehension© lessons. There are three explicit lesson plans for each Concept of Comprehension© at each grade level, for a total of 300 lesson plans. In addition to these explicit lessons, there are suggestions for developing more lesson plans for each Concept of Comprehension©. Teachers can use the explicit lesson plans to develop their own scripts with books and materials they may feel more comfortable with or have available in their classroom. The teachers at P.S. 15 received the following:

- 300 Concepts of Comprehension lessons
- All activities and assessments available for each lesson
- 500 non-fiction reading passages with comprehension questions
- Textual Analysis teacher training
- UEE Curriculum-specific teacher training

A typical Concepts of Comprehension© lesson. Each Concepts of Comprehension© lesson began with the identification of a learning outcome targeted to the specific Concept of Comprehension© for that lesson. The teacher named the Concept of Comprehension© that the students would be studying, defined it, and told the students how they would use the concept in that specific lesson. Next the teacher would often introduce a short list of vocabulary words that are critical to the Concept of Comprehension© being taught (for either the content or the structure of the book), but not found in the book.

Most lessons began with an activation of prior knowledge, which helped the students make connections to themselves, other texts, and the world and motivated them for the upcoming lesson. The class often also made predictions about the read aloud. Next, the teacher introduced

the book and reminded the students of the Concept of Comprehension© on which they are focusing for that lesson. The teacher modeled using the Concept of Comprehension© during a read aloud session in which she thought aloud about how using the Concept of Comprehension© can help her to understand the story. These think alouds included the teacher monitoring her comprehension of the story and demonstrating how she could use the Concepts of Comprehension© to repair breakdowns in comprehension. The teachers would often draw graphic organizers or other visual displays on chart paper to help make the connection between the Concept of Comprehension© and the book. The lessons identified vocabulary in the books that the students may find difficult and included strategies for helping the students learn the vocabulary.

After the read aloud was finished, the students typically engaged in a guided practice activity. Activities included writing a response or answering questions about the text that focused on the targeted Concept of Comprehension©, completing a graphic organizer, or applying what they had learned to other texts. The teacher introduced the activity to the students and then scaffolded the instruction until the students could complete the activity on their own. After this guided practice session, the students completed a similar activity independently. At the end of the lesson, the students would often share their independent work with the class and receive feedback from the teacher.

Comparison school curriculum.

Before the study. The comparison school also began using the Teachers College Reading and Writing Workshop Program (Calkins, 2000) five years prior to the start of the study (during the 2003-2004 school year), in grades K-6. In a discussion of the curriculum during an initial meeting with the research team, the principal remarked that there was inadequate training at the

launch of the program, and that the Literacy Coach had “difficulty providing professional development” for the teachers in the program. At the same time the school adopted the Teachers College program, it also implemented the Month-by-Month phonics program in the primary grades.

Beginning in the 2006-2007 school year, the comparison school implemented what they called the BEL Mondo comprehension program (Mondo Publishing, 1997) for grades K-2. The program that the comparison school used is a product of Mondo Publishing, whose research division is called the Building Essential Literacy Project (BEL). The program used by the comparison school is based on the findings of the Building Essential Literacy Project (Hill & Jaggard, 2001). The K-2 teachers were trained in the spring of 2005, and subsequently had a full day of training every other month. In addition to these training sessions, a consultant visited the school once a month to attend grade level meetings, conduct model lessons, and introduce new aspects of the program.

During the study. Description of the BEL Mondo and Teachers College Reading Programs. BEL Mondo is a research-based program aligned with the five elements identified by the National Reading Panel as important for effective reading instruction: phonemic awareness, phonics/word study, vocabulary, comprehension, and fluency. BEL Mondo adds oral language as a sixth element that it feels is crucial in teaching successful readers. The program is focused on small group instruction and includes teacher resources to differentiate instruction for all students.

The comparison school teachers received teacher’s guides and lesson plan booklets from Mondo Publishing. The teacher’s guides included explanations of effective instructional routines, a scope and sequence that coordinates whole group instruction based on standards with

differentiated instruction activities based on assessments, classroom management tools, and “reproducible master copies of independent activities.” The lesson plan booklets included oral language, shared reading, and guided reading lessons and activities.

A typical BEL Mondo lesson. A typical BEL Mondo lesson would start with the teacher introducing the read aloud book, often through a picture walk. The teacher would read the title and author of the book and think aloud about the cover, asking the students to make predictions about the book. Often, the teacher would read a short summary of the book from the back cover. Sometimes the teacher would give the students a comprehension strategy (such as character, setting, or sequence) to focus on while she read the book. The teacher would then read aloud the book for the day. During the read aloud, most teachers would stop to ask comprehension questions – usually questions assessing the students’ prior knowledge of the content and vocabulary, but sometimes questions related to the comprehension strategy she had introduced earlier – but some teachers would read straight through the book.

After the read aloud, the students would typically be divided into small groups (3-5 students each) to read books, work on computers, or complete other activities. The teacher would work with one of the groups and, again, introduce the book by looking at the front and back cover, doing a picture walk, and making predictions about the book. She would then read the book together with the group, going slowly to make sure the students were following. In kindergarten classrooms, the teacher would model how to decode words. In second grade classrooms, the teacher would model how to find the meaning of words through context.

In other classrooms, the students would complete an independent activity after the read aloud. The activities ranged from drawing a picture related to the story, filling out a graphic organizer related to the story, to writing a response to the story.

A typical Teachers College Reading and Writing Workshop lesson. The fourth grade lessons that were observed centered around the reading of a short nonfiction text. The teacher introduced the passage to students and reviewed strategies that good readers use. The teacher then modeled strategies to get information out of the text in order to answer comprehension questions.

The fifth grade lessons that were observed centered around read alouds of fiction texts. The teacher would read aloud a short section of the text and then engage students in a discussion, generally concerning the central theme of the text. After the read aloud and discussion, the student split into groups to complete an activity. The activity could be a worksheet of comprehension questions, a written response to the text, or continued reading of the text in groups.

Measures

Assessments. During Year One, all students in grades K-5 completed the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2001); grades 2-5 completed the researcher-developed Urban Education Exchange comprehension assessments (Atkins, 2008); and grades K-1 completed the Stanford 10 Listening Comprehension subtest (Harcourt Educational Measurement, 2004). In Year Two, all students in grades K-5 completed the Group Reading Assessment and Diagnostic Evaluation (GRADE) and grades 2-5 completed the researcher-developed Urban Education Exchange comprehension assessments. See Table 2 in Appendix A for a representation of the assessments completed by each grade in Year One and Year Two.

GRADE (Group Reading Assessment and Diagnostic Evaluation). The GRADE was chosen as a standardized measure of reading comprehension and was administered as a pre and

posttest in the treatment and comparison schools to grades K-5. The GRADE is a norm-referenced reading assessment that is designed to identify the specific reading skills that students have mastered and those which require additional instruction. Reliability coefficients for alternate form and test-retest are in the .90 range. The assessment is untimed, which allows it to measure comprehension, rather than reading speed or fluency. The Kindergarten classrooms were assessed on the Word Reading and Listening Comprehension subtests of GRADE, first and second grade were assessed on the Word Reading, Listening Comprehension, Sentence Comprehension, and Passage Comprehension subtests, and third through fifth grade were assessed on the Vocabulary, Listening Comprehension, Sentence Comprehension, and Passage Comprehension subtests. The GRADE assessment is untimed; most classes completed it during two sessions that lasted between 20 and 45 minutes each in kindergarten, and between 40 and 60 minutes each in the older grades.

UEE comprehension assessments. Comprehension assessments were developed for grades 2-5 to assess mastery of the UEE Concepts of Comprehension©. The tests consisted of multiple choice reading comprehension questions for fiction and non-fiction passages. Each grade was assigned four reading passages with four to five questions per passage. The passages were read aloud by the second grade teachers, while students in third through fifth grade completed the assessment as an independent reading comprehension task. The passages, which were taken from the UEE website, covered multiple Concepts of Comprehension©; that is, each passage allowed the students to practice multiple comprehension strategies and then answer questions that tested their understanding and use of these strategies. Across all the reading passages, all Concepts of Comprehension© were assessed in each grade.

Stanford 10 listening comprehension subtest. During Year One, the kindergarten and first grade students in both the treatment and comparison schools completed the listening comprehension subtest of the Stanford 10 assessment. This subsection of the Stanford 10 first required the students to listen to stories read by the teacher and then answer multiple choice questions. The second part required students to listen to words or sort phrases read by the teacher and then mark the picture that matched the word or phrase. The assessments for both kindergarten and first grade included 40 items (10 for the words section and 30 for the stories section) with three choices per item. This assessment was not administered during Year Two because of negative feedback from the teachers in Year One and also because the results were redundant with those of the GRADE listening comprehension subtest.

Classroom observation protocol. Observations of the literacy classes for grades K-5 were conducted in the treatment and comparison schools. These qualitative analyses assessed the process by which the explicit lesson units were implemented by teachers at the classroom level. The analyses investigated how lesson units are used in practice, and provided information to compare the UEE reading comprehension program with the program used in the comparison school. The guiding questions for these qualitative analyses were:

1. What reading comprehension skills and strategies are explicitly taught during reading instruction in each school?
2. Do teachers model how to use the skills and strategies appropriately before, during, and after reading and explicitly explain when and how the strategy is useful?
3. Do teachers provide the students with sufficient guidance in how to use the skills and strategies?

4. Do teachers provide the students with sufficient opportunity to use the skills and strategies independently, monitoring their understanding and providing useful feedback?

Observation protocols were developed based on coding schemes developed by Gersten et al. (2007), McCutchen et al. (2002), and Vellutino and Scanlon (1996). In these coding schemes a time-sampling procedure is used. This time-sampling procedure was first used by Vellutino and Scanlon (1996) in their observations of kindergarten language arts instruction. In the Vellutino and Scanlon (1996) study, the researchers observed the teacher for a short period (ten or twenty seconds) and then completed a checklist which consisted of a series of questions concerning the instructional group with which the teacher is interacting (individual child, pairs, small group, whole class); general focus of the activity (direct reading, indirect reading, writing, management, preparation); teacher's purpose (reading aloud, modeling, asking/answering questions, guiding practice, providing feedback); the materials being used; the focus of the instruction; and the students' expected response. The observers in this study answered these questions approximately every 30 seconds. They then calculated the percentage of time that the teacher allocated to the various activities within each classroom.

Time sampling procedures. McCutchen and colleagues (2002) used a similar time-sampling procedure in which they completed a checklist of questions every 60 seconds. These questions addressed four categories: knowledge affordance, literacy activity, textual context, and group context. Each observation lasted 15 minutes, after which the total times for the codes within each category were calculated. The time sampling procedure used by Gersten, Dimino, and Jayanthi (2007) was somewhat different from the techniques used by Vellutino and Scanlon (1996) and McCutchen and colleagues (2002). Rather than completing a checklist of questions

every 30-60 seconds, Gersten, Dimino, and Jayanthi (2007) used a fresh observation form every 15 minutes during a 90-minute observation, for a total of 6 intervals.

In the present study, every five minutes a checklist was completed that indicated the grouping arrangement (whole class, small groups, pairs, or individual), the focus of activity (direct reading, indirect reading, writing, management, or preparation activities), the teacher's purpose (reading aloud, modeling, asking/answering questions, guiding student practice, providing feedback), and student engagement (few students seem engaged, many students seem engaged much of the time, or most students seem engaged all of the time). These data gave insight into how the teacher allotted the time in the classroom, while keeping the focus of the observation on the explicitness of instruction. The first three items (grouping arrangement, focus of activity, and teacher's purpose) were found in the observation protocols developed by Scanlon and Vellutino (1996) and McCutchen and colleagues (2002), whereas the measure of student engagement was included to address the ability of the teacher to keep the students on task, the students' interest in the content of the lesson, and the students' motivation for completing the lesson.

Instances of explicit instruction. In addition to this time-sampling procedure, instances of explicit teaching were observed and catalogued. As in Gersten, Dimino, and Jayanthi (2007), the number of instances of explicit teaching was used as a proxy for quality of teaching. Given that instructional programs that include direct, explicit explanations of comprehension strategies have been found to improve the reading comprehension of students, especially at-risk students (NRP 2000; Duffy 2002), it seemed reasonable to Gersten and colleagues that quantifying these instances of explicit teaching could serve as a means of evaluating the quality of instruction.

Instances of explicit teaching could be found during the pre-reading, modeling, independent practice, and feedback sections of the lesson.

An example of explicit teaching during the *pre-reading* section could be the teacher explicitly stating the learning outcome for the lesson. An explicit *modeling* of a comprehension strategy could involve the teacher thinking aloud as she predicts what will happen next in the story. During *independent practice*, explicit instruction would allow the students to practice the comprehension concept that had been modeled. Explicit *feedback* requires the teacher to explicitly state how the student has used the skill or strategy correctly or incorrectly. These categories were adopted from Gersten, Dimino, and Jayanthi (2007), and the observation form also provided space to note which Concept of Comprehension© was being explicitly taught to make the observation form more tailored to the UEE curriculum.

Evaluative judgments. After the observation was completed, the observer answered questions designed to evaluate the lesson as a whole. There are three yes/no questions and five Likert scale questions that address the overall quality of comprehension and vocabulary instruction and the management practices of the teacher (Gersten, Dimino, & Jayanthi, 2007). (See Figure 1 in Appendix A for the observation form used in the classrooms.)

Procedure

At the beginning of Year 1, the research study was explained to the faculty in both the treatment school and control school. The teachers were told that the goal of the study was to understand how explicit instruction “looks and works” in the classroom. The quantitative (assessments) and qualitative (observations) aspects of the study were introduced and the importance of both parts was emphasized. The standardized and researcher-developed assessments were introduced, and it was stressed that because they are designed to measure

reading comprehension, teachers should provide assistance in decoding words and give the students as much time as they need to complete the assessments. The consent procedure was also introduced to the teachers.

Professional development. The treatment school received two days of professional development in September 2008, before implementation of the UEE Comprehension Program. In addition, the school received fifteen days of professional development during the school year, from September 2008 to May 2009. This schedule of professional development was repeated for Year Two of the study (2009-2010). The training included a description of the Concepts of Comprehension©, implementation techniques, and modeling of specific lessons. Teacher trainers helped teachers understand and implement UEE's Concepts of Comprehension© model lesson units and provided guidance to teachers in the development of additional lessons appropriate for their students. The teacher trainers meet with the teachers approximately once every one or two weeks to help plan a comprehension unit, either observe or model the planned lesson, and then debrief with the teacher. In addition, the teacher trainers helped teachers understand the process of comprehension and the research behind the UEE curriculum, including the importance of explicit instruction of the comprehension strategies. Because the Concepts of Comprehension© program was evaluated as a whole, including both the curriculum and the teacher training, it was not necessary to provide equal teacher training to the comparison school. As an incentive for participation, however, UEE agreed to pay for a similar period of professional development in the comparison school.

Assessments. Year 1. During Year One, standardized English Language Arts assessments were conducted in December 2008 and June 2009 using the Group Reading Assessment and Diagnostic Evaluation (GRADE) in grades K-5 of both schools. The students

completed Form A in the fall and Form B in the spring. The Kindergarten classrooms were assessed on the Word Reading and Listening Comprehension subtests of GRADE, first and second grade were assessed on the Word Reading, Listening Comprehension, Sentence Comprehension, and Passage Comprehension subtests, and third through fifth grade were assessed on the Vocabulary, Listening Comprehension, Sentence Comprehension, and Passage Comprehension subtests.

In addition, the second through fifth grade students' reading comprehension was measured with researcher-developed tests in October 2008 and June 2009. The second grade teachers read the passages and questions to their students, so this was a test of listening comprehension, whereas third, fourth, and fifth graders read the passages and questions independently. The researcher-developed tests were untimed, but each grade took approximately one hour in a single session to complete the test.

As the kindergarten and first grade students did not receive the researcher-developed tests, these grades completed the listening comprehension subsection of the Stanford 10 Achievement Test as an additional measure. This assessment was administered in November 2008 and June 2009. This subsection of the Stanford 10 required the students to listen to words and stories read by the teacher and answer multiple choice questions. The listening comprehension section consisted of 40 items with 3 choices per item for both kindergarten and first grade. This test took approximately two hours over two sessions to complete.

All tests were group-administered by the teacher in the classroom.

The teachers provided feedback on the assessments when they were collected. The K-1 teachers indicated that all the assessments had been difficult for the students to complete, especially for the Kindergarten students. The teachers felt that asking the students to listen to 12

passages and answer questions on the Stanford Listening Comprehension test was, in particular, expecting too much from Kindergarteners who had just started school. The Kindergarten students did not have any experience with whole-class testing, so it took a lot of work on the teachers' part to make them feel comfortable. Several teachers remarked that they would try to integrate some of the lessons they learned from the assessments into future instruction. These lessons included teaching students how to listen carefully, sequence events, bubble in answers, answer questions using the process of elimination, and "think through a test."

Year 2. During Year Two, the students in the treatment and comparison schools completed Form A of the GRADE in October 2009. This was followed by the administration of the researcher-developed comprehension assessment in early November 2009. In Year Two of the study, the K-1 classrooms did not complete the Stanford 10 Achievement Test (the listening comprehension subsection). This decision was made for two reasons. First, the teachers had provided feedback that this assessment, especially in the fall, was very difficult for their students and resulted in a lot of frustration, crying, and students feeling ill. Some teachers complained that the subtest took the entire day to administer and the students could not sit still and listen long enough to complete all the passages. Second, the results of this subtest did not differ dramatically from the listening comprehension subtest in the GRADE, so this assessment was seen as redundant. In May 2010, the students completed the second form of the researcher-developed comprehension assessment. Form B of the GRADE was administered to both schools in early June 2010. See Appendix A, Table 2 for the assessments completed each year in each grade.

Classroom observations. The classroom observations were designed to investigate the level of explicitness of instruction provided by each teacher. The original plan was to observe

each participating teacher twice during each of the two school years. However, several teachers in grades K-2 were observed three times a year, as observations and assessments completed early in the study indicated greater differences in instruction and student achievement between the two schools in these earlier grades. Some teachers were observed only once during the year as a result of time constraints and a lack of cooperation from teachers who had previously agreed to be observed. This was a greater issue during Year Two, when many teachers claimed that they had too much testing and other activities scheduled in April, May, and June to allow observations of their instruction. Teachers were typically notified that they would be observed the morning of the observation. For those times that the teachers were not able to accommodate the observation, an alternate date was agreed upon. All observations were conducted by the first author.

In the treatment school, the researcher observed the UEE Concepts of Comprehension© lessons. In the comparison school, the researcher observed the listening or reading comprehension lessons; these consisted of read aloud lessons, shared reading lessons, or other types of lessons in which the students were learning and practicing how to comprehend text.

During the observation, the observation form was completed, which indicated how time was allotted during the lesson as well as the explicitness of instruction. The observer looked for instances of explicit instruction in four sections of the lesson: introduction (activating prior knowledge, introducing the text, making connections between the text and students, other texts, or the world), direct reading (reading the text with modeling of strategies or think alouds), practice (both guided practice and independent practice), and feedback.

Instances of explicit instruction would manifest differently depending on the section of the lesson in which they are contained. In the *introduction*, an instance of explicit instruction

could include explicitly stating the learning outcome for the lesson, defining or giving an example of a comprehension strategy, introducing new vocabulary, looking at or reading the cover or back of the book while making predictions, or making explicit connections between the text and another text the students have read. The type of explicit instruction observed during the introduction was noted on the observation form.

During *direct reading*, an instance of explicit instruction would involve modeling or thinking aloud about a comprehension strategy for the class. This type of explicit instruction occurred when the teacher stopped reading the text and demonstrated a comprehension strategy that could be employed to help understand the text. The observation form was marked according to the comprehension strategy that was modeled and had boxes for each of the 21 Concepts of Comprehension©. The observer took notes on how the teacher modeled or explained the comprehension strategy.

During the *practice* section of the lesson, an instance of explicit instruction could include a check for student understanding, guided practice of the modeled skill, independent practice of the modeled skill, asking a literal or inferential question from the text (these types of questions were distinguished on the observation form), or asking students to justify or elaborate responses.

Instances of explicit instruction marked as *feedback* included when the teacher states clearly what a student did correctly or incorrectly about the comprehension strategy. In order to be marked as explicit feedback, the teacher had to reference the comprehension strategy being studied. General feedback that did not reference the comprehension skill or strategy being studied was not counted as explicit feedback.

After each observation, the instances of explicit instruction were totaled for each section of the lesson, and across the lesson as a whole. Therefore, for each lesson that was observed, the

teacher received an explicitness score for the introduction, direct reading, student practice, and feedback sections, as well as for the lesson as a whole.

Year 1. During the 2008-2009 school year, 12 of a possible 20 teachers in the treatment school (two per grade) and 14 of a possible 19 teachers in the comparison school (three in kindergarten, first grade and fourth grade; two in second and third grade; and one in fifth grade) agreed to allow observations of their ELA classes. A total of 43 observations were conducted in the two schools: 20 observations in the treatment school and 23 observations in the comparison school. See Appendix A, Table 3 for the number of observations conducted per grade in each school during Year One and Year Two.

The observations in the treatment school began in October, and two classrooms per grade were observed. There were a total of four observations in Kindergarten (two per classroom), four observations in first grade (two per classroom), five observations in second grade (two in one classroom and three in the other classroom), three observations in third grade (two in one classroom and one in the other classroom), two observations in fourth grade (one per classroom), and two observations in fifth grade (one per classroom).

In the comparison school, the observations began in January because the teachers declined to be observed during the fall semester. In Kindergarten, there were a total of six observations conducted across three classrooms (two per classroom). In first grade, there were a total of eight observations conducted across three classrooms (two or three per classroom). There were three observations across two classrooms in second grade, two observations across two classrooms in third grade, three observations across three classrooms in fourth grade, and one observation in one classroom in fifth grade.

Year 2. During the 2009-2010 school year, 12 of a possible 20 teachers in the treatment school (two each in kindergarten, first, second, and fifth grades; one in third grade; and three in fourth grade) and 9 of a possible 19 teachers (two each in kindergarten, third grade and fifth grade; and one in first, second, and fourth grade) agreed to the observations. A total of 37 observations were conducted in the two schools during this year: 22 in the treatment school and 15 in the comparison school. Observations in both schools lasted from November 2009 to May 2010.

In the treatment school, there were a total of four observations across two classrooms in kindergarten (two per classroom), four observations in two classrooms in first grade (two per classroom), four observations in two classrooms in second grade (two per classroom), two observations in one third grade classroom, five observations in three fourth grade classrooms (one or two per classroom), and three observations in two fifth grade classrooms (one or two per classroom).

In the comparison school, there were a total of four observations across two kindergarten classrooms (two per classroom), two observations in one first grade classroom, two observations in one second grade classroom, three observations in two third grade classrooms (one or two per classroom), one observation in one fourth grade classroom, and three observations in two fifth grade classrooms (one or two per classroom).

Interviews. At the conclusion of Year One and Year Two of the study (June 2009 and June 2010), a total of 20 participating teachers were interviewed to obtain their views on the curriculum they taught and teaching reading comprehension in general. They were asked how they felt their class responded to the materials and how they felt the curriculum could have been more effective. This interview took between 10 and 30 minutes per teacher.

Scoring of dependent measures. Both the standardized and researcher-developed measures were multiple-choice assessments. Research assistants were hired to hand-score the assessments. The research assistants entered the students' answers into a spreadsheet created for each classroom. The spreadsheet automatically scored the response as correct or incorrect, and descriptive statistics were calculated from these spreadsheets. Mean scores, standard deviations, and proportion correct for each classroom on each assessment or subtest were obtained. Independent samples t-tests were conducted to compare the treatment and comparison schools on each assessment or subtest at each grade level. There were a total of 28 comparisons in the fall of Year 1, 28 comparisons in the spring of Year 1, 26 comparisons in the fall of Year 2, and 26 comparisons in the spring of Year 2.

To answer Research Question 1 (see page 13), analyses of covariance were conducted at each grade level for each assessment, using the fall test as the covariate and the spring test as the dependent measure. The total score on the researcher-developed UEE Comprehension Assessment and on the Stanford 10 Listening subtest were examined. Each subtest of the GRADE was examined separately (Word Reading, Vocabulary, Listening Comprehension, Sentence Comprehension, and Paragraph Comprehension). Overall, 54 comparisons of the scores of the treatment school and the comparison school were conducted.

To answer Research Question 2 (see page 13), the observational data were examined. For each observation that was conducted, the number of instances of explicit instruction that were observed during the introduction, direct reading, student practice, and feedbacks sections of the lessons was summed. A grand total of instances for each lesson was also computed. The mean number of instances of explicit instruction for each teacher, each grade, and each school across all the observations that were conducted in Year 1 and Year 2 was obtained.

To answer Research Question 3 (see page 13), the correlation between the explicitness scores that were given to the teachers during the observations and the mean scores of the students in their classrooms on the standardized and researcher-developed assessments were calculated. The teachers were also grouped into high-explicit and low-explicit groups according to their explicitness scores, and the mean scores of the students in the high-explicit and low-explicit teacher groups were compared.

To answer Research Question 4 (see page 13), the teachers in the treatment school were split into two groups based on whether they were teaching the explicit program for their first year or their second year at the same grade level. The mean number of instances of explicit instruction for these two groups of treatment teachers was compared to determine if the more experienced treatment teachers demonstrated more explicit instruction. The teachers in the comparison school were also split into two groups based on whether they were new teachers (and were therefore teaching the comparison program for their first year) or they had been teaching at the comparison school (and using the comparison program) for more than one year (an undetermined number). The mean number of instances of explicit instruction for these two groups of comparison teachers was compared to determine if the more experienced comparison teachers demonstrated more explicit instruction.

To answer Research Question 5 (see page 14), the mean scores of the students of the new teachers and the mean scores of the students of the experienced teachers within each school were compared. Adjusted posttest scores (that treated the pretests as covariates) for each of the subsections of the GRADE were calculated for each classroom, and the classroom scores of the new teachers and the experienced teachers were compared.

To answer Research Question 6 (see page 14), the students of the experienced teachers in both schools were separated. The adjusted posttest scores (that treated the pretests as covariates) for each of the subsections of the GRADE were calculated for each classroom of the experienced teachers, and the classroom scores of the experienced treatment teachers and the experienced comparison teachers were compared.

To answer Research Question 7 (see page 14), the students who attended the same school for two years in a row were identified. Adjusted posttest scores for the Year 2 posttests (that treated the Year 1 pretests as covariates) were calculated for each classroom, and the scores of the two-year students in the treatment school and the two-year students in the comparison school were compared.

To answer Research Question 8 (see page 14), the two-year students in both schools who had experienced teachers during Year 2 were identified. Adjusted posttest scores for the Year 2 posttests (that treated the Year 1 pretests as covariates) were calculated for each classroom of the experienced teachers, and the scores of the two-year students with experienced teachers in the treatment school and the two-year students with experienced teachers in the comparison school were compared.

CHAPTER IV

Results

Research Question 1: Does instruction in reading comprehension that focuses on the explicit teaching of reading comprehension skills and strategies (Program T) improve performance on standardized and researcher-developed reading comprehension tests more than a traditional reading comprehension program (Program C)?

The results from Year One and Year Two will be presented together, organized by assessment (GRADE and the researcher-developed UEE Comprehension Assessment) and grade level, to facilitate comparisons between the scores in the two years. First, descriptive statistics (means, standard deviations, and proportion correct) were computed for each classroom on each assessment. Then the scores for all the classrooms at each grade level were summed and the mean scores for each grade of the two schools on each assessment were compared using independent samples t-tests.

Following these analyses, analyses of covariance were conducted to test for differences between the treatment and comparison schools on each assessment on the spring posttests for Year 1 and Year 2. For each assessment and each year, the fall test was used as the covariate and the spring test as the dependent measure. Only students that completed the pretest and posttest in the same classroom were included in these analyses. These analyses revealed differences between the treatment and comparison schools on the posttests, controlling for the pretest scores. Each subtest of the GRADE (Word Reading, Vocabulary, Listening Comprehension, Sentence Comprehension, and Passage Comprehension) was analyzed separately. The total scores on the researcher-developed UEE Comprehension Assessment were used for the analysis. Overall, 64

comparisons of the treatment school vs. the comparison school mean scores were performed (not all grades received all five subtests of the GRADE assessment).

Tables 4 and 5 in Appendix B present a summary of the findings. The mean proportion correct and standard deviations that are summarized in Tables 4 and 5 appear in Tables 6 through 29. The text that follows describes the findings in Appendix B, Tables 6 through 29.

Group Reading Assessment and Diagnostic Evaluation (GRADE)

All grades combined. Students in all grades completed the Listening Comprehension subtest, students in grades K-2 completed the Word Reading subtest, students in grades 1-5 completed the Sentence Comprehension and Passage Comprehension subtests, and students in grades 3-5 completed the Vocabulary subtest of the GRADE. The results of these assessments can be seen in Tables 6 and 7. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the difference between the treatment school ($M=.81$, $SD=.21$) and the comparison school ($M=.81$, $SD=.17$), $F(df=1, 626)=2.27$, $p=.13$, was not significant in year 1. The difference between the treatment school ($M=.85$, $SD=.12$) and the comparison school ($M=.85$, $SD=.10$) was also not significant in year 2 ($F(df=1, 644) = .01$, $p=.95$).

On the *Word Reading/Vocabulary* subtests, the difference between the treatment school ($M=.69$, $SD=.24$) and the comparison school ($M=.67$, $SD=.25$) was not significant in year 1 ($F(df=1, 621) = 1.96$, $p=.16$). The difference between the treatment school ($M=.71$, $SD=.24$) and the comparison school ($M=.71$, $SD=.24$) was also not significant in year 2 ($F(df=1, 659) = 1.98$, $p=.16$).

On the *Sentence Comprehension* subtest, the treatment school ($M=.76$, $SD=.23$) scored significantly higher than the comparison school ($M=.67$, $SD=.27$) in year 1 ($F(df=1, 529) = 34.95$, $p=.00$). In year 2, the comparison school ($M=.75$, $SD=.22$) scored significantly higher than the treatment school ($M=.74$, $SD=.23$), $F(df=1, 548) = 6.27$, $p=.01$.

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.59$, $SD=.22$) and the comparison school ($M=.54$, $SD=.22$) was not significant in year 1 ($F(df=1, 527) = .78$, $p=.38$). The treatment school ($M=.61$, $SD=.19$) scored significantly higher than the comparison school ($M=.56$, $SD=.22$) in year 2 ($F(df=1, 549) = 6.99$, $p=.01$).

Grades K-2 combined. Students in grades K-2 completed the Listening Comprehension and Word Reading subtests, and students in grades 1-2 completed the Sentence Comprehension and Passage Comprehension subtests. The results of these assessments can be seen in Tables 8 and 9. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the difference between the treatment school ($M=.87$, $SD=.17$) and the comparison school ($M=.83$, $SD=.19$), $F(df=1, 283) = .08$, $p=.78$, was not significant in year 1. In year 2, the treatment school ($M=.89$, $SD=.11$) scored significantly higher than the comparison school ($M=.87$, $SD=.09$), $F(df=1, 312) = 4.81$, $p=.03$.

On the *Word Reading* subtest, the difference between the treatment school ($M=.85$, $SD=.21$) and the comparison school ($M=.85$, $SD=.19$) was not significant in year 1 ($F(df=1, 270) = .16$, $p=.69$). The difference between the treatment school ($M=.89$, $SD=.16$) and the comparison school ($M=.87$, $SD=.17$) was also not significant in year 2 ($F(df=1, 297) = .14$, $p=.71$).

On the *Sentence Comprehension* subtest, the treatment school ($M=.74$, $SD=.25$) scored significantly higher than the comparison school ($M=.65$, $SD=.28$) in year 1 ($F(df=1, 179) = 9.12$, $p=.00$). The difference between the treatment school ($M=.72$, $SD=.26$) and the comparison school ($M=.75$, $SD=.25$) was not significant in year 2 ($F(df=1, 190) = 1.89$, $p=.17$).

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.58$, $SD=.21$) and the comparison school ($M=.53$, $SD=.23$) was not significant in year 1 ($F(df=1, 191) = .25$, $p=.62$). The treatment school ($M=.61$, $SD=.20$) scored significantly higher than the comparison school ($M=.57$, $SD=.21$) in year 2 ($F(df=1, 186) = 6.53$, $p=.01$).

Grades 3-5 combined. Students in grades 3-5 completed the Listening Comprehension, Vocabulary, Sentence Comprehension, and Passage Comprehension subtests of the GRADE. The results of these assessments can be seen in Tables 10 and 11. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the comparison school ($M=.79$, $SD=.16$) scored significantly higher than the treatment school ($M=.77$, $SD=.22$) in year 1, $F(df=1, 340) = 4.04$, $p=.05$. The difference between the treatment school ($M=.82$, $SD=.11$) and the comparison school ($M=.83$, $SD=.10$) approached significance in year 2 ($F(df=1, 329) = 3.07$, $p=.08$).

On the *Vocabulary* subtests, the difference between the treatment school ($M=.58$, $SD=.19$) and the comparison school ($M=.52$, $SD=.20$) was not significant in year 1 ($F(df=1, 348) = .03$, $p=.87$). The difference between the treatment school ($M=.58$, $SD=.19$) and the comparison school ($M=.55$, $SD=.20$) was also not significant in year 2 ($F(df=1, 359) = .02$, $p=.88$).

On the *Sentence Comprehension* subtest, the treatment school ($M=.77$, $SD=.22$) scored significantly higher than the comparison school ($M=.68$, $SD=.26$) in year 1 ($F(df=1, 347) = 30.96$, $p=.00$). In year 2, the comparison school ($M=.76$, $SD=.21$) scored significantly higher than the treatment school ($M=.75$, $SD=.21$), $F(df=1, 355) = 4.41$, $p=.04$.

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.59$, $SD=.23$) and the comparison school ($M=.54$, $SD=.21$) was not significant in year 1 ($F(df=1, 333) = .08$, $p=.78$). The treatment school ($M=.61$, $SD=.19$) scored significantly higher than the comparison school ($M=.56$, $SD=.22$) in year 2 ($F(df=1, 360) = 4.26$, $p=.04$).

Kindergarten. Kindergarten students in both schools completed the Listening Comprehension and the Word Reading subtests of the GRADE. The results of these subtests are found in Tables 12 and 13. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

The *Listening Comprehension* subtest consisted of 18 items in which the teacher read a sentence and the students chose the correct picture that corresponded to the sentence. The difference between the treatment school ($M=.89$, $SD=.17$) and the comparison school ($M=.88$, $SD=.16$) was not significant in year 1 ($F(df=1, 77) = .82$, $p=.37$). In year 2 the treatment school ($M=.94$, $SD=.07$) scored significantly higher than the comparison school ($M=.89$, $SD=.09$), $F(df=1, 94) = 11.51$, $p=.001$.

The kindergarten *Word Reading* subtest consisted of 10 items in which the students chose the word that the teacher read aloud. The difference between the treatment school ($M=.80$, $SD=.28$) and the comparison school ($M=.85$, $SD=.23$) was not significant in year 1 ($F(df=1, 70) = .29$, $p=.59$). The difference between the treatment school ($M=.84$, $SD=.23$) and the

comparison school ($M=.83$, $SD=.24$) was also not significant in year 2 ($F(df=1, 90) = .70$, $p=.40$).

First grade. First grade students in both schools completed the Listening Comprehension (17 items), Word Reading (20 items), Sentence Comprehension (19 items), and Passage Comprehension (24 items) subtests of the GRADE. The results of these assessments are found in Tables 14 and 15. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the difference between the treatment school ($M=.84$, $SD=.21$) and the comparison school ($M=.84$, $SD=.21$) was not significant in year 1 ($F(df=1, 107) = .08$, $p=.77$). The difference between the treatment school ($M=.90$, $SD=.08$) and the comparison school ($M=.91$, $SD=.07$) was also not significant in year 2 ($F(df=1, 99) = .39$, $p=.54$).

On the *Word Reading* subtest, the difference between the treatment school ($M=.81$, $SD=.24$) and the comparison school ($M=.84$, $SD=.17$) was not significant in year 1 ($F(df=1, 102) = .01$, $p=.93$). The difference between the treatment school ($M=.88$, $SD=.14$) and the comparison school ($M=.90$, $SD=.13$) was also not significant in year 2 ($F(df=1, 87) = .22$, $p=.64$).

On the *Sentence Comprehension* subtest, the treatment school ($M=.73$, $SD=.26$) scored significantly higher than the comparison school ($M=.62$, $SD=.30$) in year 1 ($F(df=1, 101) = 7.66$, $p=.01$). The difference between the treatment school ($M=.72$, $SD=.27$) and the comparison school ($M=.76$, $SD=.28$) was not significant in year 2 ($F(df=1, 72) = .02$, $p=.89$).

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.55$, $SD=.23$) and the comparison school ($M=.47$, $SD=.25$) was not significant in year 1 ($F(df=1, 102) = 2.00$, $p=.16$). In year 2, the treatment school ($M=.58$, $SD=.21$) scored significantly higher than the comparison school ($M=.56$, $SD=.25$), $F(df=1, 85) = 5.88$, $p=.02$.

Second grade. Second grade students in both schools also completed the Listening Comprehension (17 items), Word Reading (28 items), Sentence Comprehension (19 items), and Passage Comprehension (28 items) subtests of the GRADE. The results of these assessments can be seen in Tables 16 and 17. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the treatment school ($M=.88$, $SD=.10$) scored significantly higher than the comparison school ($M=.75$, $SD=.16$) in year 1, $F(df=1, 93) = 12.02$, $p=.00$. The difference between the treatment school ($M=.85$, $SD=.14$) and the comparison school ($M=.82$, $SD=.10$) approached significance in year 2, $F(df=1, 113) = 2.72$, $p=.10$.

On the *Word Reading* subtest, the difference between the treatment school ($M=.92$, $SD=.10$) and the comparison school ($M=.86$, $SD=.15$) was not significant in year 1 ($F(df=1, 92) = 2.65$, $p=.11$). In year 2, the treatment school ($M=.94$, $SD=.09$) scored significantly higher than the comparison school ($M=.88$, $SD=.12$), $F(df=1, 114) = 10.44$, $p=.00$.

On the *Sentence Comprehension* subtest, the difference between the treatment school ($M=.77$, $SD=.24$) and the comparison school ($M=.69$, $SD=.26$) was not significant in year 1 ($F(df=1, 75) = 1.38$, $p=.24$). The difference between the treatment school ($M=.72$, $SD=.25$) and the comparison school ($M=.74$, $SD=.23$) approached significance in year 2, $F(df=1, 115) = 3.39$, $p=.07$.

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.61$, $SD=.20$) and the comparison school ($M=.60$, $SD=.18$) was not significant in year 1, $F(df=1, 86) = .96$, $p=.33$. The difference between the treatment school ($M=.64$, $SD=.17$) and the comparison school ($M=.57$, $SD=.19$) was also not significant in year 2, $F(df=1, 98) = 1.20$, $p=.28$.

Third grade. Third grade students in both schools completed the Listening Comprehension (17 items), Vocabulary (30 items), Sentence Comprehension (19 items), and Passage Comprehension (28 items) subtests of the GRADE. Tables 18 and 19 display the results of these assessments. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the difference between the treatment school ($M=.83$, $SD=.10$) and the comparison school ($M=.84$, $SD=.12$) was not significant in year 1 ($F(df=1, 101) = .33$, $p=.57$). The difference between the treatment school ($M=.85$, $SD=.12$) and the comparison school ($M=.86$, $SD=.10$) was also not significant in year 2 ($F(df=1, 99) = .08$, $p=.78$).

On the *Vocabulary* subtest, the difference between the treatment school ($M=.69$, $SD=.15$) and the comparison school ($M=.67$, $SD=.18$) was not significant in year 1 ($F(df=1, 98) = 1.16$, $p=.28$). The difference between the treatment school ($M=.72$, $SD=.17$) and the comparison school ($M=.67$, $SD=.20$) was also not significant in year 2 ($F(df=1, 100) = 1.33$, $p=.25$).

On the *Sentence Comprehension* subtest, the difference between the treatment school ($M=.88$, $SD=.11$) and the comparison school ($M=.87$, $SD=.14$) was not significant in year 1 ($F(df=1, 105) = .02$, $p=.89$). The difference between the treatment school ($M=.85$, $SD=.18$) and the

comparison school ($M=.87$, $SD=.15$) was also not significant in year 2 ($F(df=1, 106) = 1.15$, $p=.29$).

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.60$, $SD=.20$) and the comparison school ($M=.54$, $SD=.23$) was not significant in year 1 ($F(df=1, 98) = .53$, $p=.47$). The difference between the treatment school ($M=.59$, $SD=.19$) and the comparison school ($M=.56$, $SD=.19$) was also not significant in year 2 ($F(df=1, 100) = .16$, $p=.69$).

Fourth grade. The fourth grade students in both schools completed the same subtests of the GRADE that the third grade students completed: Listening Comprehension (17 items), Vocabulary (35 items), Sentence Comprehension (19 items), and Passage Comprehension (28 items) subtests. The results of these assessments can be found in Tables 20 and 21. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the treatment school ($M=.67$, $SD=.32$) scored significantly lower than the comparison school ($M=.74$, $SD=.40$), $F(df=1, 124) = 4.65$, $p=.03$, in year 1. The treatment school ($M=.79$, $SD=.10$) also scored significantly lower than the comparison school ($M=.81$, $SD=.10$), in year 2, $F(df=1, 116) = 4.50$, $p=.03$.

On the *Vocabulary* subtest, the difference between the treatment school ($M=.55$, $SD=.19$) and the comparison school ($M=.46$, $SD=.21$) was not significant in year 1 ($F(df=1, 123) = .08$, $p=.78$). The difference between the treatment school ($M=.57$, $SD=.18$) and the comparison school ($M=.48$, $SD=.19$) was also not significant in year 2, $F(df=1, 115) = 1.97$, $p=.16$.

On the *Sentence Comprehension* subtest, the treatment school ($M=.71$, $SD=.23$) scored significantly higher than the comparison school ($M=.57$, $SD=.29$) in year 1, $F(df=1, 123)$

=44.57, $p=.00$. The difference between the treatment school ($M=.73$, $SD=.19$) and the comparison school ($M=.71$, $SD=.19$) approached significance in year 2, $F(df=1, 110) = 2.99$, $p=.09$.

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.51$, $SD=.25$) and the comparison school ($M=.48$, $SD=.22$) was not significant in year 1, $F(df=1, 124) = .16$, $p=.69$. The difference between the treatment school ($M=.60$, $SD=.21$) and the comparison school ($M=.55$, $SD=.21$) was also not significant in year 2, $F(df=1, 118) = 1.64$, $p=.20$.

Fifth grade. The fifth grade students in both schools completed the same subtests of the GRADE that the third and fourth grade students had completed: the Listening Comprehension (17 items), Vocabulary (35 items), Sentence Comprehension (19 items), and Passage Comprehension (30 items) subtests. Tables 22 and 23 display the results of these assessments. For each year and each subtest separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate.

On the *Listening Comprehension* subtest, the difference between the treatment school ($M=.84$, $SD=.07$) and the comparison school ($M=.82$, $SD=.09$) was not significant in year 1 ($F(df=1) = .97$, $p=.33$). The difference between the treatment school ($M=.81$, $SD=.11$) and the comparison school ($M=.83$, $SD=.09$) was also not significant in year 2 ($F(df=1) = .75$, $p=.39$).

On the *Vocabulary* subtest, the difference between the treatment school ($M=.50$, $SD=.17$) and the comparison school ($M=.47$, $SD=.15$) was not significant in year 1 ($F(df=1) = .06$, $p=.80$). The treatment school ($M=.49$, $SD=.15$) scored significantly lower than the comparison school ($M=.51$, $SD=.15$), $F(df=1) = 6.00$, $p=.02$, in year 2.

On the *Sentence Comprehension* subtest, the difference between the treatment school ($M=.73$, $SD=.24$) and the comparison school ($M=.64$, $SD=.23$) was not significant in year 1, $F(df=1, 113) = 1.93$, $p=.17$. The difference between the treatment school ($M=.68$, $SD=.23$) and the comparison school ($M=.70$, $SD=.23$) was also not significant in year 2, $F(df=1, 133) = .13$, $p=.72$.

On the *Passage Comprehension* subtest, the difference between the treatment school ($M=.67$, $SD=.20$) and the comparison school ($M=.60$, $SD=.17$) was not significant in year 1, $F(df=1, 105) = .60$, $p=.44$. The treatment school ($M=.64$, $SD=.18$) scored significantly higher than the comparison school ($M=.57$, $SD=.25$), in year 2, $F(df=1, 136) = 7.08$, $p=.01$.

UEE Comprehension Assessment

All grades (2-5) combined. Students in grades 2-5 completed the researcher-developed UEE comprehension assessments, which consisted of four comprehension passages with multiple choice comprehension questions, in year 1 and year 2. The second grade teachers read the passages to their students, whereas the students in grades 3-5 completed the assessment as an independent reading task. The results of these assessments can be seen in Table 24. For each year, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the treatment school ($M=.65$, $SD=.19$) and the comparison school ($M=.60$, $SD=.20$) was not significant in year 1, $F(df=1, 428) = .62$, $p=.43$. The difference between the treatment school ($M=.66$, $SD=.21$) and the comparison school ($M=.63$, $SD=.19$) was also not significant in year 2, $F(df=1, 411) = .05$, $p=.83$.

Grades 3-5 combined. Students in grades 3-5 completed the researcher-developed UEE *reading* comprehension assessments, which consisted of four comprehension passages with multiple choice comprehension questions, in year 1 and year 2. The results of these assessments can be seen in Table 25. For each year, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the treatment school ($M=.66$, $SD=.18$) and the comparison school ($M=.60$, $SD=.20$) was not significant in year 1, $F(df=1, 349) = 1.18$, $p=.28$. The difference between the treatment school ($M=.62$, $SD=.22$) and the comparison school ($M=.64$, $SD=.20$) was also not significant in year 2, $F(df=1, 286) = 1.29$, $p=.26$.

Second grade. The researcher-developed UEE comprehension assessment for second grade consisted of four *listening* comprehension passages and 18 items in total, with each item covering a different Concept of Comprehension©. The results of this assessment can be found in Table 26. For each year separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the treatment school ($M=.61$, $SD=.22$) and the comparison school ($M=.58$, $SD=.19$) was not significant in year 1 ($F(df=1, 76) = .13$, $p=.72$). The treatment school ($M=.74$, $SD=.15$) scored significantly higher than the comparison school ($M=.62$, $SD=.16$), in year 2, $F(df=1, 122) = 6.46$, $p=.01$.

Third grade. The researcher-developed UEE comprehension assessment for third grade consisted of four *reading* passages and 19 items, with each one covering a different Concept of Comprehension©. The results of the third grade UEE reading comprehension test are in Table 27. For each year separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the

treatment school ($M=.65$, $SD=.18$) and the comparison school ($M=.61$, $SD=.20$) was not significant in year 1 ($F(df=1, 111) = .21$, $p=.65$). The difference between the treatment school ($M=.66$, $SD=.20$) and the comparison school ($M=.60$, $SD=.19$) was also not significant in year 2 ($F(df=1, 92) = .02$, $p=.88$).

Fourth grade. The researcher-developed UEE comprehension assessment for fourth grade consisted of four *reading* passages and 20 items, with each item covering a different Concept of Comprehension©. The results of this assessment can be found in Table 28. For each year separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the treatment school ($M=.73$, $SD=.15$) and the comparison school ($M=.66$, $SD=.20$) was not significant in year 1 ($F(df=1, 116) = .93$, $p=.34$). The difference between the treatment school ($M=.70$, $SD=.20$) and the comparison school ($M=.72$, $SD=.19$) was also not significant in year 2 ($F(df=1, 107) = .61$, $p=.44$).

Fifth grade. The researcher-developed UEE comprehension assessment for fifth grade consisted of four *reading* passages and 20 items, with each item covering a different Concept of Comprehension©. The results of this assessment are in Table 29. For each year separately, the mean proportion correct on the spring posttest was subjected to analysis of covariance, using the fall pretest score as a covariate. The difference between the treatment school ($M=.60$, $SD=.18$) and the comparison school ($M=.54$, $SD=.20$) was not significant in year 1 ($F(df=1, 116) = .01$, $p=.94$). The treatment school ($M=.48$, $SD=.19$) scored significantly higher than the comparison school ($M=.58$, $SD=.20$) in year 2, $F(df=1, 81) = 4.90$, $p = .03$.

Observations

Research Question 2: Does Program T result in more instances of explicit instruction during reading comprehension lessons than Program C, as measured by classroom observations?

The observational data collected were analyzed to determine the differences between Program T and Program C in the mean number of instances of explicit instruction observed per lesson. See Appendix C. The number of instances of explicit instruction that was observed during each of the four sections of the lesson (introduction, direct reading, student practice, and feedback) and for each lesson as a whole is presented in Table 30, along with the totals across all the lessons in the treatment school and in the comparison school, in year 1 and year 2 separately, and then both years combined.

Both Years. Preliminary independent samples t-tests comparing the treatment and comparison teachers revealed that the assumption of equality of variances was violated for the student practice ($F(df=1, 73) = 14.75, p=.00$) and feedback ($F(df=1, 73) = 14.10, p=.00$) sections of the lessons as well as for the total instances of explicit instruction ($F(df=1, 73) = 22.26, p=.00$). For the introduction and direct reading sections of the lessons (for which the assumption of equality of variances was not violated), the results of the independent samples t-tests are reported. For the student practice and feedback sections of the lessons, and for the total instances of explicit instruction, the results of Poisson regression analyses are reported.

Significantly more instances of explicit instruction were observed in the treatment school classrooms (a mean of 29.60 (SD=11.17) instances of explicit instruction per lesson across Years 1 and 2) than in the comparison school classrooms (a mean of 17.26 (SD=5.60) instances of explicit instruction per lesson across Years 1 and 2), Wald Chi-Square ($df=1, 73) = 116.43$,

$p=.00$. There was a significant difference between the schools in the mean number of instances of explicit instruction observed during the *direct reading* section, with a mean of 8.93 (SD=4.74) instances of explicit instruction per lesson in the treatment school as compared to a mean of 6.40 (SD=4.20) instances of explicit instruction per lesson in the comparison school (t (df=1, 73) = 2.43, $p=.02$). The mean number of instances of explicit instruction also differed significantly during the *student practice* section, with a mean of 11.15 (SD=7.40) instances of explicit instruction in the treatment school as compared to a mean of 3.66 (SD=3.13) instances of explicit instruction in the comparison school (Wald Chi-Square (df=1, 73) = 123.59, $p=.00$). Finally, there was a significant difference between the treatment school and the comparison school in the number of instances of explicit instruction observed during the *feedback* section, with a mean of 4.30 (SD=4.44) instances of explicit instruction in the treatment school as compared to a mean of 1.54 (SD=2.41) instances of explicit instruction in the comparison school (Wald Chi-Square (df=1, 73) = 43.18, $p=.00$). The treatment school and the comparison school were similar in the instances of explicit instruction they displayed during the *introduction* section of the lesson, with a mean of 5.23 (SD=3.94) instances of explicit instruction per lesson in the treatment school as compared to a mean of 5.66 (SD=4.29) instances of explicit instruction per lesson in the comparison school (t (df=1, 73) = -.46, $p=.65$).

Year 1. Preliminary independent samples t-tests comparing the treatment and comparison teachers in Year 1 revealed that the assumption of equality of variances was violated for the introduction (F (df=1, 40) = 8.13, $p=.01$), student practice (F (df=1, 40) = 11.51, $p=.00$), and feedback (F (df=1, 40) = 5.90, $p=.02$) sections of the lessons as well as for the total instances of explicit instruction (F (df=1, 40) = 12.06, $p=.00$). For the direct reading section of the lessons (for which the assumption of equality of variances was not violated), the results of the

independent samples t-tests are reported. For the introduction, student practice, and feedback sections of the lessons, and for the total instances of explicit instruction, the results of Poisson regression analyses are reported.

The same pattern of results found in both years combined was found in Year 1, with significantly more instances of explicit instruction observed in the treatment school than the comparison school during the direct reading, student practice, and feedback sections and no difference between the teachers during the introduction section. During the *introduction* section, the treatment school averaged 5.25 (SD=2.79) instances of explicit instruction per lesson and the comparison teachers averaged 5.86 (SD=4.65) instances of explicit instruction per lesson, Wald Chi-Square (df=1, 40) = .71, p=.40. The treatment school averaged 10.25 (SD=4.51) instances of explicit instruction per lesson during the *direct reading* section, compared to 6.41 (SD=4.87) instances of explicit instruction per lesson averaged by the comparison school, $t(df=1, 40) = 2.65, p=.01$. During the *student practice* section, the treatment school averaged 9.60 (SD=6.38) instances of explicit instruction per lesson and the comparison school averaged 3.55 (SD=3.05) instances of explicit instruction per lesson, Wald Chi-Square (df=1, 40) = 55.04, p=.00. During the *feedback* section, the treatment school averaged 3.85 (SD=3.88) instances of explicit instruction, compared to 1.41 (SD=2.34) instances of explicit instruction averaged by the comparison school, Wald Chi-Square (df=1, 40) = 22.33, p=.00. Overall, across all four sections of the lessons, the treatment school averaged 28.95 (SD=10.90) instances of explicit instruction per lesson and the comparison school averaged 17.23 (SD=5.25) instances of explicit instruction per lesson, Wald Chi-Square (df=1, 40) = 61.72, p=.00.

Year 2. Preliminary independent samples t-tests comparing the treatment and comparison teachers in Year 2 revealed that the assumption of equality of variances was violated

for the student practice ($F(df=1, 31) = 5.21, p=.03$) and feedback ($F(df=1, 31) = 7.56, p=.01$) sections of the lessons as well as for the total instances of explicit instruction ($F(df=1, 31) = 7.99, p=.01$). For the introduction and direct reading sections of the lessons (for which the assumption of equality of variances was not violated), the results of the independent samples t-tests are reported. For the student practice and feedback sections of the lessons, and for the total instances of explicit instruction, the results of Poisson regression analyses are reported.

In year 2, there was a significant difference between the treatment school and the comparison school in the instances of explicit instruction demonstrated during the student practice section of the lessons, a difference that approached significance during the feedback section of the lessons, and no difference found between the teachers during the introduction and direct reading sections of the lessons. During the *introduction* section, the treatment school averaged 5.20 (SD=4.91) instances of explicit instruction per lesson and the comparison school averaged 5.31 (SD=3.75) instances of explicit instruction per lesson, $t(df=1, 31) = -.07, p=.95$. During the *direct reading* section, the treatment school averaged 7.60 (SD=4.71) instances of explicit instruction per lesson, compared to 6.38 (SD=2.90) instances of explicit instruction averaged by the comparison school, $F(df=1, 31) = .83, p=.41$. During the *student practice* section, the treatment school averaged 12.70 (SD=8.17) instances of explicit instruction per lesson and the comparison school averaged 3.85 (SD=3.39) instances of explicit instruction per lesson, Wald Chi-Square ($df=1, 31$) = 59.61, $p=.00$. During the *feedback* section, the treatment school averaged 4.75 (SD=4.99) instances of explicit instruction per lesson, compared to 1.77 (SD=2.59) instances of explicit instruction averaged by the comparison school, Wald Chi-Square ($df=1, 31$) = 18.06, $p=.00$. Overall, across all four sections of the lessons, the treatment school averaged significantly more instances of explicit instruction per lesson (30.25 (SD=11.67)) than

the comparison school (17.31 (SD=6.36) instances of explicit instruction per lesson, Wald Chi-Square (df=1, 31) = 51.13, p=.00).

Relationship between observations and assessments

Research Question 3: Do the students in classrooms of teachers whose comprehension instruction is more explicit (regardless of the school's program) outperform the students in classrooms of teachers whose comprehension instruction is less explicit?

In order to determine the relationship between the explicitness of the teachers and the students' achievement, two types of analyses were performed: correlations between the explicitness scores and achievement scores and a comparison of the achievement scores of high-explicit and low-explicit teachers. See Appendix D.

Correlations. First, mean explicitness scores were calculated for each teacher. Each lesson that was observed resulted in a number of instances of explicit instruction for each of the four sections of the lesson (introduction, direct reading, student practice, and feedback), and a total for the lesson. These instances of explicit instruction were totaled across all the lessons that were observed for that teacher in a given year, and then divided by the number of lessons that were observed. In this way, each teacher that was observed obtained a mean explicitness score for each section of the lesson (introduction, direct reading, student practice, and feedback), and a total explicitness score, for each year.

Next, adjusted posttest GRADE scores (for each section of the GRADE and the total GRADE score) were calculated for each teacher's students. There were four sections of the GRADE (listening comprehension (K-5), word reading (K-2) or vocabulary (3-5), sentence comprehension (1-5), and passage comprehension (1-5)), so each teacher obtained a total of five GRADE scores, including the total across all four GRADE sections. The analyses of covariance

conducted earlier that treated the proportion correct pretest scores as a covariate and the proportion correct posttest scores as the dependent variable resulted in estimated marginal means for each classroom. These marginal means were posttest scores that had been adjusted by treating the pretest as a covariate. These adjusted posttest scores for each classroom were used to perform subsequent analyses. Correlations were performed between each of the five explicitness scores for each teacher and each of the five adjusted posttest GRADE scores for that teacher's classroom. Tables 33 through 42 in Appendix D display the results of these correlational analyses.

Correlations for all teachers (treatment and comparison schools, all grades). Tables 33 and 34 display the correlations between explicitness scores and posttest scores for all of the observed teachers in Year 1 (Table 33) and Year 2 (Table 34).

Year 1. In Year 1, the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = .13$). Overall, there were 16 positive correlations and 8 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores (the correlation between the total explicitness score and the passage comprehension posttest score was .00), but of the 25 total correlations, only two were significant at the $p < .05$ level, one positive and one negative.

Year 2. In Year 2 (Table 34), the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.02$). Overall, there were 15 positive correlations and 10 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores. Of the 25 correlations, none were significant at the $p < .05$ level.

Correlations for treatment school teachers, all grades. Tables 35 and 36 display the correlations between explicitness scores and posttest scores for the teachers in the treatment school in Year 1 (Table 35) and Year 2 (Table 36).

Year 1. In Year 1, the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = .38$). There were 17 positive correlations and 8 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores. Of the 25 correlations, there were five significant positive correlations at the $p < .05$ level.

Year 2. In Year 2 (Table 36), the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = .18$). There were 19 positive correlations and 6 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores. Only one of the 25 correlations was significant at the $p < .05$ level.

Correlations for comparison school teachers, all grades. Tables 37 and 38 display the correlations between explicitness scores and posttest scores for the teachers in the comparison school in Year 1 (Table 37) and Year 2 (Table 38).

Year 1. In Year 1, the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.18$). There were 9 positive correlations and 16 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores. Of the 25 correlations, there were two significant correlations at the $p < .05$ level, both negative.

Year 2. In Year 2 (Table 38), the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.38$). There were 10 positive correlations and 15 negative correlations between the mean explicitness scores and the adjusted posttest

GRADE scores. Of the 25 correlations, two were significant at the $p < .05$ level, one positive and one negative.

The next set of tables (Tables 39 through 42) break this analysis down by grade level, Grades K-2 separately from Grades 3-5.

Correlations for treatment and comparison school teachers, grades K-2. Tables 39 and 40 display the correlations between explicitness scores and posttest scores for the teachers in Grades K-2 in Year 1 (Table 39) and Year 2 (Table 40).

Year 1. In Year 1, the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.18$). There were 9 positive correlations and 16 negative correlations between the mean explicitness scores and the adjusted posttest GRADE scores. Of the 25 correlations, none was significant.

Year 2. In Year 2 (Table 40), the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.19$). There were no correlations that were significant at the $p < .05$ level.

Correlations for treatment and comparison school teachers, grades 3-5. Tables 41 and 42 display the correlations between explicitness scores and posttest scores for the teachers in Grades 3-5 in Year 1 (Table 41) and Year 2 (Table 42).

Year 1. In Year 1, the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = -.26$). There were no correlations that were significant at the $p < .05$ level.

Year 2. In Year 2 (Table 42), the overall correlation between the total explicitness score and the total GRADE score was not significant ($r = .38$). Of the 25 correlations, there was one positive correlation that was significant at the $p < .05$ level.

Summary of correlations. Overall, these correlations do not indicate that there is a significant relationship between the number of instances of explicit instruction demonstrated by the teachers and the adjusted posttest GRADE scores obtained by the students. There are some trends, however, that are worth identifying. First, it appears that the overall relationship between explicitness and achievement was more positive in the treatment school than in the comparison school. Overall, there were a total of 50 correlations calculated between explicitness and achievement in the treatment school and 50 correlations calculated between explicitness and achievement in the comparison school. Thirty-six of the 50 correlations in the treatment school were positive, whereas 19 of the 50 correlations in the comparison school were positive. The difference between these proportions is significant according to a z-test for proportions, ($N = 50$, $z = 3.60$, $p = .00$). The results do not indicate that there is a difference in the relationship between explicitness and achievement in grades K-2 versus grades 3-5.

Comparison of high-explicit and low-explicit teachers. The second type of analysis that was performed to investigate the relationship between teachers' explicitness and student achievement was a comparison of the mean adjusted posttest GRADE scores for the students of high-explicit teachers and the mean adjusted posttest GRADE scores for the students of low-explicit teachers.

Comparisons of all teachers (treatment and comparison schools, all grades). Year 1. For this analysis, the teachers were ranked according to their total explicitness scores. In Year 1, there were a total of 27 teachers, with total explicitness scores that ranged from 8.00 to 47.00, with a mean across all teachers of 21.44, and a median of 19.00. The teachers were split at the median, with 14 teachers in the high-explicit group and 13 teachers in the low-explicit group. The adjusted posttest GRADE scores of the teachers in the high-explicit group and the adjusted

posttest GRADE scores of the teachers in the low-explicit group were compared using an analysis of variance. See Table 43. There were no significant differences between the high-explicit and the low-explicit teachers on the adjusted posttest scores of the listening comprehension subtest, the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.

Year 2. For the analysis of the Year 2 data, the teachers were again ranked according to their mean total explicitness scores. See Table 44. In Year 2, there were a total of 22 observed teachers with total explicitness scores that ranged from a minimum of 9.00 to a maximum of 49.00, with a mean across all teachers of 26.00 and a median of 24.00. The teachers were split at the median, with 11 teachers in the high-explicit group and 11 teachers in the low-explicit group. The adjusted posttest GRADE scores of the teachers in the high-explicit group and in the low-explicit group were compared using an analysis of variance. There were no significant differences between the high-explicit and the low-explicit teachers on the adjusted posttest scores of the listening comprehension subtest, the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.

In addition to dividing the teachers into high-explicit and low-explicit groups at the median of the total explicitness scores, the explicitness scores of the teachers for Year 2 were plotted (Figure 2) and it was apparent that there were two groups of teachers that differed in their explicitness. There were no teachers with explicitness scores between 29 and 38. Therefore, the total N of 22 was split at that point. There were 16 teachers below this point and 6 teachers (high in explicitness) above this point. The adjusted posttest GRADE scores of these two groups were compared through an analysis of variance. Table 45 presents means, standard deviations, and the results of the ANOVA. There were no significant differences between the 6 teachers in the high

explicit group and the 16 teachers in the average explicit group on the adjusted posttest scores of the listening comprehension subtest, the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.²

Comparisons of treatment school teachers, all grades. These analyses were then conducted on the scores of the treatment teachers only. The treatment school teachers were also ranked separately according to their total explicitness scores and split into two groups at the median. Then, the ANOVA comparisons of the mean adjusted posttest GRADE scores for the students of high-explicit teachers and the mean adjusted posttest GRADE scores for the students of low-explicit teachers were performed for the treatment school separately.

Year 1. In Year 1, there were a total of 14 treatment school teachers, with total explicitness scores that ranged from 11.00 to 47.00, a mean across all teachers of 25.75, and a median of 25.75. The teachers were split at the median, with 7 teachers in the high-explicit group and 7 teachers in the low-explicit group. The adjusted posttest GRADE scores of the teachers in the high-explicit group and the low-explicit group were compared using an analysis of variance. Table 47 presents means, standard deviations, and the results of the ANOVA. The students of the high-explicit teachers scored significantly higher than the students of the low-explicit teachers on the listening comprehension subtest of the GRADE ($F(df = 1, 11) = 5.13, p$

² The mean explicitness scores of the teachers during Year 1 were also plotted (Figure 3), but there were no separations that created distinct groups of teachers differing in their explicitness. Nevertheless, the teachers were split into two groups with one group obtaining total explicitness scores above 25 instances of explicit instruction and the other group obtaining total explicitness scores below 25 instances of explicit instruction. The adjusted posttest GRADE scores of these two groups were compared through an analysis of variance. The results of this ANOVA are found in Table 46. There were no significant differences between the teachers that averaged more than 25 instances of explicit instruction and the teachers that averaged fewer than 25 instances of explicit instruction on the adjusted posttest scores of the listening comprehension subtest, the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.

= .05). There were no significant differences between the mean adjusted posttest scores of the students of the high-explicit and low-explicit teachers on the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.

Year 2. In Year 2, there were a total of 13 treatment school teachers, with total explicitness scores that ranged from 15.00 to 49.00, a mean across all teachers of 31.38, and a median of 28.00. The teachers were split at the median, with 7 teachers in the high-explicit group and 6 teachers in the low-explicit group. The adjusted posttest GRADE scores of the teachers in the high-explicit group and the low-explicit group were compared using an analysis of variance. Table 48 presents means, standard deviations, and the results of the ANOVA. There were no significant differences between the high-explicit and the low-explicit teachers on the adjusted posttest scores of the listening comprehension subtest, the word reading / vocabulary subtest, the sentence comprehension subtest, the passage comprehension subtest, or the total GRADE posttest score.

Relationship between teachers' experience and their explicitness

Research Question 4: Do teachers who are teaching the same program at the same grade level two years in a row demonstrate more explicit instruction during their second year than during their first year?

In order to determine the effect of an additional year of teacher experience on the teachers' explicitness, only teachers who were observed teaching the same grade in Year 1 and Year 2 were considered. The explicitness scores of these teachers in Year 1 were compared to their explicitness scores in Year 2. The treatment teachers and comparison teachers were

considered separately in order to separate the effects of the treatment program and the comparison program. See Appendix E.

Comparisons of treatment school teachers, all grades. In the treatment school, experience was defined as whether the teachers were teaching the treatment program for their first year or for their second year at the same grade level. It was hypothesized that those teachers who taught the treatment program two years in a row would be more explicit in their teaching in their second year than in their first year, because the program is designed to help teachers become more explicit in their instruction. There were seven teachers in the treatment school who were observed teaching the program at the same grade level in both years of the study.

An analysis of variance compared the instances of explicit instruction the seven treatment teachers demonstrated in their first year and the instances of explicit instruction the seven treatment teachers demonstrated in their second year. Table 49 presents means and standard deviations of the instances of explicit instruction, and the results of the ANOVA. There were no significant differences between the mean instances of explicit instruction the treatment teachers demonstrated in their first and second years during the introduction and direct reading sections of the lessons. However, during the student practice section of the lessons, the teachers demonstrated significantly more instances of explicit instruction in their second year than in their first year ($F(df = 1, 12) = 5.63, p = .04$). The teachers also demonstrated significantly more instances of explicit instruction in their second year than in their first year during the feedback section of the lessons ($F(df = 1, 12) = 5.12, p = .04$). Across all the sections of the lessons, the teachers demonstrated significantly more instances of explicit instruction in their second year in their first year ($F(df = 1, 12) = 6.40, p = .03$).

The relationships between experience and explicitness were then investigated separately for the K-2 treatment teachers and the 3-5 treatment teachers.

Comparisons of treatment school teachers, grades K-2. An analysis of variance compared the mean number of instances of explicit instruction demonstrated by the four K-2 treatment teachers in their first year teaching the program and the mean number of instances of explicit instruction demonstrated by those teachers in their second year teaching the program. See Table 50. There were no significant differences between the mean instances of explicit instruction the K-2 treatment teachers demonstrated in their first and second years during the introduction, direct reading, student practice, and feedback sections of the lessons. Across all the sections of the lessons, there was no difference between the mean instances of explicit instruction the K-2 treatment teachers demonstrated in their first and second years.

Comparisons of treatment school teachers, grades 3-5. An analysis of variance compared the mean number of instances of explicit instruction demonstrated by the three 3-5 treatment teachers in their first year teaching the program and the mean number of instances of explicit instruction demonstrated by those teachers in their second year teaching the program. See Table 51. There were no significant differences between the mean instances of explicit instruction the 3-5 treatment teachers demonstrated in their first year and in their second year during the introduction and direct reading sections of the lessons. However, during the student practice section of the lessons, the 3-5 treatment teachers demonstrated significantly more instances of explicit instruction during their second year than during their first year ($F(df = 1, 4) = 9.11, p = .04$). The 3-5 treatment teachers also demonstrated significantly more instances of explicit instruction in their second year than in their first year during the feedback section of the lessons ($F(df = 1, 4) = 21.00, p = .01$). Across all the sections of the lessons, the 3-5 treatment

teachers demonstrated significantly more instances of explicit instruction during their second year than during their first year ($F(df = 1, 4) = 15.87, p = .02$).

Comparisons of comparison school teachers, all grades. In the comparison school, there were six teachers who were observed who taught the program at the same grade level in both years of the study.

An analysis of variance compared the instances of explicit instruction the six comparison teachers demonstrated in their first year and the instances of explicit instruction the comparison teachers demonstrated in their second year. Table 52 presents means and standard deviations of the instances of explicit instruction, and the results of the ANOVA. There were no significant differences between the mean instances of explicit instruction the comparison teachers demonstrated in their first and second years during the introduction, direct reading, student practice, and feedback sections of the lessons. Across all the sections of the lessons, there was no difference between the mean instances of explicit instruction the comparison teachers demonstrated in their first and second years.

The relationships between experience and explicitness were then investigated separately for the K-2 comparison teachers and the 3-5 comparison teachers.

Comparisons of comparison school teachers, grades K-2. An analysis of variance compared the mean number of instances of explicit instruction demonstrated by the three K-2 comparison teachers in their first year teaching the program and the mean number of instances of explicit instruction demonstrated by those teachers in their second year teaching the program. See Table 53. There were no significant differences between the mean instances of explicit instruction the K-2 comparison teachers demonstrated in their first and second years during the introduction, direct reading, student practice, and feedback sections of the lessons. Across all the

sections of the lessons, there was no difference between the mean instances of explicit instruction the K-2 comparison teachers demonstrated in their first and second years.

Comparisons of comparison school teachers, grades 3-5. An analysis of variance compared the mean number of instances of explicit instruction demonstrated by the three 3-5 comparison teachers in their first year teaching the program and the mean number of instances of explicit instruction demonstrated by those teachers in their second year teaching the program. See Table 54. There were no significant differences between the mean instances of explicit instruction the 3-5 comparison teachers demonstrated in their first year and in their second year during the introduction, student practice, and feedback sections of the lessons. However, during the direct reading section of the lessons, the 3-5 comparison teachers demonstrated significantly more instances of explicit instruction during their second year than during their first year ($F(df = 1, 4) = 9.94, p = .03$). Across all the sections of the lessons, the 3-5 comparison teachers demonstrated significantly more instances of explicit instruction during their second year than during their first year ($F(df = 1, 4) = 7.56, p = .05$).

Relationship between teachers' experience and student achievement

Research Question 5: Do the students of second-year teachers demonstrate improved performance on comprehension measures as compared to students in the first-year teachers' classrooms?

In order to determine the effect of an additional year of teacher experience on students' performance, the students were separated into two groups based on their teachers' experience. The treatment teachers and comparison teachers were considered separately. See Appendix F.

Comparisons of treatment school teachers, all grades. In the treatment school, the students in the classrooms of the seven first-year teachers were compared to the students in the

classrooms of the seven second-year teachers. (The assignment of the treatment teachers to the first-year and second-year groups is described on page 84).

An analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the seven first-year treatment teachers and the students in the classrooms of the seven second-year treatment teachers. See Table 55. There were no significant differences in the scores of the students of the first-year teachers and the students of the second-year teachers on three of the four sections of the GRADE, the listening comprehension, word reading / vocabulary, and sentence comprehension subtests. However, the students of the seven second-year teachers significantly outperformed the students of the seven first-year teachers on the passage comprehension subtest ($F(df = 1, 9) = 5.56, p = .04$). There was no difference between the groups on the total GRADE score.

The relationship between experience and achievement was then investigated separately for the K-2 treatment teachers and the 3-5 treatment teachers.

Comparisons of treatment school teachers, grades K-2. An analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the K-2 treatment teachers teaching the program for their first year and the students in the classrooms of the K-2 treatment teachers teaching the program for their second year. See Table 56. There were no significant differences in the scores of the students of the K-2 first-year teachers and the students of the K-2 second-year teachers on the listening comprehension, word reading, or sentence comprehension subtests. However, the students of the K-2 second-year teachers significantly outperformed the students of the K-2 first-year teachers on the passage comprehension subtest ($F(df = 1, 3) = 26.80, p = .01$). There was no difference between the groups on the total GRADE score.

Comparisons of treatment school teachers, grades 3-5. Finally, an analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the treatment teachers in grades 3-5 teaching the program for their first year and the students in the classrooms of the treatment teachers in grades 3-5 teaching the program for their second year. See Table 57. There were no significant differences in the scores of the students of the 3-5 first-year teachers and the students of the 3-5 second-year teachers on the listening comprehension, word reading, or sentence comprehension subtests. However, the students of the 3-5 second-year teachers significantly outperformed the students of the 3-5 first-year teachers on the passage comprehension subtest ($F(df = 1, 4) = 11.09, p = .03$). There was no difference between the groups on the total GRADE score.

Comparisons of comparison school teachers, all grades. In the comparison school, the students in the classrooms of the six teachers who were observed in Year 1 were compared to the students in the classroom of those six teachers in Year 2.

An analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the six new comparison teachers and the students in the classrooms of the six experienced comparison teachers. See Table 58. There were no significant differences between the scores of the students of the new teachers and the scores of the students of the experienced teachers on any of the four sections of the GRADE, or on the total GRADE score.

The relationship between experience and achievement was then investigated separately for the K-2 comparison teachers and the 3-5 comparison teachers.

Comparisons of comparison school teachers, grades K-2. An analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the new K-2 comparison teachers and the students in the classrooms of the experienced K-2

comparison teachers. See Table 59. There were no significant differences between the scores of the students of the new K-2 comparison teachers and the scores of the students of the experienced K-2 comparison teachers on any of the four sections of the GRADE, or on the total GRADE score.

Comparisons of comparison school teachers, grades 3-5. Finally, an analysis of variance compared the mean adjusted GRADE posttest scores between the students in the classrooms of the new comparison teachers in grades 3-5 and the students in the classrooms of the experienced comparison teachers in grades 3-5. See Table 60. There were no significant differences between the scores of the students of the new 3-5 comparison teachers and the scores of the students of the experienced 3-5 comparison teachers on any of the four sections of the GRADE, or on the total GRADE score.

Comparison of experienced treatment teachers and experienced comparison teachers

Research Question 6: Do the students of second-year teachers in the treatment school demonstrate improved performance on comprehension measures as compared to students of second-year teachers in the comparison school?

In order to equate the experience of the treatment and comparison teachers as much as possible, these analyses only examined the results of the students in Year 2 whose teachers had also taught the same program at the same grade level in Year 1. In the treatment school, there were 14 teachers in Year 2 who had taught the treatment program at the same grade level in Year 1. There were a total of 274 students in the classrooms of these 14 experienced teachers who completed both the pretest and posttest in Year 2. In the comparison school, there were 13 teachers in Year 2 who had taught the comparison program at the same grade level in Year 1.

There were a total of 241 students in the classrooms of these 13 experienced teachers who completed both the pretest and posttest in Year 2.

Analyses of covariance were conducted to test for differences between the treatment and comparison schools on each assessment on the spring posttests in Year 2. For each assessment, the fall test was used as the covariate and the spring test as the dependent measure. Each subtest of the GRADE (Word Reading, Vocabulary, Listening Comprehension, Sentence Comprehension, and Paragraph Comprehension) was analyzed separately. See Appendix G.

Comparison of experienced treatment school teachers and experienced comparison school teachers, all grades. The analyses of covariance were first calculated with classrooms from all grades combined. See Table 61. The comparison school students in the classrooms of experienced teachers (Adjusted Proportion Correct=.79, SE=.01) significantly outperformed the treatment school students in the classrooms of experienced teachers (Adjusted Proportion Correct=.74, SE=.01) on the sentence comprehension subtest of the GRADE, $F(df=1, 393) = 7.04, p=.01$. The treatment school students in the classrooms of experienced teachers (Adjusted Proportion Correct=.62, SE=.01) significantly outperformed the comparison school students in the classrooms of experienced teachers (Adjusted Proportion Correct=0.57, SE=.01) on the passage comprehension subtest of the GRADE, $F(df=1, 391) = 7.10, p=.01$. There were no differences between the treatment school students and the comparison school students on the listening comprehension or word reading subtests or on the total GRADE score.

The comparison between the experienced treatment teachers and the experienced comparison teachers was then investigated separately for the K-2 teachers and the 3-5 teachers.

Comparison of experienced treatment school teachers and experienced comparison school teachers, K-2. Analyses of covariance were calculated with the scores of students of the

K-2 teachers from both schools. See Table 62. The treatment school students in the classrooms of experienced K-2 teachers (Adjusted Proportion Correct=.62, SE=.02) significantly outperformed the comparison school students in the classrooms of experienced K-2 teachers (Adjusted Proportion Correct=.54, SE=.02) on the passage comprehension subtest of the GRADE, $F(df=1, 139) = 5.36, p=.02$. There were no differences between the treatment school K-2 students and the comparison school K-2 students on the listening comprehension, word reading, or sentence comprehension subtests or on the total GRADE score.

Comparison of experienced treatment school teachers and experienced comparison school teachers, 3-5. Analyses of covariance were calculated with the scores of students of the 3-5 teachers from both schools. See Table 63. The comparison school students in the classrooms of experienced 3-5 teachers (Adjusted Proportion Correct=.84, SE=.01) significantly outperformed the treatment school students in the classrooms of experienced 3-5 teachers (Adjusted Proportion Correct=.81, SE=.01) on the listening comprehension subtest of the GRADE, $F(df=1, 248) = 4.31, p=.04$. The comparison school students in the classrooms of experienced 3-5 teachers (Adjusted Proportion Correct=.80, SE=.01) also significantly outperformed the treatment school students in the classrooms of experienced 3-5 teachers (Adjusted Proportion Correct=.76, SE=.01) on the sentence comprehension subtest of the GRADE, $F(df=1, 247) = 5.29, p=.02$. There were no differences between the treatment school 3-5 students and the comparison school 3-5 students on the vocabulary or passage comprehension subtests or on the total GRADE score.

Comparison of experienced treatment students and experienced comparison students

Research Question 7: Do the students who have been enrolled in the treatment school for two consecutive years outperform the students who have been enrolled in the comparison school for two consecutive years?

In order to determine the effect of two consecutive years of explicit comprehension instruction, the students who had been enrolled in the treatment school for two consecutive years were compared to the students who had been enrolled in the comparison school for two consecutive years. Only the students that completed the pretests and posttests for two consecutive years in either the treatment or the comparison school were included in this analysis: 213 in the treatment school and 175 in the comparison school. See Appendix H. Table 64 presents the mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance, which was conducted to test for differences in the GRADE scores of the treatment students and the comparison students. For each subtest of the GRADE, the Year 1 pretest was used as the covariate and the Year 2 posttest was used as the dependent measure. There were no significant differences between the adjusted posttest scores of the two-year treatment students and the two-year comparison students on any of the GRADE subtests (listening comprehension, word reading / vocabulary, sentence comprehension, or passage comprehension) or on the total GRADE score.

The effect of two consecutive years of explicit comprehension instruction was then investigated separately for grades K-2 and grades 3-5, as analyses of the previous research questions had indicated that there may be differences when considering the early and later elementary grades separately. The first analysis included students who were in kindergarten and first grade in Year One (and, therefore, in first grade and second grade in Year Two). There

were 62 K-2 students in the treatment school and 61 K-2 students in the comparison school who completed the assessments in both Year 1 and Year 2 of the study. The mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance are presented in Table 65. There were no significant differences between the adjusted posttest scores of the two-year K-2 treatment students and the two-year K-2 comparison students on the listening comprehension or sentence comprehension GRADE subtests, or on the total GRADE score. However, the two-year K-2 treatment students significantly outperformed the two-year K-2 comparison students on the word reading GRADE subtest ($F(df = 1, 101) = 5.22, p = .02$). The two-year K-2 treatment students also obtained a higher mean score than the two-year K-2 comparison students on the passage comprehension GRADE subtest, with the difference approaching significance ($F(df = 1, 51) = 3.42, p = .07$).

The same analysis was conducted on the data for the later elementary grades. An analysis of covariance was conducted to test for differences in the GRADE scores of the two-year 3-5 treatment students versus the two-year 3-5 comparison students. Table 66 presents the mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance. These analyses included students that were in third grade and fourth grade in Year One (and, therefore, in fourth grade and fifth grade in Year Two). There were 111 third through fifth grade students in the treatment school and 81 third through fifth grade students in the comparison school. There were no significant differences between the adjusted posttest scores of the two-year 3-5 treatment students and the two-year 3-5 comparison students on any of the GRADE subtests or on the total GRADE score.

The effect of two consecutive years of explicit comprehension instruction was then investigated separately for each grade. A summary of the results for the total GRADE score is

presented in Table 67. Students who were in Grade 2 in Year 2 were the only ones that showed a significant difference between the treatment and the comparison schools. Thus the data at this grade level were subjected to an analysis of each portion of the GRADE separately. This analysis is presented in Table 68. There were forty-two 1-2 students in the treatment school and thirty 1-2 students in the comparison school. The two-year 1-2 treatment students significantly outperformed the two-year 1-2 comparison students on the listening comprehension ($F(df = 1, 62) = 4.90, p = .03$) and word reading ($F(df = 1, 64) = 10.59, p = .00$) GRADE subtests, and on the total GRADE score ($F(df = 1, 70) = 9.51, p = .00$). The two-year 1-2 treatment school students also obtained a higher score than the two-year 1-2 comparison school students on the passage comprehension GRADE subtest, and this difference approached significance ($F(df = 1, 51) = 3.73, p = .06$). There was no significant difference between the scores of the two-year 1-2 treatment students and the two-year 1-2 comparison students on the sentence comprehension GRADE subtest.

Appendix I presents similar analyses for the other grades, none of which showed an significant differences. In summary, the results of the analyses designed to address Research Question 7 found very few differences between the scores of students who had been exposed to two years of explicit comprehension instruction in the treatment program and the scores of students who had been exposed to two years of traditional comprehension instruction in the comparison program. The only differences between the treatment students and the comparison students were found in those students who had been in first grade in Year One and second grade in Year Two. The first and second grade treatment students scored significantly higher than the first and second grade comparison students on the listening comprehension and word reading subtests of the GRADE, and had a significantly higher total GRADE score. In addition, the

difference between the schools on the passage comprehension subtest of the GRADE approached significance, with the treatment school again outperforming the comparison school.

Comparison of experienced treatment students with experienced teachers and experienced comparison students with experienced teachers

Research Question 8: Do the students who have been enrolled for two consecutive years and have second year teachers in the treatment school outperform the students who have been enrolled for two consecutive years and have second year teachers in the comparison school?

The final research question examined the difference between the experienced students with experienced teachers in the treatment school and the experienced students with experienced teachers in the comparison school. Only the students that completed the pretests and posttests for two consecutive years in either the treatment or the comparison school and whose teachers in Year 2 were experienced with the comprehension program they taught were included in this analysis: 159 in the treatment school and 121 in the comparison school. See Appendix J. Table 73 presents the mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance. For each subtest of the GRADE, the Year 1 pretest was used as the covariate and the Year 2 posttest was used as the dependent measure. The two-year treatment students with experienced teachers (Adjusted proportion correct = .73, SE = .01) significantly outperformed the two-year comparison students with experienced teachers (Adjusted proportion correct = .68, SE = .02) on the word reading / vocabulary subtest of the GRADE, $F(df=1, 239) = 4.98, p=.03$. There were no significant differences between the adjusted posttest scores of the two-year treatment students with experienced teachers and the two-year comparison students with experienced teachers on the listening comprehension,

sentence comprehension, or passage comprehension subtests of the GRADE, or on the total GRADE score.

The effect of two consecutive years of explicit comprehension instruction was then investigated separately for grades K-2 and grades 3-5, as analyses of the previous research questions had indicated that there may be differences when considering the early and later elementary grades separately. The first analysis included students who were in kindergarten and first grade in Year One (and, therefore, in first grade and second grade in Year Two). There were 62 K-2 students in the treatment school and 37 K-2 students in the comparison school who completed the assessments in both Year 1 and Year 2 of the study. The mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance are presented in Table 74. There were no significant differences between the adjusted posttest scores of the two-year K-2 treatment students with experienced teachers and the two-year K-2 comparison students with experienced teachers on any of the GRADE subtests (listening comprehension, word reading / vocabulary, sentence comprehension, or passage comprehension) or on the total GRADE score.

The same analysis was conducted on the data for the later elementary grades. An analysis of covariance was conducted to test for differences in the GRADE scores of the two-year 3-5 treatment students with experienced teachers versus the two-year 3-5 comparison students with experienced teachers. Table 75 presents the mean adjusted Year 2 posttest score and standard deviation for each school, and the results of the analysis of covariance. These analyses included students that were in third grade and fourth grade in Year One (and, therefore, in fourth grade and fifth grade in Year Two). There were 69 third through fifth grade students in the treatment school and 51 third through fifth grade students in the comparison school. There

were no significant differences between the adjusted posttest scores of the two-year 3-5 treatment students with experienced teachers and the two-year 3-5 comparison students with experienced teachers on the vocabulary, sentence comprehension, or passage comprehension GRADE subtests, or on the total GRADE score. However, the two-year 3-5 comparison students with experienced teachers (Adjusted proportion correct = .83, SE = .02) significantly outperformed the two-year 3-5 treatment students with experienced teachers (Adjusted proportion correct = .79, SE = .01) on the listening comprehension GRADE subtest ($F(df = 1, 107) = 5.77, p = .02$).

CHAPTER V

Discussion

This study investigated the effect of the explicit teaching of reading comprehension skills and strategies, through the Urban Education Exchange Concepts of Comprehension program[©], on the reading and listening comprehension of students in grades K-5. The implementation (measured through observations) and results (measured through assessments) of the UEE comprehension program at one school were compared to the implementation and results of a widely-used comprehension program at a demographically similar school. The population of both schools included a substantial proportion of students at risk for reading failure.

Research Question 1: Does instruction in reading comprehension that focuses on the explicit teaching of reading comprehension skills and strategies (Program T) improve performance on standardized and researcher-developed reading comprehension tests more than a traditional reading comprehension program (Program C)?

The evidence from this study that addresses this question is inconclusive. The treatment school, which utilized the explicit comprehension curriculum, significantly outperformed the comparison school on 37 of the 108 assessments that were administered across the two years (compared to only 3 assessments on which the comparison school outperformed the treatment school) on independent samples t-tests comparing mean scores. The difference between these proportions is significant according to a z-test for proportions, ($N = 108$, $z = 6.37$, $p = .00$). However, more than half (19) of these significant differences were found in either the fall of Year One or the fall of Year Two. These data can be found in Appendix B. The superiority of the treatment school on the fall assessments suggests that the students in the treatment school began each school year with a significant advantage over the comparison school. On the other

hand, this advantage may have been the result of the instruction, different in each school, which occurred before the initial assessments each year. In Year One, the students in both schools had been receiving either the treatment or the comparison comprehension instruction for two to three months before the initial assessments were administered. In Year Two, a majority of the students in both schools were returning students and, therefore, had been exposed to the instructional program specific to that school during Year One.

Analysis of covariance is a statistical technique to correct for different baselines; it adjusts scores to “equate” performance at the beginning of treatment. A total of 52 analyses of covariance (26 each in Year 1 and Year 2) were conducted that treated the fall test as the covariate and the spring test as the dependent measure, in order to ascertain the differences between the schools that had occurred as a result of the instruction during that year. See Appendix B, Table 4. In Year 2, these analyses resulted in 5 significant differences out of 26 comparisons in which the treatment school outperformed the comparison school and 3 significant differences out of 26 comparisons in which the comparison school outperformed the treatment school. The difference in these proportions is not significant according to a z-test for proportions ($N = 26, z = .76, p = .45$). The assessments on which the treatment school obtained higher mean scores than the comparison school were concentrated in kindergarten, first, and second grade; in these grades, in Year 2, there were a total of 11 comparisons and the treatment school outperformed the comparison school 4 out of 11 times, whereas none of the 11 comparisons favored the comparison school. The difference between these proportions is significant according to a z-test for proportions, ($N = 11, z = 2.39, p = .02$). The results favoring the comparison school were primarily in the third, fourth and fifth grades; in these grades, in Year 2, there were a total of 15 comparisons and the comparison school outperformed the treatment

school on 3 of the 15 comparisons, whereas 1 of the 15 comparisons favored the treatment school. The difference in these proportions is not significant according to a z-test for proportions ($N = 15, z = 1.06, p = .29$). Taken as a whole, the results of this study suggest that the treatment school comprehension program is comparable to the comparison school comprehension program in terms of student achievement.

However, the pattern of results suggests that there are differences between the schools when the lower grades are considered separately from the higher grades. The differences found between the grade levels are not surprising. What distinguishes Urban Education Exchange from other reading comprehension programs is its focus on explicit reading comprehension instruction in grades K-2. Many widely-available and utilized programs do not even begin comprehension instruction until the third grade, reflecting their approach that is based in the belief that children must “learn to read” before they can “read to learn.” By beginning comprehension instruction in kindergarten, Urban Education Exchange seeks to help children understand what they read from the very beginning of instruction. Thus, one explanation for the differing results in the early and later grades is that explicit instruction in comprehension strategies may be more important and more effective in the early primary grades and less important in the upper elementary grades. Once the students reach the later grades, they may require less explicit instruction in comprehension strategies, thus negating the advantage of an explicit comprehension program for these children.

Another explanation could be that the comparison school utilized different curricula in the early and later grades: the Mondo Bookshop program in grades K-3 and the Teachers College Reading and Writing Workshop program in grades 4-5. The differences in these two programs may have contributed to the differences found between the treatment school and the comparison

school. It could be that the Teachers College Reading and Writing Workshop program is a more effective program than the Mondo Bookshop program. However, it is interesting to note that the comparison school had previously been using the TC Reading and Writing Workshop program in grades K-3, but adopted the Mondo Bookshop program because the teachers and administration felt that the TC program was ineffective for the younger grades. They did feel, however, that the TC Reading and Writing Workshop program was effective for their fourth and fifth grade students.

Because this study covered only two years of instruction, we cannot know how the fourth and fifth grade students in the treatment school would have responded if they had received the explicit UEE comprehension instruction when they were in grades K-3. If explicit instruction is indeed more important and effective in the early elementary grades, then the fourth and fifth grade students in the treatment school were at a disadvantage because they did not receive this type of instruction in grades K-3. The comparison school adopted the Mondo Bookshop program beginning in the 2006-2007 school year, so that the majority of the comparison school students in the study had received the Mondo Bookshop program. Observations revealed that this program contains some limited amount of explicit comprehension instruction, though not to the extent of the UEE program. Thus, the students who were in fourth grade in Year 1 and fifth grade in Year 2 had received some explicit comprehension instruction via the Mondo Bookshop program in second and third grade. This early explicit instruction may have provided a strong base that continued to help them in their reading comprehension in later years.

Another factor to take into consideration is the teachers' familiarity with the reading comprehension program they were using. The fourth and fifth grade teachers in the comparison school had been using the Teachers College Reading and Writing Workshop program for several

years before the start of this study (they began using this program during the 2003-2004 school year). This familiarity with the program may have allowed them to be more effective teachers than if they had only started using it at the beginning of the study. The teachers in the treatment school, on the other hand, were using the UEE comprehension program for the first time during Year 1. Research has suggested that teachers are more effective in their instruction once they have been using the same program for several years (Fullan, 2001; Gersten, Carnine, Zoref, & Cronin, 1986; Hall & Hord, 2001; James-Burdumy et. al., 2012).

Thus, the findings concerning the first research question are, as would be expected, inconclusive as to whether the reading comprehension performance of the students who received the UEE Comprehension Program is different from that of students who received the Mondo Bookshop and/or the Teachers College Programs. However, a case certainly can be made that the UEE Comprehension Program is comparable in effectiveness to the other programs. Such a conclusion is not without meaning: the Mondo Bookshop Program and the Teachers College Program are both well-established instructional programs with large followings. Over the last few years the field has seen how difficult it is to find high-quality empirical studies of instructional programs in real-life school settings that show significant differences in favor of any particular program over others. One needs larger studies that involve many schools to demonstrate significant differences on end-of-year achievement tests (Dole, Duffy, Roehler, & Pearson, 1991; Pressley, 1998; Slavin et al., 2009). Typically, consensus by experts, not empirical evidence, is usually the basis for judging the value of an instructional program.

Research Question 2: Did Program T result in more instances of explicit instruction during reading comprehension lessons than Program C, as measured by classroom observations?

Observations of the comprehension lessons in both the treatment and comparison schools were conducted, and the instances of explicit instruction that were observed during the introduction, direct reading, student practice, and feedback sections of the lessons were tallied. The results of these observations can be found in Appendix C, Tables 30-32. Treatment school teachers demonstrated significantly more instances of explicit instruction (a mean of 29.6 instances of explicit instruction across Years 1 and 2) than the comparison school (a mean of 17.3 instances of explicit instruction across Years 1 and 2). These differences were observed during the direct reading, student practice, and explicit feedback sections of the lesson. These results indicate that teachers in the treatment school provided more explicit instruction during 3 of the 4 sections of reading comprehension lessons than their counterparts in the comparison school.

The UEE program trains teachers to think aloud and model the Concepts of Comprehension© while reading aloud, and it includes suggestions for modeling and thinking aloud in the lesson plans. It is to be expected, therefore, that teachers using the UEE curriculum demonstrated more instances of explicit instruction during direct reading than teachers in the comparison school. This demonstrates fidelity to the UEE curriculum. In fact, the first three sections of the lesson (introduction, direct reading, and practice) were largely determined by the program itself. The last section, feedback, however, was largely dependent on the teacher. The fact that teachers in the treatment school displayed close to three times as many instances of explicit feedback during the feedback section as the teachers in the comparison school suggests that the explicit instruction modeled in the rest of the lesson had shaped the way that these teachers provide feedback. These data suggest that the UEE curriculum may have helped to make the treatment school teachers more explicit in their instruction beyond what is found in the

lesson plans. Given the consensus that explicit and structured reading comprehension instruction is effective (National Reading Panel, 2000) and the dearth of empirical evidence that an explicit comprehension program actually results in more explicit instruction, the finding that teachers demonstrate more explicit instruction within a regular classroom situation when using a specific instructional program is an important one.

Research Question 3: Did the students in classrooms of teachers whose comprehension instruction was more explicit (regardless of the school's program) outperform the students in classrooms of teachers whose comprehension instruction was less explicit?

This question was addressed in two ways, first with a correlational analysis. The explicitness scores given to the teachers during the observations were correlated with the students' scores on the standardized assessments. There was no relationship between the explicitness of the teachers and their students' performance. In Year 1, the correlation between the total explicitness score and the total GRADE score was .13. In Year 2, this same correlation was -.02. Looking at correlations of explicitness on specific sections of the lessons and outcomes on specific sections of the GRADE did not support that there was a relationship.

In addition to the correlational analysis, the teachers were split into two groups based on the median explicitness score, and then the relationship between these groups and student outcomes was examined using ANOVA. This analysis, too, did not reveal any differences in performance between the students of the more explicit teachers and the students of the less explicit teachers. In Year 1, the adjusted proportion correct on the GRADE for students in classrooms whose teachers were above the median was .72, for students in classrooms whose teachers were below the median it was also .72. In Year 2, the proportions correct for these groups of students were .73 and .74, respectively. Additional analyses that split the teachers into

groups of unequal size that maximized the groups' difference in explicitness scores did not reveal any significant differences as a function of explicitness.

Taken together, the correlational analyses and the comparisons of explicit and less-explicit teachers did not indicate a relationship between teacher explicitness and student achievement. One reason for the lack of a significant relationship between teacher explicitness and student achievement could be the relatively small number of teachers involved. Perhaps a larger sample, which would allow examinations of specific sub-groups, such as teachers at different grade levels and teachers that used different instructional programs, would show a significant relationship. Another possibility is that the way in which teacher explicitness and/or student achievement were operationalized in this study did not allow for any actual relationship to appear. The observations may have been conducted too infrequently to obtain a valid measure of teacher explicitness. The observational protocol may have lacked reliability. Different observational protocols and comprehension assessments could be used to test this idea. Of course it may be that there is, in fact, no relationship between teacher explicitness and student achievement. Because of the widespread idea (without specific research evidence) that teacher explicitness is valuable, further questions were asked.

Research Question 4: Did teachers who taught the same program at the same grade level two years in a row demonstrate more explicit instruction during their second year than teachers who were teaching the program for their first year?

The second-year treatment teachers were more explicit during the student practice and feedback sections of the lessons, as well as in the total number of instances of explicit instruction observed across the whole lesson. Thus, the second-year treatment teachers demonstrated more explicit instruction than the first-year treatment teachers. The second-year comparison teachers,

however, did not demonstrate more explicit instruction during any of the sections of the lessons that were observed, or across the whole lesson than the first-year comparison teachers.

The fact that the treatment teachers were more explicit during their second year using the explicit program supports our earlier observation and previous research (Fullan, 2001; Gersten, Carnine, Zoref, & Cronin, 1986; Hall & Hord, 2001; James-Burdumy et. al., 2010) that teachers need time to get used to and master a reading comprehension (or any educational) program. The treatment program focuses on the explicit instruction of comprehension strategies, and the teachers that have an extra year of experience teaching this program are more explicit in their instruction. They have demonstrated increased mastery of this goal of the comprehension program. The comparison program, however, has less of a focus on the explicit teaching of comprehension strategies and, therefore, it is not surprising that an extra year of experience using this program does not result in more explicit instruction.

The fact that the differences in the first-year and second-year treatment teachers were found primarily in the student practice and feedback sections of the lessons could suggest that explicit instruction during these sections of the lesson is more dependent on the teachers' experience teaching the program. The introduction and direct reading sections of the lesson are more scripted and, therefore, the amount of explicit instruction during these sections may have less opportunity to increase with more experience teaching the program. Increased familiarity with the scripted sections of the lesson may allow the teachers to focus more of their attention on the other sections (the student practice and feedback sections), and to increase the amount of explicit instruction they provide for those sections. An alternative explanation is that, if the teachers were observed for longer than two years, they might also show more explicitness during the introduction and direct reading sections of the lessons.

Research Question 5: Did the students of second-year teachers demonstrate improved performance on comprehension measures as compared to students in the first-year teachers' classrooms?

The students in the second-year treatment teachers' classrooms scored significantly higher on the passage comprehension GRADE subtest than the students in the first-year treatment teachers' classrooms. There were no significant differences between the scores of the students in first-year treatment teachers' classrooms and the scores of the students in second-year treatment teachers' classrooms on any of the other GRADE subtests. There were also no differences between the scores of the students in first-year comparison teachers' classrooms and the scores of the students in second-year comparison teachers' classrooms on any of the GRADE subtests.

One could argue that, as passage comprehension is the most similar to reading tasks the students are likely to perform in school, the passage comprehension subtest of the GRADE is the most important subtest. The finding that the students of teachers with more experience teaching the treatment program improved their passage comprehension more than the students of teachers with no experience could be seen as evidence that experience teaching the treatment program is related to improved reading comprehension. The goal of the explicit teaching of reading comprehension strategies, after all, is to improve performance in reading comprehension. Replication of this potentially important finding is necessary, however, before any definitive conclusion can be reached about the relationship between teacher experience and student achievement.

Research Question 6: Did the students of second-year teachers in the treatment school demonstrate improved performance on comprehension measures as compared to students of second-year teachers in the comparison school?

The students of second-year teachers in the comparison school scored significantly higher on the sentence comprehension GRADE subtest than the students in the second-year treatment teachers' classrooms. The students of second-year teachers in the treatment school scored significantly higher on the passage comprehension GRADE subtest than the students in the second-year comparison teachers' classrooms. There were no significant differences between the scores of the students in second-year treatment teachers' classrooms and the scores of the students in second-year comparison teachers' classrooms on any of the other GRADE subtests.

In grades K-2, the students of second-year teachers in the treatment school scored significantly higher on the passage comprehension GRADE subtest than the students in the second-year comparison teachers' classrooms. There were no significant differences on any of the other GRADE subtests. In grades 3-5, the students of second-year teachers in the comparison school scored significantly higher on the listening comprehension and the sentence comprehension GRADE subtests than the students in the second-year treatment teachers' classrooms. There were no significant differences on any of the other GRADE subtests.

When examined together, the results of these analyses mirror previous results that suggest that the treatment school students outperformed the comparison school students in grades K-2 (particularly on tests of passage comprehension) and that the comparison school students outperformed the treatment school students in grades 3-5 (particularly on tests of sentence comprehension). As discussed previously, the passage comprehension test most closely approximates the actual act of reading comprehension. The explicit comprehension program that

the treatment teachers used includes teacher modeling and student practice of comprehension strategies while reading authentic fiction and nonfiction texts. This type of instruction can also be found in other comprehension programs, of course, but it often does not begin until the third grade. With an additional year of experience using the explicit program, the treatment teachers may have been more comfortable modeling and practicing the comprehension strategies with the students and thus provided more effective comprehension instruction. The performance of the K-2 students with experienced teachers on the passage comprehension test suggests that the explicit comprehension instruction with authentic texts that they received improved their reading comprehension.

Research Question 7: Did the students who were enrolled in the treatment school for two consecutive years outperform the students who were enrolled in the comparison school for two consecutive years?

Significant differences between the two-year treatment students and the two-year comparison students appeared only among those students that were first graders in Year One and second graders in Year Two. In this cohort, the treatment school students significantly outperformed the comparison school students on the listening comprehension and word reading subtests, and on the total GRADE score, and there was a trend toward significance on the passage comprehension subtest.

There are a number of possible explanations for these differences. One explanation is that this group of students in the treatment school was, for whatever reason, particularly strong. However, this explanation is unlikely because no differences were found between the scores of the first grade treatment school students and the first grade comparison school students on any of the GRADE subtests or the total GRADE score on the Year One pretest.

Another explanation is that the first and second grade teachers in the treatment school were, as a whole, superior teachers to the first and second grade teachers in the comparison school. The first grade treatment teachers did demonstrate significantly more explicit instruction than the first grade comparison teachers in Year 1. The second grade treatment teachers also demonstrated significantly more explicit instruction than the second grade comparison teachers in Year 2. However, these differences in explicit instruction between the treatment and comparison school teachers were also found in kindergarten in Year 1, in second grade in Year 1, and in third grade in Year 2. The treatment cohorts that were in the classrooms of these teachers, however, did not significantly outperform their comparison cohorts.

It is difficult to separate the effects of teachers from the effects of the program that the teachers use for instruction. It should be noted, however, that the most unique feature of the UEE comprehension program is its focus on comprehension instruction in the early primary grades: kindergarten, first grade, and second grade. Thus, another explanation for the differences between this cohort of treatment and comparison school students is that the explicit comprehension program was more effective in the early elementary grades than the traditional comprehension program for students who have had sufficient experience with the program. Further support for this explanation comes from the Year One first grade sentence comprehension subtest, the Year Two first grade passage comprehension subtest, the Year One second grade listening comprehension subtest, and the Year Two second grade word reading subtest, all of which significantly favored the treatment school. While these results may not be overwhelming enough to declare that the treatment program is clearly more effective in these grades, they do provide evidence that further research would be worthwhile.

But why were more differences in favor of the treatment school were not found in kindergarten? One possibility is that kindergarten children are, in fact, too young to benefit from explicit comprehension instruction. Another explanation is that the listening comprehension and word reading assessments that were administered to the kindergarten students did not accurately assess their differences in comprehension ability. Again, further research would be useful in examining these alternative explanations.

Research Question 8: Do the students who have been enrolled for two consecutive years and have second year teachers in the treatment school outperform the students who have been enrolled for two consecutive years and have second year teachers in the comparison school?

The comparison of the experienced students with experienced teachers in the treatment school versus the experienced students with experienced teachers in the comparison school revealed a significant difference between the schools on the word reading / vocabulary subtests, with the treatment school outperforming the comparison school. When investigating the word reading subtest (administered to grades K-2) and the vocabulary subtest (administered to grades 3-5) separately, however, there were no differences between the schools. The comparison school outperformed the treatment school on the listening comprehension subtest in grades 3-5.

The lack of significant differences found when comparing experienced students with experienced teachers in both schools (particularly in grades K-2) is somewhat surprising, given the differences found when comparing experienced teachers and experienced students separately. It may be an issue of small sample size, as only 99 students were included in the K-2 sample and 120 students were included in the 3-5 sample. Additional studies utilizing a larger sample of

students and teachers that are experienced in using an explicit comprehension program would be worthwhile.

Limitations and Further Analyses

In this section, we discuss possible reasons for the inconclusive results with respect to the general hypothesis of whether explicit teaching of reading comprehension skills and strategies results in improved listening and reading comprehension, and we suggest further analyses that may help to clarify the relationship between explicit teaching of comprehension skills and strategies and reading comprehension performance.

Limitations of the present study.

This study, designed as an initial investigation of the effectiveness of the Urban Education Exchange Comprehension Program, involved a relatively small number of participants. Small-scale studies mean low statistical power (i.e., less opportunity to detect significant differences between the treatment and comparison groups). Moreover, a single school in its entirety was assigned to each experimental condition, having volunteered to participate in that particular condition.

The fall tests in both Year 1 and Year 2 were administered late as a result of complications at the schools. Importantly, however, each assessment was administered within the same week at both schools. Thus, although the fall tests may not provide true pretest scores for the schools, they could still be used effectively as covariates to control for any preexisting differences in the schools.

An additional limitation was that all the observations of the reading comprehension lessons in the treatment and comparison schools were conducted by one researcher. This decision was made as a result of a lack of time and funding to hire and train an additional observer. Because of this

limitation, the reliability of the observation protocol was not measured. The observation protocol was based on the Reading Comprehension and Vocabulary (RCV) Observation Measure developed by Gersten, Dimino, and Jayanthi (2007). The inter-observer reliability of the RCV Observation Measure was 84.49% on average for the vocabulary scale and 90.89% on average for the comprehension scale. The reliability (Cronbach's alpha) of the vocabulary scale was .70; the internal consistency coefficient for the comprehension scale was .69 (Gersten, Dimino, & Jayanthi, 2007). Future research should assess the reliability and validity of the adapted observation protocol used in the present study.

Issues observed in the comparison school.

There were indications in the comparison school that the principal and teachers were not of the same opinion regarding the comprehension curricula used in the school. During the initial meeting with the research team, the principal remarked that she felt that the structure of the BEL Mondo program was good for the primary grade students and that she thought the program provided effective instruction in comprehension skills and strategies. The principal also thought that the fourth graders in the school struggled largely because they had moved from the structured BEL Mondo program to the less-structured Teachers College program. The principal said that she would like to implement the BEL Mondo program in grades K-6, but that the teachers in the upper grades were resistant.

Through discussions with the teachers who were using the BEL Mondo program, however, we found that they were less excited about the program than the principal was. The teachers did not feel as if there was enough rigor in the program to keep the students on task. They also complained that the goals for the lessons were not always made clear and that many of the independent activities felt like "busy work without a purpose." An important component of

the BEL Mondo program is ongoing assessments, which are meant to drive instruction and student progress. The teachers, however, were overwhelmed by the constant assessments and either did not administer them or did not feel that the assessments helped to guide their instruction.

Because an important factor in the success of any program is the “buy-in” of the teachers, these issues suggest that the teachers may not have been utilizing the Mondo comprehension program as effectively as possible. While these types of issues are not limited to the comparison school in this study, and can certainly be found in many schools, we feel that it is important to acknowledge the tension between the teachers and principal regarding the program. This lack of buy-in could have negatively affected the instruction of the teachers and the performance of the students.

Issues observed in the treatment school.

We did not observe the same differences in the opinions of the principals and the teachers concerning the UEE curriculum in the treatment school. That being said, there were discrepancies among the teachers regarding their relationships with the principal of the school and with Urban Education Exchange. Most of the teachers bought into the principal’s choice of curriculum and were amenable to the professional development that was administered by UEE. There was a significant minority, however, who neither believed in the principal’s plan for the school nor cooperated with the UEE teacher trainers. These teachers did not consent to allow us to observe their classrooms, and it is not clear how much (if any) of the UEE program they followed.

In April of Year 2, there was an incident – unrelated to this project – in the treatment school which resulted in the dismissal of a fourth grade teacher. Following this incident, the

mood of the school (particularly in the fourth grade) changed noticeably. The teachers were less willing to allow me to observe their instruction and told me they did not think their students would be able to perform very well on their final assessments. This incident, therefore, may have negatively affected the scores of some students in the treatment school on the assessments administered in the spring of Year 2.

Further analyses.

This study resulted in the development of an observation protocol that can be used to determine the number of instances of explicit instruction that are demonstrated by teachers during reading comprehension lessons. This instrument could be used in future studies to investigate the relationship between explicitness and student achievement with more rigor. A larger sample utilizing more schools and teachers would be useful in both further validating the observation protocol and examining in more detail the relationship between teacher explicitness and reading comprehension achievement.

Conclusions

The consensus among educators and educational researchers is that explicit instruction is important and does, in fact, lead to better reading comprehension. Many of these educators and researchers cite the work of the National Reading Panel (2000) as evidence that explicit instruction will lead to better student outcomes. Others cite research involving evaluations of instructional programs featuring explicit instruction (Duffy et. al., 2002; Pressley, 2006). While the empirical evidence indicates that explicit instructional *programs* can improve reading comprehension, researchers have not directly examined the relationship between *teacher* explicitness and student achievement. That is, research has not yet demonstrated that those teachers who are more explicit have students with better achievement outcomes.

Overall, our findings are inconclusive as to whether explicit reading comprehension instruction, such as is found in the UEE Comprehension Program, results in improved reading comprehension achievement as compared to a more traditional approach to reading comprehension instruction, such as is found in the BEL Mondo and/or Teachers College Programs. Comparisons of student performance in these two schools on reading assessments did not result in differences across grades K-5. This is not unexpected, given the previous literature on this topic that indicates the difficulty in demonstrating such differences in classroom settings. There was a suggestion, however, that the treatment program was more effective for grades K-2.

In this study, we found evidence that individual teachers, regardless of the instructional program they are using, vary widely in the explicitness of their instruction. We also found that teaching an explicit program makes a difference in how explicit an individual teacher's teaching is: First, the teachers who were trained in teaching the explicit program were more explicit, even in parts of the program where explicitness was not required, than teachers untrained in the explicit program. Second, teachers with more experience teaching the explicit program were more explicit than teachers teaching the explicit program for the first time.

We did find small indications of differences between the programs. The hints of superiority of the treatment program detected earlier in grades K-2 were supported when only experienced teachers and experienced students are examined. We found, across all grades, but most strongly in grades K-2, that the students of experienced treatment teachers performed significantly better on the GRADE passage comprehension subtest than did students of experienced comparison teachers. (Moreover, it may not be only the experience of the teachers that matters. We also found that the experienced treatment students in the second grade

performed better on most of the GRADE subtests than did the experienced comparison students in the second grade.)

These findings suggest that there is a link between experience and explicitness that is important in determining student achievement outcomes. It may be that the relationship of explicit instruction and student achievement is not straightforward. The present findings suggest that the relationship may be moderated by teacher experience, which may provide a benefit because of the increased explicitness that comes with experience.

We conclude that there is a complex relationship between teacher explicitness, teacher and student experience, and student achievement. This study offers some indication that a program that features explicit instruction in the early elementary grades can improve student reading comprehension outcomes if the teachers have had experience with the program. Such effects, however, do not always appear and, when they do, they are not large. But given the importance of this question in the literature and in the professional development of teachers, there is good reason to pursue further research on the relationship between teacher explicitness, teacher experience, student experience, and student achievement in classroom settings.

REFERENCES

- Afflerbach, P., Pearson, P.D., & Paris, S.G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, *61*(5), 364-373.
- Anagnostopoulos, P.M. (2010). School allocation memorandum no. 9, FY11. New York City Department of Education.
- Armbruster, B., Lehr, F., & Osborn, J. (2001). *Putting reading first: The research building blocks for teaching children to read*. Publication developed by the Center for the Improvement of Early Reading Achievement (CIERA).
- Atkins, J.G. (2008). Questions for ReadWorks Comprehension Assessments, Grades 2-5.
- Baker, S.K., Gersten, R., Haager, D., & Dingle, M. (2006). Teaching practice and the reading growth of first-grade English learners: Validation of an observation instrument. *The Elementary School Journal*, *107*(2), 199-219.
- Best, R. M., Rowe, M, Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, *25*, 62 - 80.
- Block, C. C., & Lacina, J. (2009). Comprehension instruction in kindergarten through grade three. In S. Israel & G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 494–509). New York: Routledge.
- Block, C. & Pressley, M. (2003). Best practices in comprehension instruction. In L. M. Morrow, L.G. Gambrell, & M. Pressley (Eds.), *Best practices in literacy instruction* (pp. 111-126). New York: Guilford Press.
- Brown, R. (2008). Teacher's attempts to teach comprehension strategies explicitly during core instruction. *Journal of School Connections*, *1*(1), 87-104.
- Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second grade readers. *Journal of Educational Psychology*, *88*, 18-37.
- Calfee, R. & Sperling, M. (2010). *On mixed methods: Approaches to language and literacy research*. New York, NY: Teachers College Press.
- Calkins, L.M. (2000). *The Art of Teaching Reading*. Boston, MA: Allyn & Bacon.
- Chamberlain, A., Daniels, C., Madden, N.A., & Slavin, R.E. (2009). A randomized evaluation of the Success for All middle school reading program. In D.L. Hough (Ed.), *Middle*

- grades research: Exemplary studies linking theory to practice* (pp. 21-41). Charlotte, NC: Information Age Publishing, Inc.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Correnti, R. & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal*, 44(2), 298-338.
- CTB/McGraw-Hill (2010). *New York State Testing Program 2010: English Language Arts, Grades 3–8 Technical Report*. Monterey, CA: The McGraw-Hill Companies, Inc.
- Dewitz, P., Jones, J., & Leahy, S. (2009). Comprehension strategy instruction in core reading programs. *Reading Research Quarterly*, 44(2), 102-126.
- Dole, J.A., Duffy, G.G., Roehler, L.R., & Pearson, P.D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research*, 61, 239-264.
- Duffy, G.G. (2002). The case for direct explanation of strategies. In C.C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 28-41). New York: Guilford Press.
- Duffy, G.G., Roehler, L.R., Meloth, M.S., Vavrus, L.G., Book, C., Putnam, J., & Wesselman, R. (1986). The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: A study of reading teacher effects. *Reading Research Quarterly*, 21(3), 237-252.
- Duffy, G.G., Roehler, L.R., Sivan, E., Rackliffe, G., Book, C., Meloth, M.S., Vavrus, L.G., Wesselman, R., Putnam, J., & Bassiri, D. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly*, 21(3), 237-252.
- Durkin, D. (1978/1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, 14(4), 481-533.
- Fullan, M. (2001). *The new meaning of education change* (3rd ed.). New York, NY: Teachers College Press.
- Gamse, B.C., Tepper-Jacob, R., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First impact study: Final report*. Washington, DC: Institute for Education Sciences, U.S. Department of Education.
- Gersten, R. (1996). The double demands of teaching English language learners. *Educational Leadership*, 53, 18-22.

- Gersten, R., Carnine, D., Zoref, L., & Cronin, D. (1986). A multifaceted study of change in seven inner-city schools. *The Elementary School Journal*, 86, 257-276.
- Gersten, R., Dimino, J., & Jayanthi, M. (2007). Toward the development of a nuanced classroom observational system for studying comprehension and vocabulary instruction. In B.M. Taylor & J.E. Ysseldyke (Eds.), *Effective instruction for struggling readers, K-6* (pp. 196-215). New York: Teachers College Press.
- Gersten, R., Dimino, J., Jayanthi, M., Kim, J.S., & Santoro, L.E. (2010). Teacher study group: Impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal*, 47(3), 694-739.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71(2), 279-320.
- Hall, G., & Hord, S. (2001). *Implementing change: Patterns, principles, and potholes*. Boston, MA: Allyn & Bacon.
- Harcourt Educational Measurement (2004). *Stanford technical data report*. San Antonio, TX: Author.
- Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 27(1), 4-9.
- Hill, P.W. & Jaggard, A.M. (2003). An evaluation report on the impact of Mondo Publishing's Building Essential Literacy (BEL) design and Bookshop Reading Program 1998-1999 and 1999-2000 school years. Mondo Publishing.
- Hirsch, Jr., E.D. (2003), Reading comprehension requires knowledge – of words and of the world. *American Educator* 27(1), 10-13, 16-22, 28-29, 48.
- James-Burdumy, S., Deke, J., Lugo-Gil, J., Carey, N., Hershey, A., Gersten, R., Newman-Gonchar, R., Dimino, J., & Haymond, K. (2010). Effectiveness of selected supplemental reading comprehension interventions: Findings from two student cohorts (NCEE 2010-4015). Washington, DC: Institute for Education Sciences, U.S. Department of Education.
- James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., et al. (2009). Effectiveness of selected supplemental reading comprehension interventions: Impacts on a first cohort of fifth-grade students (NCEE 2009-4032). Washington, DC: Institute for Education Sciences, U.S. Department of Education.
- Jitendra, A.K., Hoppes, M.K., & Xin, Y.P. (2000). Enhancing main idea comprehension for students with learning problems. *The Journal of Special Education*, 34(3), 127-139.

- Johnson, R.B. & Onwuegbuzie, A.J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Klinger, J.K., Urbach, J., Golos, D., Brownwell, M., & Menon, S. (2010). Teaching reading in the 21st century: A glimpse at how special education teacher promote reading comprehension. *Learning Disability Quarterly*, 33, 59-74.
- Kucan, L., & Beck, I. L. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research*, 67(3), 271-299.
- Liang, L.A. & Dole, J.A. (2006). Help with teaching reading comprehension: Comprehension instructional frameworks. *The Reading Teacher*, 59(8), 742-753.
- McCutchen, D, Abbott, R.D., Green, L.B., Beretvas, S.N., Cox, S., Potter, N.S., Quiroga, T., & Gray, A.L. (2002). Beginning Literacy: Links Among Teacher Knowledge, Teacher Practice, and Student Learning. *Journal of Learning Disabilities*, 35(1), 69-86.
- Mondo Publishing. (1997). Bookshop: A reading program. Greenvale, New York: Author.
- Moss, M., Fountain, A.R., Boulay, B., Horst, M., Rodger, C., & Brown-Lyons, M. (2008). *Reading First implementation evaluation: Final report*. Cambridge, MA: Abt Associates.
- Nation, K. & Norbury, C.F. (2005). Reading comprehension's new look: Influences of theory and technology on practice. *Topics in Language Disorders* 25(1), 21-32.
- National Reading Panel, National Institute of Child Health and Human Development. (2000). Report of the National research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Ness, M. (2011). Explicit reading comprehension instruction in elementary classrooms: teacher use of reading comprehension strategies. *Journal of Research in Childhood Education*, 25, 98-117.
- New York City Department of Education (2010). School Demographics and Accountability Snapshot (Version 2010-1B - April 2010).
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction*, 2, 117-175.
- Pearson, P.D. (2009). The Roots of Reading Comprehension Instruction. In S.E. Israel & G.G. Duffy (Eds.), *Handbook of Research on Reading Comprehension* (pp. 3-31). New York: Routledge.

- Pearson, P.D. & Duke, N.K. (2002). Comprehension instruction in the primary grades. In C.C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 28-41). New York: Guilford Press.
- Perfetti, C.A., Marron, M.A., & Foltz, P.W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi and J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and difficulties*, (pp.853-70). Mahwah, NJ: Erlbaum.
- Pressley, M. (1998). Comprehension strategies instruction. In J. Osborn and F. Lehr (Eds.), *Literacy for All: Issues in Teaching and Learning* (pp. 113-133).
- Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 545-561). Mahwah, NJ: Erlbaum.
- Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching* (3rd ed.). New York: Guilford.
- Raphael, T.E., & Wonnacott, C.A. (1985). Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly*, 20(3), 282-296.
- Raphael, T.E., Wonnacott, C.A., & Pearson (1983). Increasing students' sensitivity to sources of information: An instructional study in question-answer relationships. *Center for the Study of Reading: Technical Report No. 284*.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, 66(2), 181-221.
- Sammons, P. (2010). The contribution of mixed methods to recent research on educational effectiveness. In A. Tashakkori, & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social and Behavioral Research: Second Edition* (pp. 697-723). Los Angeles, CA: Sage Publications.
- Scanlon, D.M. & Vellutino, F.R. (1996). Prerequisite skills, early instruction, and success in first-grade reading: Selected results from a longitudinal study. *Mental Retardation & Developmental Disabilities Research Reviews*, 2(1), 54-63.
- Slavin, R.E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research*, 79(4), 1391-1466.
- Snow, C.E. (2002). *Reading for understanding: Toward an R&D program in reading*. RAND: Santa Monica, CA.

- Swanson, H.L. (1999). Reading research for students with LD. *Journal of Learning Disabilities*, 32(6), 504-532
- Teddlie, C. & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, 13(1), 12-28.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2011 Reading Assessment.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89, 420-432.
- Williams, J.P. (1998). Improving the comprehension of disabled readers. *Annals of Dyslexia* 48(1), 213-238.
- Williams, J.P. (2002). Reading comprehension strategies and teacher preparation. In A.E. Farstrup & S.J. Samuels (Eds.), *What Research has to Say About Reading Instruction* (pp. 243-260). Newark, DE: International Reading Association.
- Williams, J.P. & Atkins, J.G. (2009). The role of metacognition in teaching reading comprehension to primary students. In D.J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education*. Mahwah, N.J.: Erlbaum.\
- Williams, J.P. & Atkins, J.G. (2011). A preliminary evaluation of the ReadWorks comprehension program. Unpublished manuscript.
- Williams, J.P., Stafford, K.B., Lauer, K.D., Hall, K.M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology*, 101(1), 1-20.
- Williams, K.T. (2001). *Group Reading and Diagnostic Evaluation*. Circle Pines, MN: American Guidance Service.

APPENDIX A

Table 1: School Demographics

	Treatment		Comparison	
	2008-2009	2009-2010	2008-2009	2009-2010
Grades	K – 5	K - 5	Pre-K - 6	Pre-K - 5
Student Population	449	438	598	581
	Percentage of Students		Percentage of Students	
Gender				
Male	52.1	52.1	51.2	50.8
Female	47.9	47.9	48.8	49.2
Ethnicity				
American Indian	0.4	0.2	0.3	0.3
Black	89.5	88.1	89.6	89.3
Hispanic	4.2	5.0	6.4	6.4
Asian / Pac. Isl	0.7	1.1	1.0	1.2
White	2.7	3.4	1.7	1.7
% of days students attended	93.7	NA	92.1	NA
Student stability: % of enrollment	90.4		88.5	
Poverty rate: % of enrollment	55.3	75.1	65.8	84.9
Special Education	Number of Students		Number of Students	
General Education Classes	16	15	33	31
Collaborative Team Teaching Classes	31	27	0	3
Self Contained Classes	15	21	61	37
English Language Learners	4	6	15	17
Temporary housing: Total # of students	19	NA	20	NA
Recent immigrants: Total # of students	1	0	2	0

APPENDIX A (continued)

Table 1: School Demographics (continued)

	Treatment		Comparison	
	2008-2009	2009-2010	2008-2009	2009-2010
	Percentage of Teachers		Percentage of Teachers	
Teacher Qualifications				
Fully licensed and permanently assigned to school	97.1	100.0	100.0	100.0
More than two years at the school	50.0	NA	78.6	NA
More than five years teaching anywhere	61.8	NA	83.3	NA
Masters Degree or higher	76.0	NA	95.0	NA
Core classes taught by “highly qualified” teachers	93.8	NA	94.3	NA

APPENDIX A (continued)

Table 2: Assessments Conducted

Year One

Assessment	Grade					
	K	1	2	3	4	5
UEE Comprehension						
Listening Comprehension			X			
Reading Comprehension				X	X	X
GRADE						
Word Reading	X	X	X			
Vocabulary				X	X	X
Listening Comprehension	X	X	X	X	X	X
Sentence Comprehension		X	X	X	X	X
Passage Comprehension		X	X	X	X	X
Stanford 10						
Listening Comprehension	X	X				

Year Two

Assessment	Grade					
	K	1	2	3	4	5
UEE Comprehension						
Listening Comprehension			X			
Reading Comprehension				X	X	X
GRADE						
Word Reading	X	X	X			
Vocabulary				X	X	X
Listening Comprehension	X	X	X	X	X	X
Sentence Comprehension		X	X	X	X	X
Passage Comprehension		X	X	X	X	X
Stanford 10						
Listening Comprehension						

APPENDIX A (continued)

Table 3: Observations Conducted

Year One

School	Dates	Grade	Teachers (n)	Observations (n)
Treatment	10/10/08 – 05/15/09	K	3	4
		1	2	4
		2	2	5
		3	3	3
		4	2	2
		5	2	2
		Total	14	20
Comparison	01/30/09 – 05/29/09	K	3	6
		1	3	8
		2	2	3
		3	2	2
		4	3	3
		5	1	1
		Total	14	23

Year Two

School	Dates	Grade	Teachers (n)	Observations (n)
Treatment	11/17/09 – 05/24/10	K	2	4
		1	2	4
		2	2	4
		3	2	2
		4	3	5
		5	2	3
		Total	13	22
Comparison	11/23/09 – 05/27/10	K	2	4
		1	1	2
		2	1	2
		3	2	3
		4	1	1
		5	2	3
		Total	9	15

APPENDIX A (continued)

Figure 1: Classroom Observation Form

Teacher: _____ School: _____ Date: _____
 Begin lesson: _____ End lesson: _____

A. Explicitness of instruction	Tally	Total	Notes
Prior to reading, teacher			
1. Conducts preparatory activities: relating text to experiences, other texts, background knowledge, browsing (book cover, spine, TOC, pictures)			
2. Explicitly states the learning outcome	Y	N	
3. Introduces new vocabulary			
4. Transitions from activation of prior knowledge into direct teaching	Y	N	
During or after reading, teacher			
2. Models the following skills (including think-alouds)			(Did teacher explicitly state when and how to use the strategy?)
a. Finding explicit information in the text			
b. Using prior knowledge to draw conclusions			
c. Using strategies to identify vocabulary in context			
d. Understanding figurative language			
e. Identifying the genre of a text			
f. Understanding the sequence of events			
g. Identifies character elements: traits/relationships			
h. Identifying the setting of the passage			
i. Identifying the problem and solution in the plot			
j. Identifying cause/effect text structure			
k. Predicting outcomes			
l. Identifying the main idea and supporting details			
m. Classifying and categorizing information			
n. Identifying facts and opinions			
o. Identifying compare/contrast text structure			
p. Identifying the point of view in the passage			
q. Identifying the voice: how it is written			
r. Identifying the author's purpose			
s. Identifying the theme			

B. Student practice	Tally	Total	Notes
Prior to independent practice, teacher			
1. Checked for student understanding of the modeled skill			
2. Provided appropriate guided practice that supported the skill and prepared them for independent practice			
During or after independent practice, teacher			
3. Provided opportunity for students to independently apply skill to provide teacher with evidence of student mastery			
4. Asks students to answer literal recall questions from the text (specific questions)			
5. Asks students questions requiring inferences based on text			
6. Asks students to justify or elaborate responses and explain thinking			
7. Keeps students thinking for 2+ seconds before calling on a student for a response			
8. Gives practice in answering comprehension questions or applying comprehension strategies with expected product			
9. Monitored student work to check for understanding (teachable moments, assessment, scaffolding)			
10. Differentiated instruction and activities for additional support and enrichment purposes			
11. Appropriately closes or ended the lesson			

C. Corrective feedback: Teacher	Tally	Total	Notes
1. Communicates clearly what student/s did correctly about the strategy			
2. Reinstructs when student makes a mistake by encouraging child to try again or reminding student about comprehension strategy			
D. Uses a graphic organizer before, during, or after lesson	Y	N	

Code every 5 minutes	Grouping arrangements 1. Whole class 2. Small groups 3. Pairs 4. Individual	Focus of activity 1. Direct reading 2. Indirect reading 3. Writing 4. Management 5. Prep activities	Teacher's purpose 1. Reading aloud 2. Modeling 3. Asking/answering questions 4. Guiding student practice 5. Providing feedback	Student engagement 1. Few students seem engaged 2. Many students seem engaged much of the time 3. Most students seem engaged all of the time
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3

Answer the following at the end of the observation:

A. During comprehension instruction:

1. Teacher gave inaccurate and/or confusing explanations while modeling strategies	Y N	
2. Teacher missed opportunity to correct or address error, or provided confusing or inaccurate feedback	Y N	
3. Teacher took advantage of opportunities to teach vocabulary and increase general knowledge about topics from the passage	Y N	
3. Teacher called individually on about half or more of the students	Y N	
4. Teacher differentiated pre-reading, during reading, and after reading strategies	Y N	

B. Based on your overall judgment, how would you rate each domain you observed?

Comprehension Excellent	Not Observed	Minimal/Erratic	Partially Effective	Good
Vocabulary Excellent	Not Observed	Minimal/Erratic	Partially Effective	Good

C. Please rate the management/responsiveness to students on a 4-point scale

1. The instructional routines appear to be	Minimal	Fair	Good	Excellent
2. The teacher maximizes the amount of time available for instruction	Minimal	Fair	Good	Excellent
3. The teacher manages student behavior effectively in order to avoid disruptions and to provide productive learning environments	Minimal	Fair	Good	Excellent

Additional notes:

APPENDIX B

Table 4: Significant differences between the treatment school and the comparison school (ANCOVA results)

Year 1

Assessment	Combined Grades				Individual Grades						
	All		K-2	3-5		K	1	2	3	4	5
UEE Comprehension											
GRADE											
Word Reading											
Vocabulary											
Listening Comprehension				C				T		C	
Sentence Comprehension	T		T	T			T			T	
Passage Comprehension											

Year 2

Assessment	Combined Grades				Individual Grades						
	All		K-2	3-5		K	1	2	3	4	5
UEE Comprehension								T			C
GRADE											
Word Reading								T			
Vocabulary											C
Listening Comprehension			T			T				C	
Sentence Comprehension	C			C							
Passage Comprehension	T		T	T			T				T

Shaded cell: The assessment was not administered in this grade

Blank cell: There was no difference between the schools

T: The treatment school significantly outperformed the comparison school (p<.05)

C: The comparison school significantly outperformed the treatment school (p<.05)

APPENDIX B (continued)

Table 5: Differences between the treatment school and the comparison school (ANCOVA results)

Year 1

Assessment	Combined Grades				Individual Grades						
	All		K-2	3-5		K	1	2	3	4	5
UEE Comprehension	T			T				C	T	T	N
GRADE											
Word Reading	C		N			C	C	T			
Vocabulary	C			N					T	C	C
Listening Comprehension	C		T	C*		C	C	T*	C	C*	T
Sentence Comprehension	T*		T*	T*			T*	T	N	T*	T
Passage Comprehension	T		T	T			T	C	T	C	T

Year 2

Assessment	Combined Grades				Individual Grades						
	All		K-2	3-5		K	1	2	3	4	5
UEE Comprehension	C			C				T*	N	C	C*
GRADE											
Word Reading	C		N			C	C	T*			
Vocabulary	C			C					T	T	C*
Listening Comprehension	N		T*	C [†]		T*	C	T [†]	C	C*	C
Sentence Comprehension	C*		C	C*			N	C [†]	C	C [†]	C
Passage Comprehension	T*		T*	T*			T*	T	C	T	T*

T: The treatment school outperformed the comparison school
 C: The comparison school outperformed the treatment school
 N: The treatment and comparison schools had the same mean score

* The difference between the schools was significant (p<.05)
[†] The difference between the schools was near-significant (p<.10)

APPENDIX B (continued)

Table 6: GRADE – All Grades – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Listening comprehension	T	301	.80	.20	.81	.21	.80	.01	1, 626	2.27	.13
	C	328	.75	.21	.81	.17	.82	.01			
Word Reading / Vocabulary	T	307	.62	.24	.69	.24	.67	.01	1, 621	1.96	.16
	C	317	.57	.25	.67	.25	.69	.01			
Sentence comprehension	T	266	.66	.27	.76	.23	.77	.01	1, 529	34.95	.00*
	C	266	.68	.32	.67	.27	.67	.01			
Passage comprehension	T	278	.52	.23	.59	.22	.57	.01	1, 527	.78	.38
	C	252	.47	.21	.54	.22	.56	.01			
Total	T	323	.65	.18	.71	.19	.70	.01	1, 661	.06	.81
	C	341	.62	.19	.69	.19	.70	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the treatment school were significantly above those of the comparison school

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 6a: GRADE – All Grades – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared [†]
Listening Comprehension	Pretest	6.60	1	6.60	263.29**	.30
	Condition	.06	1	.06	2.27	.00
	Error	15.70	626	.03		
Word Reading / Vocabulary	Pretest	19.16	1	19.16	632.87**	.51
	Condition	.06	1	.06	1.96	.00
	Error	18.80	621	.03		
Sentence comprehension	Pretest	11.39	1	11.39	279.47**	.35
	Condition	1.42	1	1.42	34.95**	.06
	Error	21.55	529	.04		
Passage comprehension	Pretest	11.54	1	11.54	425.74**	.45
	Condition	.02	1	.02	.78	.00
	Error	14.29	527	.03		
Total	Pretest	13.16	1	13.16	796.65**	.55
	Condition	.00	1	.00	.06	.00
	Error	10.92	661	.02		

*p<0.05

**p < 0.01

$$\eta^2_{\text{partial}} = \frac{SS_{\text{effect}}}{SS_{\text{effect}} + SS_{\text{error}}}$$

[†] Partial Eta Squared is used as an estimate of effect size. Partial eta², or η^2_{partial} , is defined as follows:

Cohen (1988) established conventions for small ($\eta^2_{\text{partial}}=.01$), medium ($\eta^2_{\text{partial}}=.06$), and large ($\eta^2_{\text{partial}}=.14$) effects.

APPENDIX B (continued)

Table 7: GRADE – All Grades – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	312	.81	.14	.85	.12	.85	.01	1, 644	.01	.95
	C	335	.81	.13	.85	.10	.85	.01			
Word Reading / Vocabulary	T	337	.58	.24	.71	.24	.70	.01	1, 659	1.98	.16
	C	325	.55	.25	.71	.24	.72	.01			
Sentence comprehension	T	282	.61	.25	.74	.23	.73	.01	1, 548	6.27	.01*
	C	269	.57	.27	.75	.22	.77	.01			
Passage comprehension	T	273	.49	.22	.61	.19	.61	.01	1, 549	6.99	.01*
	C	279	.47	.20	.56	.22	.57	.01			
Total	T	354	.63	.17	.74	.17	.73	.01	1, 706	1.20	.27
	C	355	.60	.16	.73	.17	.74	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the comparison school were significantly above those of the treatment school

Passage Comprehension: Scores of the treatment school were significantly above those of the comparison school

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 7a: GRADE – All Grades – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.44	1	.44	39.87**	.06
	Condition	.00	1	.00	.00	.00
	Error	7.06	644	.01		
Word Reading / Vocabulary	Pretest	16.57	1	16.57	513.51**	.44
	Condition	.06	1	.06	1.98	.00
	Error	21.26	659	.03		
Sentence comprehension	Pretest	9.76	1	9.76	291.23**	.35
	Condition	.21	1	.21	6.27*	.01
	Error	18.36	548	.03		
Passage comprehension	Pretest	8.47	1	8.47	312.45**	.36
	Condition	.19	1	.19	6.99**	.01
	Error	14.88	549	.03		
Total	Pretest	8.93	1	8.93	579.46**	.45
	Condition	.02	1	.02	1.20	.00
	Error	10.88	706	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 8: GRADE – Grades K-2 – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	128	.83	.15	.87	.17	.85	.01	1, 283	.08	.78
	C	158	.77	.18	.83	.19	.84	.01			
Word Reading	T	124	.74	.27	.85	.21	.85	.01	1, 270	.16	.69
	C	149	.70	.24	.85	.19	.85	.01			
Sentence comprehension	T	85	.53	.33	.74	.25	.74	.02	1, 179	9.12	.00*
	C	97	.53	.29	.65	.28	.65	.02			
Passage comprehension	T	107	.45	.25	.58	.21	.56	.02	1, 191	.25	.62
	C	87	.39	.22	.53	.23	.55	.02			
Total	T	133	.65	.21	.77	.19	.76	.01	1, 298	.10	.76
	C	168	.63	.20	.76	.20	.76	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the treatment school were significantly above those of the comparison school

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 8a: GRADE – Grades K-2 – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	1.91	1	1.91	76.54**	.21
	Condition	.00	1	.00	.08	.00
	Error	7.08	283	.03		
Word Reading	Pretest	3.69	1	3.69	145.49**	.35
	Condition	.00	1	.00	.16	.00
	Error	6.85	270	.03		
Sentence comprehension	Pretest	5.18	1	5.18	118.36**	.40
	Condition	.40	1	.40	9.12**	.05
	Error	7.84	179	.04		
Passage comprehension	Pretest	3.99	1	3.99	142.36**	.43
	Condition	.01	1	.01	.25	.00
	Error	5.35	191	.03		
Total	Pretest	6.04	1	6.04	340.99**	.53
	Condition	.00	1	.00	.10	.00
	Error	5.28	298	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 9: GRADE – Grades K-2 – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	143	.79	.17	.89	.11	.89	.01	1, 312	4.81	.03*
	C	172	.80	.14	.87	.09	.87	.01			
Word Reading	T	141	.69	.27	.89	.16	.88	.01	1, 297	.14	.71
	C	159	.64	.27	.87	.17	.88	.01			
Sentence comprehension	T	87	.55	.30	.72	.26	.71	.02	1, 190	1.89	.17
	C	106	.52	.28	.75	.25	.75	.02			
Passage comprehension	T	78	.38	.21	.61	.20	.62	.02	1, 186	6.53	.01*
	C	111	.41	.18	.57	.21	.56	.02			
Total	T	150	.64	.19	.82	.15	.81	.01	1, 327	.51	.47
	C	180	.61	.17	.79	.16	.80	.01			

ANCOVA RESULTS

Listening Comprehension: Scores of the treatment school were significantly above those of the comparison school

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: Scores of the treatment school were significantly above those of the comparison school

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 9a: GRADE – Grades K-2 – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.22	1	.22	22.95**	.07
	Condition	.05	1	.05	4.81*	.02
	Error	2.96	312	.01		
Word Reading	Pretest	2.07	1	2.07	99.30**	.25
	Condition	.00	1	.00	.14	.00
	Error	6.19	297	.02		
Sentence comprehension	Pretest	3.02	1	3.02	61.83**	.25
	Condition	.09	1	.09	1.89	.01
	Error	9.28	190	.05		
Passage comprehension	Pretest	2.57	1	2.57	88.68**	.32
	Condition	.19	1	.19	6.53*	.03
	Error	5.40	186	.03		
Total	Pretest	2.80	1	2.80	180.86**	.36
	Condition	.01	1	.01	.51	.00
	Error	5.07	327	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 10: GRADE – Grades 3-5 – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Listening comprehension	T	173	.77	.23	.77	.22	.77	.01	1, 340	4.04	.05*
	C	170	.74	.24	.79	.16	.80	.01			
Vocabulary	T	183	.54	.18	.58	.19	.55	.01	1, 348	.03	.87
	C	168	.46	.19	.52	.20	.55	.01			
Sentence comprehension	T	181	.72	.21	.77	.22	.79	.01	1, 347	30.96	.00*
	C	169	.76	.30	.68	.26	.67	.01			
Passage comprehension	T	171	.56	.21	.59	.23	.57	.01	1, 333	.08	.78
	C	165	.51	.20	.54	.21	.56	.01			
Total	T	190	.64	.16	.67	.18	.66	.01	1, 360	1.89	.17
	C	173	.62	.17	.64	.17	.64	.01			

ANCOVA RESULTS

Listening Comprehension: Scores of the comparison school were significantly above those of the treatment school

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the treatment school were significantly above those of the comparison school

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 10a: GRADE – Grades 3-5 – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	4.29	1	4.29	176.08**	.34
	Condition	.10	1	.10	4.04*	.01
	Error	8.28	340	.02		
Vocabulary	Pretest	5.83	1	5.83	272.13**	.44
	Condition	.00	1	.00	.03	.00
	Error	7.46	348	.02		
Sentence comprehension	Pretest	6.84	1	6.84	182.85**	.35
	Condition	1.16	1	1.16	30.96**	.08
	Error	12.97	347	.04		
Passage comprehension	Pretest	8.26	1	8.26	334.71**	.50
	Condition	.00	1	.00	.08	.00
	Error	8.22	333	.03		
Total	Pretest	6.99	1	6.99	677.02**	.65
	Condition	.02	1	.02	1.89	.01
	Error	3.72	360	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 11: GRADE – Grade 3-5 – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	169	.83	.10	.82	.11	.81	.01	1, 329	3.07	.08 [†]
	C	163	.81	.11	.83	.10	.83	.01			
Vocabulary	T	196	.50	.18	.58	.19	.56	.01	1, 359	.02	.88
	C	166	.47	.20	.55	.20	.57	.01			
Sentence comprehension	T	195	.64	.22	.75	.21	.74	.01	1, 355	4.41	.04*
	C	163	.60	.25	.76	.21	.77	.01			
Passage comprehension	T	195	.53	.21	.61	.19	.60	.01	1, 360	4.26	.04*
	C	168	.50	.21	.56	.22	.57	.01			
Total	T	204	.61	.16	.68	.15	.67	.01	1, 376	.79	.38
	C	175	.59	.16	.67	.16	.68	.01			

ANCOVA RESULTS

Listening Comprehension: The treatment school outperformed the comparison school, with the difference approaching significance

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the comparison school were significantly above those of the treatment school

Passage Comprehension: Scores of the treatment school were significantly above those of the comparison school

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 11a: GRADE – Grade 3-5 – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.39	1	.39	38.72**	.10
	Condition	.03	1	.03	3.07	.01
	Error	3.33	329	.01		
Vocabulary	Pretest	7.45	1	7.45	471.65**	.57
	Condition	.00	1	.00	.02	.00
	Error	5.67	359	.02		
Sentence comprehension	Pretest	7.00	1	7.00	283.87**	.44
	Condition	.11	1	.11	4.41*	.01
	Error	8.75	355	.03		
Passage comprehension	Pretest	6.63	1	6.63	273.25**	.43
	Condition	.10	1	.10	4.26*	.01
	Error	8.74	360	.02		
Total	Pretest	5.73	1	5.73	661.61**	.64
	Condition	.01	1	.01	.79	.00
	Error	3.26	376	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 12: GRADE – Kindergarten – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	23	.84	.13	.89	.17	.86	.03	1, 77	.82	.37
	C	57	.78	.17	.88	.16	.89	.02			
Word reading	T	22	.59	.30	.80	.28	.81	.04	1, 70	.29	.59
	C	51	.61	.26	.85	.23	.84	.03			
Total	T	23	.72	.19	.85	.19	.83	.03	1, 81	.31	.58
	C	61	.69	.20	.85	.19	.85	.02			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 12a: GRADE – Kindergarten – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.73	1	.73	40.85**	.35
	Condition	.02	1	.02	.82	.01
	Error	1.37	77	.02		
Word Reading	Pretest	1.23	1	1.23	29.07**	.29
	Condition	.01	1	.01	.29	.00
	Error	2.97	70	.04		
Total	Pretest	1.45	1	1.45	73.21**	.48
	Condition	.01	1	.01	.31	.00
	Error	1.60	81	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 13: GRADE – Kindergarten – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	42	.80	.13	.94	.07	.94	.01	1, 94	11.51	.00*
	C	55	.80	.15	.89	.09	.89	.01			
Word reading	T	39	.52	.28	.84	.23	.82	.04	1, 90	.70	.40
	C	54	.39	.22	.83	.24	.85	.03			
Total	T	42	.67	.17	.89	.13	.87	.02	1, 96	.00	.99
	C	57	.60	.16	.86	.16	.87	.02			

ANCOVA RESULTS

Listening Comprehension: Scores of the treatment school were significantly above those of the comparison school

Word reading: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 13a: GRADE – Kindergarten – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.26	1	.26	54.55**	.37
	Condition	.06	1	.06	11.51**	.11
	Error	.45	94	.01		
Word Reading	Pretest	.81	1	.81	17.61**	.16
	Condition	.03	1	.03	.70	.01
	Error	4.16	90	.05		
Total	Pretest	.58	1	.58	37.18**	.28
	Condition	.00	1	.00	.00	.00
	Error	1.49	96	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 14: GRADE – First Grade – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	57	.82	.19	.84	.21	.84	.03	1, 107	.08	.77
	C	53	.80	.19	.84	.21	.85	.03			
Word reading	T	54	.67	.28	.81	.24	.82	.02	1, 102	.01	.93
	C	51	.70	.21	.84	.17	.83	.02			
Sentence comprehension	T	53	.47	.32	.70	.29	.73	.03	1, 101	7.66	.01*
	C	51	.49	.31	.62	.30	.61	.03			
Passage comprehension	T	58	.40	.27	.54	.24	.54	.02	1, 102	2.00	.16
	C	47	.34	.20	.47	.25	.49	.03			
Total	T	60	.57	.23	.72	.20	.72	.02	1, 111	1.30	.26
	C	54	.58	.19	.69	.21	.69	.02			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the treatment school were significantly above those of the comparison school

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 14a: GRADE – First Grade – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	1.06	1	1.06	31.02**	.23
	Condition	.00	1	.00	.08	.00
	Error	3.66	107	.03		
Word Reading	Pretest	2.05	1	2.05	86.77**	.46
	Condition	.00	1	.00	.01	.00
	Error	2.41	102	.02		
Sentence comprehension	Pretest	3.08	1	3.08	63.14**	.39
	Condition	.37	1	.37	7.66**	.07
	Error	4.92	101	.05		
Passage comprehension	Pretest	2.32	1	2.32	69.56**	.41
	Condition	.07	1	.07	2.00	.02
	Error	3.40	102	.03		
Total	Pretest	2.55	1	2.55	132.21**	.54
	Condition	.03	1	.03	1.30	.01
	Error	2.14	111	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 15: GRADE – First Grade – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	48	.79	.22	.90	.08	.90	.01	1, 99	.39	.54
	C	54	.80	.16	.91	.07	.91	.01			
Word reading	T	48	.67	.25	.88	.14	.88	.02	1, 87	.22	.64
	C	42	.68	.22	.90	.13	.90	.02			
Sentence comprehension	T	32	.40	.33	.72	.27	.74	.04	1, 72	.02	.89
	C	43	.47	.29	.76	.28	.74	.04			
Passage comprehension	T	43	.29	.17	.58	.21	.62	.03	1, 85	5.88	.02*
	C	45	.40	.16	.56	.25	.52	.03			
Total	T	52	.56	.20	.78	.16	.79	.02	1, 104	.70	.41
	C	55	.61	.18	.78	.18	.77	.02			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Word reading: No difference between the treatment and comparison schools

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: Scores of the treatment school were significantly above those of the comparison school

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 15a: GRADE – First Grade – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.00	1	.00	.41	.00
	Condition	.00	1	.00	.39	.00
	Error	.54	99	.01		
Word Reading	Pretest	.56	1	.56	45.19**	.34
	Condition	.00	1	.00	.22	.00
	Error	1.07	87	.01		
Sentence comprehension	Pretest	1.24	1	1.24	20.58**	.22
	Condition	.00	1	.00	.02	.00
	Error	4.34	72	.06		
Passage comprehension	Pretest	1.28	1	1.28	32.76**	.28
	Condition	.23	1	.23	5.88*	.07
	Error	3.33	85	.04		
Total	Pretest	1.06	1	1.06	58.53**	.36
	Condition	.01	1	.01	.70	.01
	Error	1.88	104	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 16: GRADE – Second Grade – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	48	.84	.11	.88	.10	.87	.02	1, 93	12.02	.00*
	C	48	.72	.20	.75	.16	.77	.02			
Word reading	T	48	.88	.16	.92	.10	.91	.02	1, 92	2.65	.11
	C	47	.78	.22	.86	.15	.87	.02			
Sentence comprehension	T	32	.64	.31	.77	.24	.75	.04	1, 75	1.38	.24
	C	46	.58	.27	.69	.26	.70	.03			
Passage comprehension	T	49	.52	.22	.61	.20	.59	.02	1, 86	.96	.33
	C	40	.44	.23	.60	.18	.62	.02			
Total	T	50	.72	.16	.79	.14	.76	.02	1, 100	.22	.64
	C	53	.61	.21	.72	.16	.75	.02			

ANCOVA RESULTS

Listening Comprehension: Scores of the treatment school were significantly above those of the comparison school

Word Reading: No difference between the treatment and comparison schools

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 16a: GRADE – Second Grade – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.12	1	.12	6.96*	.07
	Condition	.20	1	.20	12.02**	.11
	Error	1.58	93	.02		
Word Reading	Pretest	.32	1	.32	24.93**	.21
	Condition	.03	1	.03	2.65	.03
	Error	1.18	92	.01		
Sentence comprehension	Pretest	2.03	1	2.03	53.42**	.42
	Condition	.05	1	.05	1.38	.02
	Error	2.86	75	.04		
Passage comprehension	Pretest	1.31	1	1.31	62.11**	.42
	Condition	.02	1	.02	.96	.01
	Error	1.82	86	.02		
Total	Pretest	1.21	1	1.21	98.48**	.50
	Condition	.00	1	.00	.22	.00
	Error	1.23	100	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 17: GRADE – Second Grade – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	53	.79	.14	.85	.14	.86	.02	1, 113	2.72	.10 [†]
	C	63	.80	.11	.82	.10	.82	.01			
Word reading	T	54	.84	.19	.94	.09	.93	.01	1, 114	10.44	.00*
	C	63	.81	.16	.88	.12	.89	.01			
Sentence comprehension	T	55	.64	.25	.72	.25	.69	.03	1, 115	3.39	.07 [†]
	C	63	.56	.28	.74	.23	.76	.03			
Passage comprehension	T	35	.48	.21	.64	.17	.62	.02	1, 98	1.20	.28
	C	66	.42	.20	.57	.19	.59	.02			
Total	T	56	.71	.16	.80	.13	.77	.01	1, 121	.41	.52
	C	68	.63	.17	.74	.14	.76	.01			

ANCOVA RESULTS

Listening Comprehension: The treatment school outperformed the comparison school, with the difference approaching significance

Word Reading: Scores of the treatment school were significantly above those of the comparison school

Sentence Comprehension: The comparison school outperformed the treatment school, with the difference approaching significance

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 17a: GRADE – Second Grade – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.11	1	.11	8.87**	.07
	Condition	.03	1	.03	2.72	.02
	Error	1.42	113	.01		
Word Reading	Pretest	.65	1	.65	110.44**	.49
	Condition	.06	1	.06	10.44**	.08
	Error	.67	114	.01		
Sentence comprehension	Pretest	2.12	1	2.12	53.06**	.32
	Condition	.14	1	.14	3.39	.03
	Error	4.60	115	.04		
Passage comprehension	Pretest	1.33	1	1.33	66.28**	.40
	Condition	.02	1	.02	1.20	.01
	Error	1.96	98	.02		
Total	Pretest	1.33	1	1.33	206.85**	.63
	Condition	.00	1	.00	.41	.00
	Error	.78	121	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 18: GRADE – Third Grade – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	56	.80	.10	.83	.10	.83	.02	1, 101	.33	.57
	C	48	.80	.13	.84	.12	.84	.02			
Vocabulary	T	54	.65	.16	.69	.15	.69	.02	1, 98	1.16	.28
	C	47	.65	.19	.67	.18	.67	.02			
Sentence comprehension	T	59	.84	.12	.88	.11	.88	.01	1, 105	.02	.89
	C	49	.80	.17	.87	.14	.88	.02			
Passage comprehension	T	54	.58	.16	.60	.20	.59	.02	1, 98	.53	.47
	C	47	.54	.18	.54	.23	.56	.03			
Total	T	59	.72	.10	.76	.12	.75	.01	1, 106	.31	.58
	C	50	.70	.13	.73	.13	.74	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Vocabulary: No difference between the treatment and comparison schools

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 18a: GRADE – Third Grade – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.06	1	.06	4.46*	.04
	Condition	.00	1	.00	.33	.00
	Error	1.27	101	.01		
Vocabulary	Pretest	1.23	1	1.23	85.36**	.47
	Condition	.02	1	.02	1.16	.01
	Error	1.42	98	.01		
Sentence comprehension	Pretest	.44	1	.44	38.54**	.27
	Condition	.00	1	.00	.02	.00
	Error	1.21	105	.01		
Passage comprehension	Pretest	1.57	1	1.57	52.30**	.35
	Condition	.02	1	.02	.53	.01
	Error	2.94	98	.03		
Total	Pretest	.84	1	.84	98.47**	.48
	Condition	.00	1	.00	.31	.00
	Error	.90	106	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 19: GRADE – Third Grade – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	55	.81	.10	.85	.12	.85	.02	1, 99	.08	.78
	C	47	.81	.11	.86	.10	.86	.02			
Vocabulary	T	51	.59	.22	.72	.17	.71	.02	1, 100	1.33	.25
	C	52	.56	.23	.67	.20	.68	.02			
Sentence comprehension	T	56	.70	.24	.85	.18	.85	.02	1, 106	1.15	.29
	C	53	.68	.22	.87	.15	.87	.02			
Passage comprehension	T	53	.49	.23	.59	.19	.57	.02	1, 100	.16	.69
	C	50	.43	.20	.56	.19	.58	.02			
Total	T	57	.65	.18	.76	.14	.75	.01	1, 108	.02	.88
	C	54	.62	.16	.74	.14	.75	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Vocabulary: No difference between the treatment and comparison schools

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 19a: GRADE – Third Grade – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.09	1	.09	7.83**	.07
	Condition	.00	1	.00	.08	.00
	Error	1.16	99	.01		
Vocabulary	Pretest	1.87	1	1.87	113.10**	.53
	Condition	.02	1	.02	1.33	.01
	Error	1.66	100	.02		
Sentence comprehension	Pretest	1.32	1	1.32	90.37**	.46
	Condition	.02	1	.02	1.15	.01
	Error	1.54	106	.02		
Passage comprehension	Pretest	1.66	1	1.66	78.57**	.44
	Condition	.00	1	.00	.16	.00
	Error	2.12	100	.02		
Total	Pretest	1.29	1	1.29	178.23**	.62
	Condition	.00	1	.00	.02	.00
	Error	.78	108	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 20: GRADE – Fourth Grade – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	65	.66	.32	.67	.32	.66	.03	1, 124	4.65	.03*
	C	62	.62	.33	.74	.40	.75	.03			
Vocabulary	T	64	.53	.17	.55	.19	.50	.02	1, 123	.08	.78
	C	62	.37	.11	.46	.21	.51	.02			
Sentence comprehension	T	64	.66	.22	.71	.23	.77	.03	1, 123	44.57	.00*
	C	62	.89	.36	.57	.29	.51	.03			
Passage comprehension	T	65	.52	.25	.51	.25	.49	.02	1, 124	.16	.69
	C	62	.47	.21	.48	.22	.50	.02			
Total	T	65	.59	.16	.61	.15	.61	.01	1, 124	7.72	.01*
	C	62	.59	.21	.56	.17	.57	.01			

ANCOVA RESULTS

Listening Comprehension: Scores of the comparison school were significantly above those of the treatment school

Vocabulary: No difference between the treatment and comparison schools

Sentence Comprehension: Scores of the treatment school were significantly above those of the comparison school

Passage Comprehension: No difference between the treatment and comparison schools

Total: Scores of the treatment school were significantly above those of the comparison school

APPENDIX B (continued)

Table 20a: GRADE – Fourth Grade – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	3.25	1	3.25	68.52**	.36
	Condition	.22	1	.22	4.65*	.04
	Error	5.88	124	.05		
Vocabulary	Pretest	1.09	1	1.09	35.59**	.22
	Condition	.00	1	.00	.08	.00
	Error	3.77	123	.03		
Sentence comprehension	Pretest	3.26	1	3.26	79.08**	.39
	Condition	1.84	1	1.84	44.57**	.27
	Error	5.08	123	.04		
Passage comprehension	Pretest	4.55	1	4.55	213.28**	.63
	Condition	.00	1	.00	.16	.00
	Error	2.65	124	.02		
Total	Pretest	2.34	1	2.34	286.51**	.70
	Condition	.06	1	.06	7.72**	.06
	Error	1.01	124	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 21: GRADE – Fourth Grade – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	64	.84	.12	.79	.10	.78	.01	1, 116	4.50	.04*
	C	55	.78	.13	.81	.10	.82	.01			
Vocabulary	T	65	.52	.16	.57	.18	.54	.02	1, 115	1.97	.16
	C	53	.46	.17	.48	.19	.51	.02			
Sentence comprehension	T	64	.67	.19	.73	.19	.70	.02	1, 110	2.99	.09 [†]
	C	49	.54	.26	.71	.19	.75	.02			
Passage comprehension	T	64	.53	.21	.60	.21	.59	.02	1, 118	1.64	.20
	C	57	.51	.21	.55	.21	.55	.02			
Total	T	65	.64	.14	.67	.14	.64	.01	1, 121	.89	.35
	C	59	.57	.16	.63	.15	.66	.01			

ANCOVA RESULTS

Listening Comprehension: Scores of the comparison school were significantly above those of the treatment school

Vocabulary: No difference between the treatment and comparison schools

Sentence Comprehension: The comparison school outperformed the treatment school, with the difference approaching significance

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 21a: GRADE – Fourth Grade – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.20	1	.20	22.56**	.16
	Condition	.04	1	.04	4.50*	.04
	Error	1.03	116	.01		
Vocabulary	Pretest	2.15	1	2.15	143.79**	.56
	Condition	.03	1	.03	1.97	.02
	Error	1.72	115	.02		
Sentence comprehension	Pretest	1.50	1	1.50	65.66**	.37
	Condition	.07	1	.07	2.99	.03
	Error	2.51	110	.02		
Passage comprehension	Pretest	2.00	1	2.00	73.29**	.38
	Condition	.05	1	.05	1.64	.01
	Error	3.23	118	.03		
Total	Pretest	1.59	1	1.59	200.47**	.62
	Condition	.01	1	.01	.89	.01
	Error	.96	121	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 22: GRADE – Fifth Grade – Year 1

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	52	.88	.09	.84	.07	.84	.01	1, 109	.97	.33
	C	60	.83	.09	.82	.09	.82	.01			
Vocabulary	T	65	.46	.14	.50	.17	.48	.02	1, 121	.06	.80
	C	59	.40	.14	.47	.15	.49	.02			
Sentence comprehension	T	58	.66	.22	.73	.24	.71	.02	1, 113	1.93	.17
	C	58	.60	.24	.64	.23	.67	.02			
Passage comprehension	T	52	.60	.19	.67	.20	.65	.02	1, 105	.60	.44
	C	56	.52	.20	.60	.17	.63	.02			
Total	T	66	.61	.17	.63	.20	.63	.01	1, 124	.47	.50
	C	61	.59	.13	.63	.15	.64	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Vocabulary: Scores of the comparison school were significantly above those of the treatment school

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: No difference between the treatment and comparison schools

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 22a: GRADE – Fifth Grade – Year 1

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.03	1	.03	4.40*	.04
	Condition	.01	1	.01	.97	.01
	Error	.67	109	.01		
Vocabulary	Pretest	1.14	1	1.14	65.46**	.35
	Condition	.00	1	.00	.06	.00
	Error	2.11	121	.02		
Sentence comprehension	Pretest	3.52	1	3.52	146.40**	.56
	Condition	.05	1	.05	1.93	.02
	Error	2.72	113	.02		
Passage comprehension	Pretest	1.57	1	1.57	81.16**	.44
	Condition	.01	1	.01	.60	.01
	Error	2.03	105	.02		
Total	Pretest	2.52	1	2.52	210.15**	.63
	Condition	.01	1	.01	.47	.00
	Error	1.48	124	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 23: GRADE – Fifth Grade – Year 2

Assessment	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct				
			M	SD	M	SD	M	SE	df	F	p
Listening comprehension	T	50	.83	.08	.81	.11	.81	.01	1, 108	.75	.39
	C	61	.85	.08	.83	.09	.83	.01			
Vocabulary	T	80	.44	.14	.49	.15	.48	.01	1, 138	6.00	.02*
	C	61	.41	.17	.51	.15	.53	.01			
Sentence comprehension	T	75	.57	.22	.68	.23	.69	.02	1, 133	.13	.72
	C	61	.59	.27	.70	.23	.70	.02			
Passage comprehension	T	78	.56	.19	.64	.18	.64	.02	1, 136	7.08	.01*
	C	61	.56	.20	.57	.25	.57	.02			
Total	T	82	.57	.15	.63	.16	.64	.01	1, 141	.04	.84
	C	62	.60	.16	.65	.16	.64	.01			

ANCOVA RESULTS

Listening Comprehension: No difference between the treatment and comparison schools

Vocabulary: Scores of the comparison school were significantly above those of the treatment school

Sentence Comprehension: No difference between the treatment and comparison schools

Passage Comprehension: Scores of the treatment school were significantly above those of the comparison school

Total: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 23a: GRADE – Fifth Grade – Year 2

Analysis of Covariance Summary

Assessment	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Listening Comprehension	Pretest	.15	1	.15	16.97**	.14
	Condition	.01	1	.01	.75	.01
	Error	.92	108	.01		
Vocabulary	Pretest	1.68	1	1.68	172.04**	.55
	Condition	.06	1	.06	6.00*	.04
	Error	1.35	138	.01		
Sentence comprehension	Pretest	3.13	1	3.13	107.95**	.45
	Condition	.00	1	.00	.13	.00
	Error	3.86	133	.03		
Passage comprehension	Pretest	3.16	1	3.16	139.00**	.51
	Condition	.16	1	.16	7.08**	.05
	Error	3.09	136	.02		
Total	Pretest	2.44	1	2.44	325.05**	.70
	Condition	.00	1	.00	.04	.00
	Error	1.06	141	.01		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 24: UEE Comprehension Assessment – All Grades – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	249	.55	.17	.65	.19	.63	.01	1, 428	.62	.43
	C	182	.48	.18	.60	.20	.62	.01			
Year 2	T	205	.52	.18	.66	.21	.64	.01	1, 411	.05	.83
	C	209	.44	.22	.63	.19	.65	.01			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: Scores of the treatment school were significantly above those of the comparison school

APPENDIX B (continued)

Table 24a: UEE Comprehension Assessment – All Grades – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	3.17	1	3.17	102.60**	.19
	Condition	.02	1	.02	.62	.00
	Error	13.24	428	.03		
Year 2	Pretest	2.28	1	2.28	67.41**	.14
	Condition	.00	1	.00	.05	.00
	Error	13.87	411	.03		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 25: UEE Comprehension Assessment – Grades 3-5 – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	197	.56	.17	.66	.18	.64	.01	1, 349	1.18	.28
	C	155	.49	.19	.60	.20	.62	.01			
Year 2	T	145	.53	.19	.62	.22	.62	.02	1, 286	1.29	.26
	C	144	.51	.19	.64	.20	.64	.02			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: Scores of the treatment school were significantly above those of the comparison school

APPENDIX B (continued)

Table 25a: UEE Comprehension Assessment – Grades 3-5 – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	2.31	1	2.31	76.07**	.18
	Condition	.04	1	.04	1.18	.00
	Error	10.58	349	.03		
Year 2	Pretest	2.62	1	2.62	75.51**	.21
	Condition	.05	1	.05	1.29	.00
	Error	9.91	286	.04		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 26: UEE Listening Comprehension Assessment (18 items) – Second Grade – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	52	.50	.14	.61	.22	.60	.03	1, 76	.13	.72
	C	27	.45	.12	.58	.19	.61	.04			
Year 2	T	60	.51	.16	.74	.15	.72	.02	1, 122	6.46	.01*
	C	65	.29	.20	.62	.16	.64	.02			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: Scores of the treatment school were significantly above those of the comparison school

APPENDIX B (continued)

Table 26a: UEE Listening Comprehension Assessment – Second Grade – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	.93	1	.93	28.51**	.27
	Condition	.00	1	.00	.13	.00
	Error	2.49	76	.03		
Year 2	Pretest	.17	1	.17	7.26**	.06
	Condition	.15	1	.15	6.46*	.05
	Error	2.82	122	.02		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 27: UEE Reading Comprehension Assessment (20 items) – Third Grade – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	56	.49	.19	.65	.18	.64	.02	1, 111	.21	.65
	C	58	.44	.20	.61	.20	.63	.02			
Year 2	T	46	.55	.22	.66	.20	.63	.03	1, 92	.02	.88
	C	49	.46	.19	.60	.19	.63	.02			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 27a: UEE Reading Comprehension Assessment – Third Grade – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	1.61	1	1.61	67.23**	.38
	Condition	.01	1	.01	.21	.00
	Error	2.65	111	.02		
Year 2	Pretest	1.02	1	1.02	37.93**	.29
	Condition	.00	1	.00	.02	.00
	Error	2.48	92	.03		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 28: UEE Reading Comprehension Assessment (20 items) – Fourth Grade – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	70	.53	.17	.73	.15	.71	.02	1, 116	.93	.34
	C	49	.46	.16	.66	.20	.68	.02			
Year 2	T	59	.51	.19	.70	.20	.70	.02	1, 107	.61	.44
	C	51	.49	.18	.72	.19	.72	.03			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: No difference between the treatment and comparison schools

APPENDIX B (continued)

Table 28a: UEE Reading Comprehension Assessment – Fourth Grade – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	1.00	1	1.00	46.40**	.29
	Condition	.02	1	.02	.93	.01
	Error	2.50	116	.02		
Year 2	Pretest	.81	1	.81	25.40**	.19
	Condition	.02	1	.02	.61	.01
	Error	3.40	107	.03		

*p<0.05

**p < 0.01

APPENDIX B (continued)

Table 29: UEE Reading Comprehension Assessment (20 items) – Fifth Grade – Years 1 and 2

Year	School	n	Pretest		Posttest				ANCOVA		
			Raw proportion correct		Raw proportion correct		Adjusted proportion correct		df	F	p
			M	SD	M	SD	M	SE			
Year 1	T	71	.65	.12	.60	.18	.57	.02	1, 116	.01	.94
	C	48	.58	.17	.54	.20	.57	.02			
Year 2	T	40	.48	.19	.48	.19	.49	.02	1, 81	4.90	.03*
	C	44	.58	.20	.58	.20	.57	.02			

ANCOVA: Year 1: No difference between the treatment and comparison schools

Year 2: Scores of the comparison school were significantly above those of the treatment school

APPENDIX B (continued)

Table 29a: UEE Reading Comprehension Assessment – Fifth Grade – Years 1 and 2

Analysis of Covariance Summary

Year	Source	Type III Sum of Squares	df	Mean Square	F	Partial Eta Squared
Year 1	Pretest	1.17	1	1.17	45.43**	.28
	Condition	.00	1	.00	.01	.00
	Error	2.99	116	.03		
Year 2	Pretest	1.18	1	1.18	51.14**	.39
	Condition	.11	1	.11	4.90*	.06
	Error	1.87	81	.02		

*p<0.05

**p < 0.01

APPENDIX C

Table 30: Number of instances of explicit instruction observed per lesson – Years 1 and 2

All grades

Year	Grade	School	# of observations	Number of instances of explicit instruction									
				Introduction		Reading		Practice		Feedback		Total	
				M	SD	M	SD	M	SD	M	SD	M	SD
1	All grades	T	20	5.25	2.79	10.25*	4.51	9.60*	6.38	3.85*	3.88	28.95*	10.90
1	All grades	C	22	5.86	4.65	6.41*	4.87	3.55*	3.05	1.41*	2.34	17.23*	5.25
2	All grades	T	20	5.20	4.91	7.60	4.71	12.70*	8.17	4.75[†]	4.99	30.25*	11.67
2	All grades	C	13	5.31	3.75	6.38	2.90	3.85*	3.39	1.77[†]	2.59	17.31*	6.36
Both	All grades	T	40	5.23	3.94	8.93*	4.74	11.15*	7.40	4.30*	4.44	29.60*	11.17
Both	All grades	C	35	5.66	4.29	6.40*	4.20	3.66*	3.13	1.54*	2.41	17.26*	5.60

* Difference is significant at the 0.05 level

[†] Difference is significant at the 0.10 level

APPENDIX C (continued)

Table 31: Number of instances of explicit instruction observed per lesson – Years 1 and 2

Grades K-2

Year	Grade	School	# of observations	Number of instances of explicit instruction									
				Introduction		Reading		Practice		Feedback		Total	
				M	SD	M	SD	M	SD	M	SD	M	SD
1	K	T	4	3.75	2.50	11.75*	3.86	11.75*	7.27	4.50*	2.38	31.75*	13.94
1	K	C	6	7.50	4.76	5.33*	3.44	3.33*	2.94	0.50*	1.22	16.67*	6.22
2	K	T	3	5.33	6.81	4.00	5.29	12.00	17.32	0.33	0.58	21.67	10.69
2	K	C	4	6.50	5.26	4.00	3.37	1.75	1.50	1.00	2.00	13.25	6.34
1	1	T	4	6.25	2.22	9.00	4.32	14.50*	4.20	6.50*	1.73	36.25*	5.44
1	1	C	8	7.00	5.40	8.25	5.68	3.38*	3.29	1.25*	2.43	19.88*	4.61
2	1	T	4	8.25	8.18	8.75	4.43	7.75	7.37	4.50	7.14	29.25	16.32
2	1	C	2	2.50	0.71	9.00	1.41	3.50	4.95	0.00	0.00	15.00	2.83
1	2	T	5	3.80	2.05	13.40	4.56	11.20*	3.56	5.20	5.54	33.60*	9.37
1	2	C	3	2.67	3.79	8.67	7.23	2.00*	1.73	0.00	0.00	13.33*	4.51
2	2	T	3	4.00	3.46	11.00	6.24	20.67*	1.53	7.33*	1.53	43.00*	5.20
2	2	C	2	7.00	4.24	6.50	0.71	4.50*	2.12	1.50*	2.12	19.50*	9.19

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX C (continued)

Table 32: Number of instances of explicit instruction observed per lesson – Years 1 and 2

Grades 3-5

Year	Grade	School	# of lessons	Number of instances of explicit instruction									
				Introduction		Reading		Practice		Feedback		Total	
				M	SD	M	SD	M	SD	M	SD	M	SD
1	3	T	3	9.00[†]	3.60	5.33	3.06	2.67	4.62	0.00*	0.00	17.00	4.36
1	3	C	2	2.00[†]	0.00	3.50	0.71	7.00	4.24	5.00*	2.83	17.50	0.71
2	3	T	2	8.00	5.66	11.00	4.24	13.00	2.83	7.50	10.61	39.50*	2.12
2	3	C	2	6.00	2.83	9.00	0.00	6.00	2.83	2.00	2.83	23.00*	2.83
1	4	T	2	4.00	1.41	7.00	1.41	6.50	9.19	3.50	4.95	21.00	14.14
1	4	C	2	6.00	4.24	3.50	3.54	5.50	0.71	3.50	3.54	18.50	2.12
2	4	T	5	3.60	1.14	8.60[†]	1.67	13.00	6.67	5.20	4.09	30.40	9.34
2	4	C	1	1.00		4.00[†]		4.00		4.00		13.00	
1	5	T	2	5.50	0.71	12.50	3.54	5.00	7.07	0.00	0.00	23.00	4.24
1	5	C	1	4.00		3.00		0.00		1.00		8.00	
2	5	T	3	3.00	3.00	2.33	2.08	11.33	3.06	4.33	4.51	21.00	2.65
2	5	C	2	5.50	3.54	7.00	2.83	5.50	7.78	4.00	5.66	22.00	7.07

* Difference is significant at the 0.05 level

[†] Difference is significant at the 0.10 level

APPENDIX D

Table 33: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 1: All Cases (N=27 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.36 [†]									
	Student Practice	-.15	.37 [†]								
	Feedback	-.31	.19	.64*							
	Total	-.02	.64*	.89*	.64*						
Adjusted mean GRADE posttest score	Listening Comp	.04	.07	.40*	.22	.34 [†]					
	Word Reading / Vocabulary	-.04	.17	.22	.24	.26	.47*				
	Sentence Comp	-.12	-.17	-.02	.03	-.14	.29	.29			
	Passage Comp	-.45*	.16	.18	-.06	.00	.29	.32	.33		
	Total	-.07	.04	.17	.14	.13	.67*	.80*	.76*	.60*	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 34: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 2: All Cases (N=22 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	.15									
	Student Practice	-.10	.24								
	Feedback	-.14	.01	.40 [†]							
	Total	.27	.55*	.81*	.58*						
Adjusted mean GRADE posttest score	Listening Comp	.14	.03	-.03	-.23	-.04					
	Word Reading / Vocabulary	.28	.19	.07	-.17	.14	.67*				
	Sentence Comp	.43 [†]	.39	-.10	-.09	.15	.06	.09			
	Passage Comp	.20	-.05	.34	.47 [†]	.43 [†]	.16	.04	.19		
	Total	.23	.01	-.01	-.24	-.02	.75*	.85*	.42 [†]	.26	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 35: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 1: Treatment School (N=14 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.34									
	Student Practice	-.32	.54*								
	Feedback	-.38	.41	.73*							
	Total	-.18	.73*	.93*	.77*						
Adjusted mean GRADE posttest score	Listening Comp	-.34	.54 [†]	.69*	.51 [†]	.67*					
	Word Reading / Vocabulary	-.28	.25	.52 [†]	.78*	.55*	.74*				
	Sentence Comp	.12	-.29	-.23	-.04	-.24	.31	.57 [†]			
	Passage Comp	-.52	.67*	.36	.15	.39	.71*	.38	.21		
	Total	-.28	.25	.35	.52 [†]	.38	.77*	.87*	.87*	.59 [†]	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 36: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 2: Treatment School (N=13 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	.22									
	Student Practice	-.12	.17								
	Feedback	-.10	-.01	.15							
	Total	.35	.61*	.68*	.48 [†]						
Adjusted mean GRADE posttest score	Listening Comp	.32	.01	-.16	-.48	-.18					
	Word Reading / Vocabulary	.36	.32	.27	-.06	.40	.78*				
	Sentence Comp	.67*	.55 [†]	.06	.01	.46	-.18	-.10			
	Passage Comp	.55 [†]	.12	-.06	.19	.34	.04	.13	.31		
	Total	.34	.18	.16	-.27	.18	.86*	.84*	.22	.25	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 37: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 1: Comparison School (N=13 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.48 [†]									
	Student Practice	-.05	-.32								
	Feedback	-.30	-.20	.44							
	Total	.15	.35	.49 [†]	.35						
Adjusted mean GRADE posttest score	Listening Comp	.33	-.26	.06	-.08	-.01					
	Word Reading / Vocabulary	.20	.32	-.22	-.54 [†]	.09	.18				
	Sentence Comp	-.39	-.25	.13	.14	-.47	.45	.27			
	Passage Comp	-.40	-.12	.03	-.28	-.65*	-.08	.26	.55 [†]		
	Total	.12	-.06	-.05	-.35	-.18	.60*	.72*	.87*	.63 [†]	

* Correlation is significant at the 0.05 level (2-tailed)

† Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 38: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 2: Comparison School (N=9 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading / Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.03									
	Student Practice	-.18	-.02								
	Feedback	-.32	-.40	.64 [†]							
	Total	.33	.22	.79*	.49						
Adjusted mean GRADE posttest score	Listening Comp	-.20	-.03	-.13	.20	-.11					
	Word Reading / Vocabulary	.14	-.13	-.27	-.44	-.35	.52				
	Sentence Comp	-.03	.24	.25	.03	.28	.40	.43			
	Passage Comp	-.28	-.70 [†]	.44	.77*	.30	.25	.00	.25		
	Total	.06	-.36	-.26	-.20	-.38	.64 [†]	.88*	.73 [†]	.33	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 39: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 1: Grades K-2 (N=15 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.60*									
	Student Practice	-.06	.44 [†]								
	Feedback	-.11	.32	.80*							
	Total	-.03	.59*	.95*	.82*						
Adjusted mean GRADE posttest score	Listening Comp	.50 [†]	-.28	.21	.07	.16					
	Word Reading	-.12	-.26	-.35	-.08	-.37	.06				
	Sentence Comp	-.10	-.17	.41	.23	.20	.25	.72*			
	Passage Comp	-.47	-.10	.01	-.17	-.27	-.25	.77*	.65 [†]		
	Total	.06	-.29	-.05	-.15	-.18	.44 [†]	.47 [†]	.90*	.77*	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 40: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 2: Grades K-2 (N=10 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Word Reading	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.20									
	Student Practice	-.09	.25								
	Feedback	.44	.36	.61 [†]							
	Total	.32	.53	.81*	.89*						
Adjusted mean GRADE posttest score	Listening Comp	-.06	-.23	.32	-.18	.04					
	Word Reading	.14	-.03	.54	.42	.46	-.17				
	Sentence Comp	.64	.43	-.49	-.49	-.15	-.36	-.04			
	Passage Comp	.77	.15	.37	.79	.68	-.49	.04	.52		
	Total	.14	-.52	.07	-.37	-.19	.23	.20	.80 [†]	.38	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 41: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 1: Grades 3-5 (N=12 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	-.02									
	Student Practice	-.33	-.12								
	Feedback	-.59*	-.26	.26							
	Total	.05	.52 [†]	.64*	.14						
Adjusted mean GRADE posttest score	Listening Comp	-.42	.10	.44	.25	.30					
	Vocabulary	.05	-.41	.04	.20	-.15	.47				
	Sentence Comp	-.14	-.08	-.21	-.03	-.33	.44	.76*			
	Passage Comp	-.43	.35	.35	-.02	.27	.73*	.18	.19		
	Total	-.18	-.22	-.07	.16	-.26	.77*	.90*	.90*	.49	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 42: Correlations between teachers' explicitness scores and the adjusted posttest GRADE scores of their students

Year 2: Grades 3-5 (N=12 teachers)

		Explicitness during observations					Adjusted mean GRADE posttest score				
		Intro	Direct Reading	Student Practice	Feedback	Total	Listening Comp	Vocabulary	Sentence Comp	Passage Comp	Total
Explicitness during observations	Intro										
	Direct Reading	.71*									
	Student Practice	-.09	.23								
	Feedback	-.54	-.24	.27							
	Total	.24	.60*	.82*	.40						
Adjusted mean GRADE posttest score	Listening Comp	.18	.31	-.12	-.06	.05					
	Vocabulary	.49	.66*	.05	.00	.41	.50				
	Sentence Comp	.35	.55 [†]	.18	-.02	.40	.48	.85*			
	Passage Comp	-.40	-.27	.35	.44	.22	.30	-.16	.09		
	Total	.23	.49 [†]	.11	.13	.38	.65*	.86*	.93*	.28	

* Correlation is significant at the 0.05 level (2-tailed)

[†] Correlation is significant at the 0.10 level (2-tailed)

APPENDIX D (continued)

Table 43: Comparison of High-Explicit vs. Low-Explicit Teachers (split at the median score*)

Year 1

Assessment	Explicitness	Number of teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Low	14	.80	.10	1, 24	1.33	.26
	High	12	.83	.05			
Word Reading / Vocabulary	Low	14	.71	.16	1, 25	.09	.77
	High	13	.73	.20			
Sentence comprehension	Low	12	.75	.13	1, 18	1.37	.26
	High	8	.68	.13			
Passage comprehension	Low	12	.55	.10	1, 18	.08	.79
	High	8	.56	.07			
Total	Low	14	.72	.10	1, 25	.00	.99
	High	13	.72	.13			

* High-explicit teachers averaged more than 19 instances of explicit instruction per lesson. Low-explicit teachers averaged 19 or fewer instances of explicit instruction per lesson.

APPENDIX D (continued)

Table 44: Comparison of High-Explicit vs. Low-Explicit Teachers (split at the median score*)

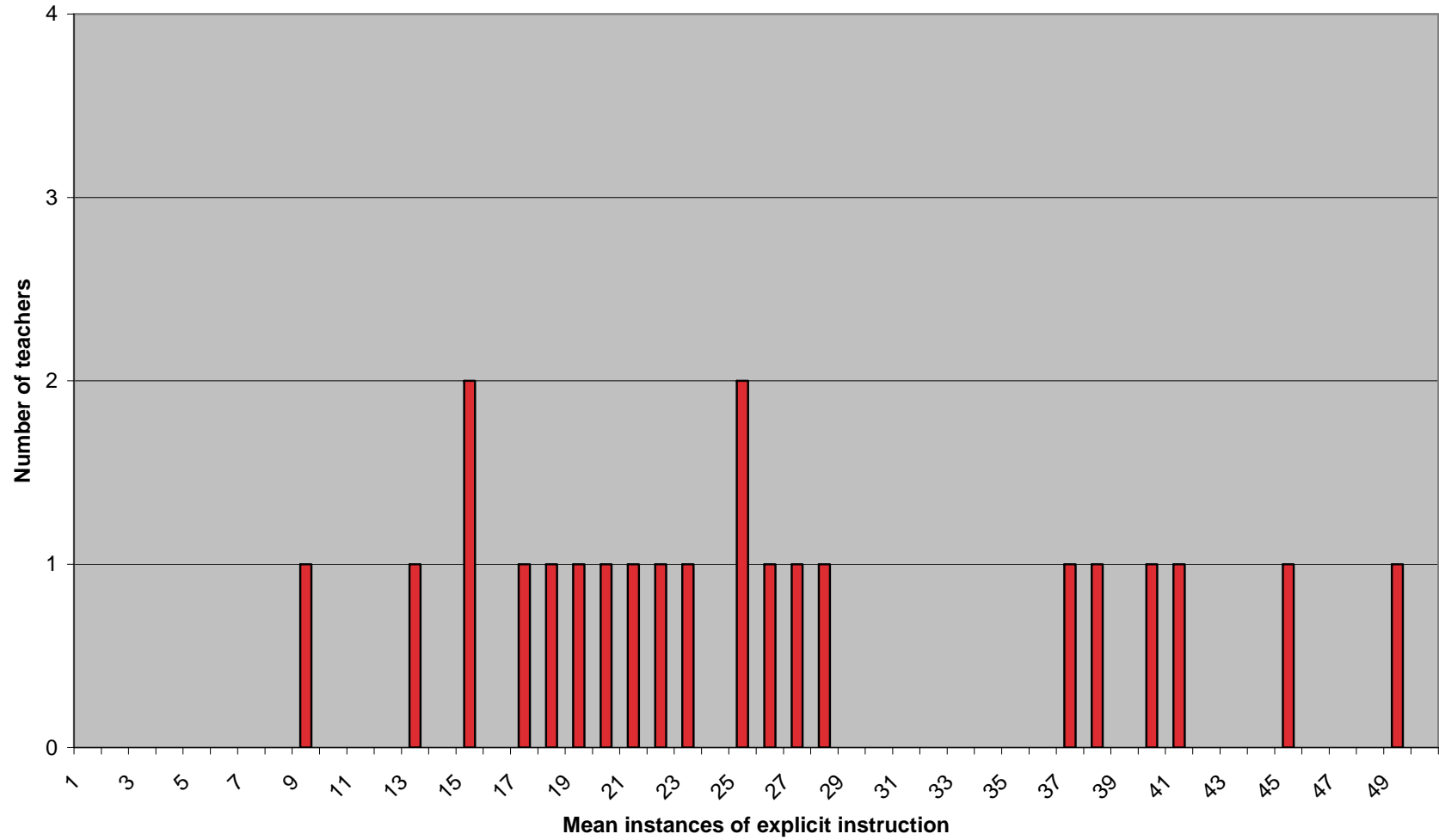
Year 2

Assessment	Explicitness	Number of teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Low	10	.85	.06	1, 19	.24	.63
	High	11	.84	.06			
Word Reading / Vocabulary	Low	11	.70	.18	1, 20	.06	.80
	High	11	.72	.17			
Sentence comprehension	Low	8	.73	.09	1, 15	.12	.74
	High	9	.75	.10			
Passage comprehension	Low	8	.56	.09	1, 15	1.05	.32
	High	9	.60	.08			
Total	Low	11	.74	.10	1, 20	.03	.88
	High	11	.73	.09			

* High-explicit teachers averaged more than 25 instances of explicit instruction per lesson. Low-explicit teachers averaged fewer than 25 instances of explicit instruction per lesson.

APPENDIX D (continued)

**Figure 2: Frequencies of Mean Instances of Explicit Instruction (All teachers)
Year 2**



APPENDIX D (continued)

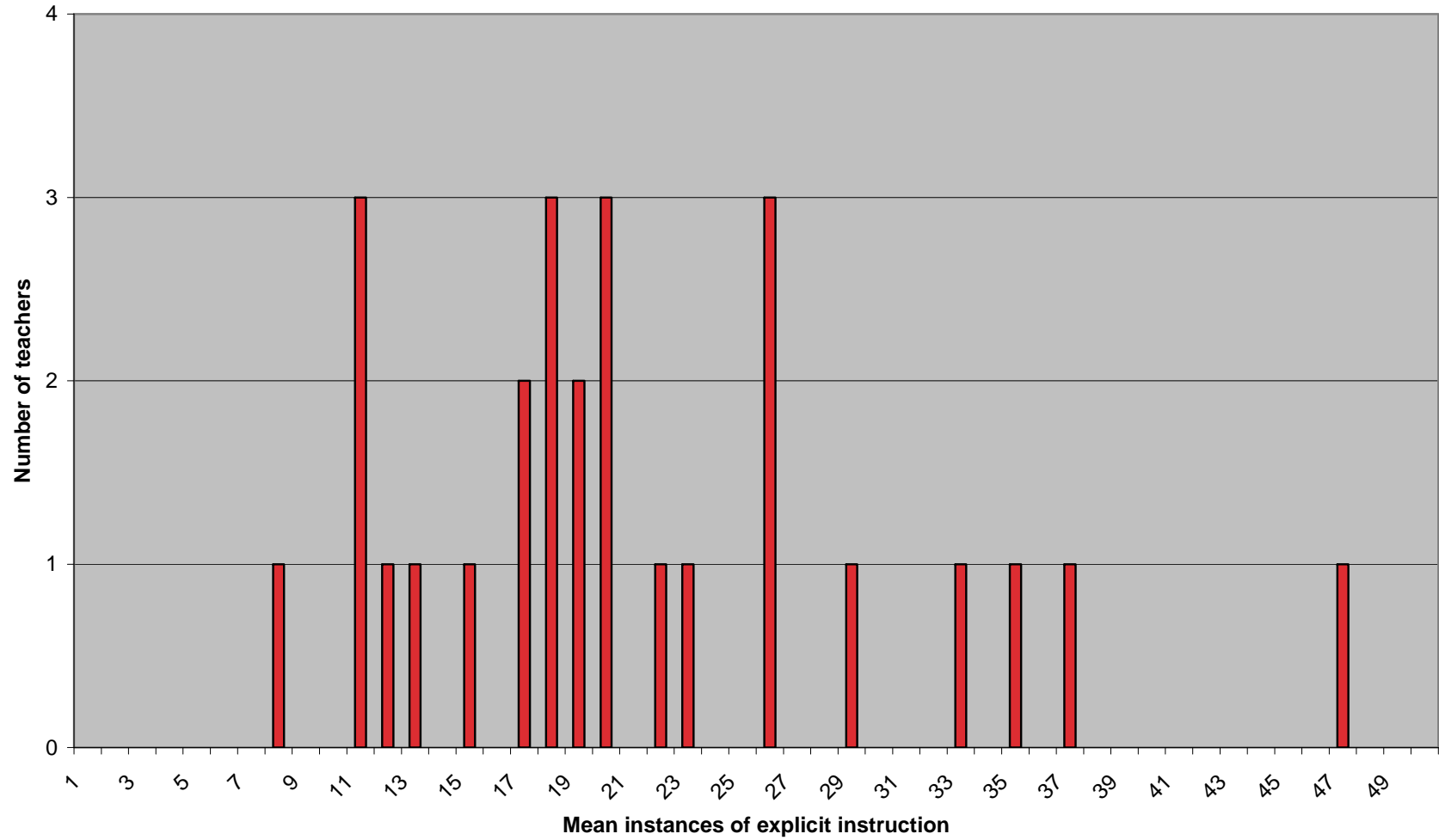
Table 45: Comparison of teachers that averaged more than 35 instances of explicit instruction vs. teachers that averaged fewer than 35 instances of explicit instruction (Split at empirical gap in scores*)
Year 2

Assessment	Explicitness	n	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Fewer than 35	15	.84	.06	1, 19	.01	.98
	More than 35	6	.85	.05			
Word Reading / Vocabulary	Fewer than 35	16	.68	.17	1, 20	1.62	.22
	More than 35	6	.79	.16			
Sentence comprehension	Fewer than 35	12	.73	.08	1, 15	.27	.61
	More than 35	5	.76	.13			
Passage comprehension	Fewer than 35	12	.56	.09	1, 15	1.87	.19
	More than 35	5	.62	.06			
Total	Fewer than 35	16	.73	.10	1, 20	.15	.71
	More than 35	6	.75	.06			

* High-explicit teachers averaged more than 35 instances of explicit instruction per lesson. Average-explicit teachers averaged fewer than 35 instances of explicit instruction per lesson.

APPENDIX D (continued)

**Figure 3: Frequencies of Mean Instances of Explicit Instruction (All teachers)
Year 1**



APPENDIX D (continued)

Table 46: Comparison of teachers that averaged more than 25 instances of explicit instruction vs. teachers that averaged fewer than 25 instances of explicit instruction (Split at empirical gap in scores*)
Year 1

Assessment	Explicitness	n	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Fewer than 25	19	.80	.09	1, 24	2.36	.14
	More than 25	7	.85	.05			
Word Reading / Vocabulary	Fewer than 25	19	.70	.16	1, 25	.77	.39
	More than 25	8	.76	.20			
Sentence comprehension	Fewer than 25	16	.73	.13	1, 18	.26	.62
	More than 25	4	.69	.15			
Passage comprehension	Fewer than 25	16	.55	.10	1, 18	.11	.75
	More than 25	4	.57	.06			
Total	Fewer than 25	19	.71	.10	1, 25	.14	.71
	More than 25	8	.73	.14			

* High-explicit teachers averaged more than 25 instances of explicit instruction per lesson. Average-explicit teachers averaged fewer than 25 instances of explicit instruction per lesson.

APPENDIX D (continued)

Table 47: Comparison of High-Explicit vs. Low-Explicit Teachers in the Treatment School (Year 1)*

Assessment	Explicitness	Number of teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Low	7	.77	.08	1, 11	5.13	.05*
	High	6	.86	.05			
Word Reading / Vocabulary	Low	7	.63	.16	1, 12	1.37	.27
	High	7	.74	.21			
Sentence comprehension	Low	7	.80	.06	1, 8	3.20	.11
	High	7	.69	.15			
Passage comprehension	Low	7	.53	.10	1, 8	.53	.49
	High	7	.57	.06			
Total	Low	7	.70	.10	1, 12	.01	.94
	High	7	.71	.13			

* High-explicit teachers averaged more than 25.75 instances of explicit instruction per lesson. Low-explicit teachers averaged fewer than 25.75 instances of explicit instruction per lesson.

APPENDIX D (continued)

Table 48: Comparison of High-Explicit vs. Low-Explicit Teachers in the Treatment School (Year 2)*

Assessment	Explicitness	Number of teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	Low	6	.86	.08	1, 10	.10	.76
	High	6	.85	.05			
Word Reading / Vocabulary	Low	7	.63	.17	1, 11	2.62	.13
	High	6	.79	.16			
Sentence comprehension	Low	5	.68	.03	1, 8	1.67	.23
	High	5	.76	.13			
Passage comprehension	Low	5	.59	.06	1, 8	.75	.41
	High	5	.62	.06			
Total	Low	7	.72	.11	1, 11	.40	.54
	High	6	.75	.06			

* High-explicit teachers averaged more than 28 instances of explicit instruction per lesson. Low-explicit teachers averaged 28 or fewer instances of explicit instruction per lesson.

APPENDIX E

Table 49: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year

Observations (Treatment School)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	7	5.33	3.14	1, 12	.05	.83
	Second	7	5.79	4.36			
Direct Reading	First	7	9.10	3.10	1, 12	.04	.84
	Second	7	8.64	5.01			
Student Practice	First	7	6.79	6.78	1, 12	5.63	.04*
	Second	7	15.36	6.74			
Feedback	First	7	3.14	3.06	1, 12	5.12	.04*
	Second	7	7.64	4.28			
Total	First	7	24.36	10.64	1, 12	6.40	.03*
	Second	7	37.43	8.58			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX E (continued)

Table 50: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year
Observations (Treatment School, Grades K-2)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	4	3.83	1.35	1, 6	2.03	.21
	Second	4	7.63	5.15			
Direct Reading	First	4	10.92	2.83	1, 6	.37	.57
	Second	4	8.63	7.02			
Student Practice	First	4	11.88	3.35	1, 6	1.24	.31
	Second	4	16.00	6.60			
Feedback	First	4	5.50	1.22	1, 6	.00	1.00
	Second	4	5.50	3.44			
Total	First	4	32.13	4.53	1, 6	1.07	.34
	Second	4	37.75	9.91			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX E (continued)

Table 51: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year
Observations (Treatment School, Grades 3-5)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	3	7.33	4.04	1, 4	2.72	.18
	Second	3	3.33	1.15			
Direct Reading	First	3	6.67	1.15	1, 4	4.50	.10
	Second	3	8.67	1.15			
Student Practice	First	3	.00	.00	1, 4	9.11	.04*
	Second	3	14.50	8.32			
Feedback	First	3	.00	.00	1, 4	21.00	.01*
	Second	3	10.50	3.97			
Total	First	3	14.00	5.20	1, 4	15.87	.02*
	Second	3	37.00	8.54			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX E (continued)

Table 52: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year

Observations (Comparison School)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	6	4.39	2.83	1, 10	.14	.72
	Second	6	5.08	3.61			
Direct Reading	First	6	4.94	3.92	1, 10	.78	.40
	Second	6	6.67	2.75			
Student Practice	First	6	3.58	3.44	1, 10	.47	.51
	Second	6	5.00	3.73			
Feedback	First	6	1.83	2.79	1, 10	.08	.78
	Second	6	2.33	3.20			
Total	First	6	14.75	3.92	1, 10	1.89	.20
	Second	6	19.08	6.67			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX E (continued)

Table 53: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year

Observations (Comparison School, Grades K-2)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	3	6.11	3.13	1, 4	.08	.80
	Second	3	5.17	5.06			
Direct Reading	First	3	6.56	5.50	1, 4	.06	.82
	Second	3	5.67	3.25			
Student Practice	First	3	2.50	.87	1, 4	.05	.84
	Second	3	2.33	1.04			
Feedback	First	3	.00	.00	1, 4	1.00	.37
	Second	3	.67	1.15			
Total	First	3	15.17	2.75	1, 4	.20	.68
	Second	3	13.83	4.37			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX E (continued)

Table 54: Comparison of Teachers using the Curriculum for their First Year vs. Their Second Year

Observations (Comparison School, Grades 3-5)

Assessment	First or Second Year	Number of Teachers	Instances of Explicit Instruction		ANOVA		
			M	SD	df	F	p
Introduction	First	3	2.67	1.15	1, 4	1.96	.23
	Second	3	5.00	2.65			
Direct Reading	First	3	3.33	.58	1, 4	9.94	.03*
	Second	3	7.67	2.31			
Student Practice	First	3	4.67	5.03	1, 4	.72	.45
	Second	3	7.67	3.51			
Feedback	First	3	3.67	3.06	1, 4	.01	.91
	Second	3	4.00	4.00			
Total	First	3	14.33	5.51	1, 4	7.56	.05*
	Second	3	24.33	3.06			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F

Table 55: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year

Assessments (Treatment School)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	7	.80	.10	1, 12	.91	.36
	Second	7	.85	.07			
Word Reading / Vocabulary	First	7	.74	.21	1, 12	.13	.73
	Second	7	.78	.18			
Sentence comprehension	First	5	.78	.02	1, 8	1.12	.32
	Second	5	.73	.11			
Passage comprehension	First	6	.53	.08	1, 9	5.56	.04*
	Second	5	.63	.06			
Total	First	7	.72	.10	1, 12	.58	.46
	Second	7	.76	.09			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F (continued)

Table 56: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year Assessments (Treatment School, Grades K-2)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	4	.88	.03	1, 6	.71	.43
	Second	4	.90	.04			
Word Reading	First	4	.90	.05	1, 6	.19	.68
	Second	4	.91	.04			
Sentence comprehension	First	2	.78	.01	1, 2	.68	.50
	Second	2	.67	.19			
Passage comprehension	First	3	.60	.01	1, 3	26.80	.01*
	Second	2	.68	.03			
Total	First	4	.79	.07	1, 6	.31	.60
	Second	4	.82	.08			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F (continued)

Table 57: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year
Assessments (Treatment School, Grades 3-5)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	3	.70	.07	1, 4	3.37	.14
	Second	3	.78	.02			
Vocabulary	First	3	.52	.11	1, 4	.65	.47
	Second	3	.59	.10			
Sentence comprehension	First	3	.78	.03	1, 4	.26	.64
	Second	3	.76	.04			
Passage comprehension	First	3	.47	.04	1, 4	11.09	.03*
	Second	3	.60	.06			
Total	First	3	.63	.06	1, 4	2.25	.21
	Second	3	.68	.03			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F (continued)

Table 58: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year Assessments (Comparison School)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	6	.87	.06	1, 10	.10	.76
	Second	6	.85	.06			
Word Reading / Vocabulary	First	6	.74	.12	1, 10	.03	.87
	Second	6	.75	.15			
Sentence comprehension	First	4	.79	.10	1, 6	.01	.92
	Second	4	.80	.08			
Passage comprehension	First	4	.57	.08	1, 6	.25	.63
	Second	4	.53	.11			
Total	First	6	.77	.08	1, 10	.00	.97
	Second	6	.77	.09			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F (continued)

Table 59: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year
Assessments (Comparison School, Grades K-2)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	3	.88	.05	1, 4	.04	.85
	Second	3	.87	.02			
Word Reading	First	3	.84	.01	1, 4	4.48	.10
	Second	3	.88	.03			
Sentence comprehension	First	1	.69	--	1, 0	--	--
	Second	1	.75	--			
Passage comprehension	First	1	.50	--	1, 0	--	--
	Second	1	.42	--			
Total	First	3	.81	.08	1, 4	.27	.63
	Second	3	.84	.05			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX F (continued)

Table 60: Comparison of Students' Achievement of Teachers using the Curriculum for their First Year vs. Their Second Year
Assessments (Comparison School, Grades 3-5)

Assessment	First or Second Year	Number of Teachers	Adjusted proportion correct		ANOVA		
			M	SD	df	F	p
Listening comprehension	First	3	.85	.07	1, 4	.05	.84
	Second	3	.84	.08			
Vocabulary	First	3	.64	.08	1, 4	.06	.82
	Second	3	.63	.08			
Sentence comprehension	First	3	.83	.09	1, 4	.02	.89
	Second	3	.82	.10			
Passage comprehension	First	3	.59	.09	1, 4	.06	.82
	Second	3	.57	.10			
Total	First	3	.73	.06	1, 4	.24	.65
	Second	3	.71	.07			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX G

Table 61: Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers

All grades (Using the Year 2 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	263	.85	.01	1, 485	.46	.50
	Comparison	225	.86	.01			
Word Reading / Vocabulary	Treatment	260	.73	.01	1, 479	.03	.87
	Comparison	222	.73	.01			
Sentence comprehension	Treatment	215	.74	.01	1, 393	7.04	.01*
	Comparison	181	.79	.01			
Passage comprehension	Treatment	203	.62	.01	1, 391	7.10	.01*
	Comparison	191	.57	.01			
Total	Treatment	274	.75	.01	1, 512	.16	.69
	Comparison	241	.75	.01			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX G (continued)

Table 62: Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers
Students in Grades K - 2 in Year 2 (Using the Year 2 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	133	.89	.01	1, 234	.24	.62
	Comparison	104	.89	.01			
Word Reading	Treatment	132	.90	.01	1, 227	.00	.95
	Comparison	98	.90	.01			
Sentence comprehension	Treatment	86	.71	.03	1, 143	2.33	.13
	Comparison	60	.77	.03			
Passage comprehension	Treatment	77	.62	.02	1, 139	5.36	.02*
	Comparison	65	.54	.02			
Total	Treatment	140	.81	.01	1, 246	.18	.67
	Comparison	109	.81	.01			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX G (continued)

Table 63: Comparison of Treatment Students with Experienced Teachers vs. Comparison Students with Experienced Teachers
 Students in Grades 3 - 5 in Year 2 (Using the Year 2 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	130	.81	.01	1, 248	4.31	.04*
	Comparison	121	.84	.01			
Vocabulary	Treatment	128	.58	.01	1, 249	.00	.97
	Comparison	124	.58	.01			
Sentence comprehension	Treatment	129	.76	.01	1, 247	5.29	.02*
	Comparison	121	.80	.01			
Passage comprehension	Treatment	126	.62	.01	1, 249	1.47	.23
	Comparison	126	.59	.01			
Total	Treatment	134	.69	.01	1, 263	1.41	.24
	Comparison	132	.70	.01			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX H

Table 64: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students

All grades (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	174	.84	.01	1, 327	.60	.44
	Comparison	156	.85	.01			
Word Reading / Vocabulary	Treatment	198	.68	.01	1, 341	.02	.89
	Comparison	146	.67	.02			
Sentence comprehension	Treatment	167	.74	.02	1, 286	.25	.62
	Comparison	122	.75	.02			
Passage comprehension	Treatment	163	.60	.01	1, 287	.23	.63
	Comparison	127	.60	.02			
Total	Treatment	213	.71	.01	1, 385	.88	.35
	Comparison	175	.72	.01			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX H (continued)

Table 65: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (K-2):
Students in Grades 1 and 2 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	57	.87	.01	1, 109	.18	.68
	Comparison	55	.86	.01			
Word Reading	Treatment	57	.93	.02	1, 101	5.22	.02*
	Comparison	47	.88	.02			
Sentence comprehension	Treatment	39	.72	.04	1, 63	.87	.36
	Comparison	27	.67	.05			
Passage comprehension	Treatment	27	.62	.03	1, 51	3.42	.07 [†]
	Comparison	27	.54	.03			
Total	Treatment	62	.81	.02	1, 121	1.40	.24
	Comparison	61	.78	.02			

* Difference is significant at the 0.05 level

[†] Difference is significant at the 0.10 level

APPENDIX H (continued)

Table 66: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (3-5):
Students in Grades 4 and 5 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	78	.80	.01	1, 151	2.22	.14
	Comparison	75	.83	.01			
Vocabulary	Treatment	104	.57	.02	1, 174	.16	.69
	Comparison	72	.56	.02			
Sentence comprehension	Treatment	101	.74	.02	1, 168	.01	.91
	Comparison	69	.75	.02			
Passage comprehension	Treatment	99	.66	.02	1, 171	1.37	.24
	Comparison	74	.62	.03			
Total	Treatment	111	.67	.01	1, 190	.17	.68
	Comparison	81	.68	.02			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX H (continued)

Table 67: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students in Each Grade

(Using the Year 1 pretests as a covariate)

Total GRADE score	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Students in Grade 1 in Year 2	Treatment	20	.79	.03	1, 49	.41	.52
	Comparison	31	.82	.03			
Students in Grade 2 in Year 2	Treatment	42	.82	.02	1, 70	9.51	.00*
	Comparison	30	.72	.02			
Students in Grade 3 in Year 2	Treatment	40	.79	.02	1, 71	.69	.41
	Comparison	33	.77	.02			
Students in Grade 4 in Year 2	Treatment	52	.68	.02	1, 87	.21	.65
	Comparison	37	.67	.03			
Students in Grade 5 in Year 2	Treatment	59	.66	.02	1, 101	1.09	.30
	Comparison	44	.69	.02			
All Students	Treatment	213	.71	.01	1, 385	.88	.35
	Comparison	175	.72	.01			

* Difference is significant at the 0.01 level

APPENDIX H (continued)

Table 68: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students

Students in Grade 2 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	38	.87	.01	1, 62	4.90	.03*
	Comparison	26	.82	.02			
Word Reading	Treatment	39	.96	.02	1, 64	10.59	.00*
	Comparison	27	.87	.02			
Sentence comprehension	Treatment	39	.79	.04	1, 64	.34	.56
	Comparison	27	.76	.05			
Passage comprehension	Treatment	27	.67	.03	1, 51	3.73	.06 [†]
	Comparison	26	.57	.04			
Total	Treatment	42	.82	.02	1, 70	9.51	.00*
	Comparison	30	.72	.02			

* Difference is significant at the 0.05 level

[†] Difference is significant at the 0.10 level

APPENDIX I

The first analyses included students that were in kindergarten in Year One (and, therefore, in first grade in Year Two). There were a total of 20 K-1 students in the treatment school and 31 K-1 students in the comparison school who completed the assessments in both Year 1 and Year 2 of the study. The overall results of the ANCOVAs can be found in Table 51. There were no significant differences between the adjusted posttest scores of the two-year K-1 treatment students and the two-year K-1 comparison students on any of the GRADE subtests or on the total GRADE score.

The next analyses included students that were in second grade in Year One (and, therefore, in third grade in Year Two). There were a total of forty 2-3 students in the treatment school and thirty-three 2-3 students in the comparison school. The overall results of the ANCOVAs can be found in Table 53. There were no significant differences between the adjusted posttest scores of the two-year 2-3 treatment students and the two-year 2-3 comparison students on any of the GRADE subtests or on the total GRADE score.

Then, analyses were conducted that included students that were in third grade in Year One (and, therefore, in fourth grade in Year Two). There were a total of fifty-two 3-4 students in the treatment school and thirty-seven 3-4 students in the comparison school. The overall results of the ANCOVAs can be found in Table 54. There were no significant differences between the adjusted posttest scores of the two-year 3-4 treatment students and the two-year 3-4 comparison students on any of the GRADE subtests or on the total GRADE score.

The final analyses included students that were in fourth grade in Year One (and, therefore, in fifth grade in Year Two). There were a total of fifty-nine 4-5 students in the treatment school and forty-four 4-5 students in the comparison school. The overall results of the

ANCOVAs can be found in Table 55. There were no significant differences between the adjusted posttest scores of the two-year 4-5 treatment students and the two-year 4-5 comparison students on any of the GRADE subtests or on the total GRADE score.

APPENDIX I (continued)

Table 69: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Single Grade):

Students in Grade 1 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	18	.90	.02	1, 45	.03	.87
	Comparison	29	.90	.01			
Word Reading	Treatment	18	.91	.03	1, 36	1.21	.28
	Comparison	20	.94	.02			
Total	Treatment	20	.79	.03	1, 49	.41	.52
	Comparison	31	.82	.03			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX I (continued)

Table 70: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Single Grade):
Students in Grade 3 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	39	.87	.01	1, 63	.52	.47
	Comparison	26	.85	.02			
Vocabulary	Treatment	37	.74	.02	1, 62	.84	.36
	Comparison	27	.71	.03			
Sentence comprehension	Treatment	27	.93	.02	1, 51	.37	.55
	Comparison	26	.91	.02			
Passage comprehension	Treatment	37	.64	.03	1, 61	.04	.84
	Comparison	26	.65	.04			
Total	Treatment	40	.79	.02	1, 71	.69	.41
	Comparison	33	.77	.02			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX I (continued)

Table 71: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Single Grade):
Students in Grade 4 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	49	.80	.02	1, 81	.99	.32
	Comparison	34	.82	.02			
Vocabulary	Treatment	48	.61	.03	1, 76	1.80	.18
	Comparison	30	.55	.03			
Sentence comprehension	Treatment	51	.75	.03	1, 80	.03	.86
	Comparison	31	.74	.03			
Passage comprehension	Treatment	48	.64	.03	1, 80	.05	.82
	Comparison	34	.63	.04			
Total	Treatment	52	.68	.02	1, 87	.21	.65
	Comparison	37	.67	.03			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX I (continued)

Table 72: Comparison of Two-Year Treatment Students vs. Two-Year Comparison Students (Single Grade):

Students in Grade 5 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	29	.82	.02	1, 68	.55	.46
	Comparison	41	.83	.02			
Vocabulary	Treatment	56	.52	.02	1, 96	1.28	.26
	Comparison	42	.56	.02			
Sentence comprehension	Treatment	50	.74	.03	1, 86	.12	.74
	Comparison	38	.76	.03			
Passage comprehension	Treatment	51	.68	.03	1, 89	1.97	.16
	Comparison	40	.61	.03			
Total	Treatment	59	.66	.02	1, 101	1.09	.30
	Comparison	44	.69	.02			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX J

Table 73: Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers

All grades (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	146	.84	.01	1, 250	1.93	.17
	Comparison	107	.86	.01			
Word Reading / Vocabulary	Treatment	144	.73	.01	1, 239	4.98	.03*
	Comparison	98	.68	.02			
Sentence comprehension	Treatment	121	.76	.02	1, 196	1.14	.29
	Comparison	78	.79	.02			
Passage comprehension	Treatment	113	.63	.02	1, 194	.06	.80
	Comparison	84	.62	.02			
Total	Treatment	159	.73	.01	1, 277	.00	.99
	Comparison	121	.73	.01			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX J (continued)

Table 74: Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers

Students in Grades 1 and 2 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	57	.87	.01	1, 87	.61	.44
	Comparison	33	.89	.02			
Word Reading	Treatment	57	.93	.01	1, 82	1.43	.24
	Comparison	28	.90	.02			
Sentence comprehension	Treatment	39	.72	.04	1, 44	.06	.81
	Comparison	8	.70	.09			
Passage comprehension	Treatment	27	.63	.03	1, 34	2.33	.14
	Comparison	10	.52	.06			
Total	Treatment	62	.79	.02	1, 96	1.07	.30
	Comparison	37	.76	.02			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX J (continued)

Table 75: Comparison of Two-Year Treatment Students with Experienced Teachers vs. Two-Year Comparison Students with Experienced Teachers

Students in Grades 4 and 5 in Year 2 (Using the Year 1 pretests as a covariate)

Assessment	School	Number of Students	Year 2 posttest Adjusted proportion correct		ANCOVA		
			M	SE	df	F	p
Listening comprehension	Treatment	62	.79	.01	1, 107	5.77	.02*
	Comparison	48	.83	.02			
Vocabulary	Treatment	63	.54	.02	1, 103	.58	.45
	Comparison	43	.52	.02			
Sentence comprehension	Treatment	63	.73	.02	1, 104	.55	.46
	Comparison	44	.75	.03			
Passage comprehension	Treatment	61	.62	.02	1, 106	.68	.41
	Comparison	48	.65	.02			
Total	Treatment	69	.66	.01	1, 117	.88	.35
	Comparison	51	.68	.02			

* Difference is significant at the 0.05 level

† Difference is significant at the 0.10 level

APPENDIX K

Issues with informed consent procedures.

Before beginning the study in Year 1, we received approval from an independent IRB for our research plan. We asked for approval to administer the standardized and researcher-developed pre and posttests, to observe the reading comprehension lessons of teachers who provided consent, and to administer a reading motivation questionnaire. It was our understanding that all students would complete the assessments, as these assessments were part of the curriculum in the treatment school. We would then score the tests of the students who had parental/guardian consent and destroy the assessments completed by students who did not have consent.

After administering the researcher-developed tests in Year One, a teacher at the treatment school contacted the IRB to notify them that she was uncomfortable administering the test to students who had not provided consent. This particular teacher had been opposed to the research project from the beginning of the school year, and did not want her students to be tested. We were contacted by the IRB and discussed our options for continuing the study in a manner that would protect the participants while not compromising the research design.

We applied for, and were granted, a waiver of consent. In this waiver, we explained that assessment is part of the treatment school curriculum, therefore, students do not need to receive consent to be tested. In order to avoid issues with differential consent between the treatment and control groups, informed consent was also waived for the comparison school. We updated our confidentiality plan to ensure that the identities of both the teachers and the students participating in the study would not be revealed. The teachers who had agreed to allow us to observe their lessons provided written consent.