A META-ANALYSIS OF COMPREHENSION STRATEGY INSTRUCTION FOR UPPER ELEMENTARY AND MIDDLE SCHOOL STUDENTS

By

Dennis S. Davis

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Teaching, Learning, and Diversity

August, 2010

Nashville, Tennessee

Approved:

Professor Carin Neitzel

Professor Robert Jiménez

Professor David Dickinson

Professor Bridget Dalton

Professor Mark Lipsey

To little Mirabel Lucy, who I saw for the first time

in an ultrasound picture the same day

I finished this study.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the generous support of Peabody College's Experimental Education Research Training (ExpERT) program and the National Academy of Education's Adolescent Literacy Predoctoral Fellowship. I am also indebted to the members of my dissertation committee for their thoughtful guidance and encouragement, and in particular, to Dr. Neitzel, for her five years of mentorship.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Comprehension strategies are mental tools that readers use to aid their understanding of a text before, during, and after reading, such as previewing, stopping to summarize, and generating questions (Duke & Pearson, 2002; National Institute of Child Health and Human Development [NICHD], 2000). Instructional frameworks designed to teach students to use these tools have been studied extensively over the past thirty years. The genesis of these frameworks is often attributed to the cognitive revolution of the 1960s and 1970s when researchers began using introspective methods to identify the cognitive moves, or strategies, used by proficient readers (Kucan & Beck, 1997; Pearson, 2009; Pressley & Afflerbach, 1995). Around the same time, the reading research community was becoming increasingly concerned with the status of reading comprehension instruction in U.S. schools. In a now classic study, Durkin (1978-1979) found that teachers spent very little time explicitly teaching students to comprehend; instead they devoted large chunks of time to assessing comprehension through interrogation and completion of worksheet-type assignments.

This alarming finding led researchers to search for ways to improve the teaching of reading comprehension. By the 1980s, researchers were beginning to understand the benefits of explicitly teaching the strategies associated with proficient reading (Duffy et al., 1987). Many in the field agreed, for example, that through teacher modeling of these strategies, novice readers would be able to see "expert" comprehension processes in

action and could eventually come to mimic these processes (Collins & Smith, 1982). This instructional approach, which came to be known as comprehension strategy instruction, is based on a view of reading as a problem-solving activity and of readers as active agents in pursuit of comprehension goals (Duffy, 1993).

In early studies of comprehension strategy instruction, researchers examined the role of individual strategies in improving comprehension achievement (e.g., Baumann, 1984; Pressley, 1976). In more recent studies, the field has moved beyond thinking of strategic reading as the use of isolated strategies (NICHD, 2000; Raphael, George, Weber, & Nies, 2009). Current definitions of comprehension regard it as a complex act requiring the careful coordination of multiple strategic moves (RAND, 2002). When reading this paragraph, for example, a skilled reader will make predictions about upcoming content, link what is being read here to information read in other paragraphs and other articles, actively search for the main take-away points, all the while evaluating whether or not the paragraph content is adequately understood. Instructional programs that teach students to use repertoires of strategies—often referred to as multiple comprehension strategies instruction (MCSI)—have supplanted single strategy instruction as the gold standard in strategic reading pedagogy.

Overview of the Problem

Researchers who study adolescent literacy consistently name strategy instruction as a high-leverage instructional practice for students in the middle grades, particularly those who struggle with comprehension (Biancarosa & Snow, 2004). Numerous reviews of the comprehension literature have reached this conclusion (e.g., Block & Duffy, 2008;

Dole, Nokes, & Drits, 2009; Pressley, Johnson, Symons, McGoldrick, Kurita, 1989; Rosenshine & Meister, 1994), including the widely cited National Reading Panel Report (NICHD, 2000). As a result of overwhelmingly positive research support, comprehension strategies have found their way into curriculum materials and educational standards nationwide. For example, the language arts standards for fifth-grade students in Tennessee contain two strands related to comprehension strategies (Tennessee Department of Education, n.d.). Teachers are expected to help students learn a repertoire of pre-reading strategies, including making predictions about text content, organizing background knowledge using graphic organizers, and establishing a reading purpose (reading accomplishment 5.1.07). Also, they are expected to teach a series of strategies for students to use during and after reading, such as visualizing text content, reacting and relating to text, asking questions to clarify understanding, identifying main ideas, and monitoring their own strategy use (accomplishment 5.1.08ab). Similar objectives can be found in the curriculum standards in every state, in the standards documents produced by national organizations (e.g., NCTE/IRA, 1996), and in popular practitioner texts (e.g., Harvey and Goudvis, 2000; Keene & Zimmerman, 1997; Wilhelm, 2001).

The prevalence of strategies in research and curriculum materials does not mean that multiple comprehension strategy instruction is pervasive in classrooms. Researchers have argued quite convincingly that strategy instruction has not made its way into large-scale practice (Durkin, 1978-79; Pressley, Wharton-McDonald, Mistretta-Hampton, & Echevarria, 1998). As Pressley (2008) remarked in his final address to the International Reading Association, "…there is no evidence of much comprehension strategies instruction occurring extensively now and certainly no evidence of children being taught

such strategies to the point that they use them in a self-regulated fashion…" (p. 406). This statement aptly summarizes the prevailing belief among researchers that although teachers and students often use the terminology of strategy instruction (i.e., they talk about "making connections" and "activating background knowledge"), the instructional methods described in the literature are rarely seen in typical classroom practice.

A major factor contributing to the difficulty in translating MCSI research into practice is that researchers who study strategic reading pedagogy operate with varying understandings for how strategies should be selected, taught, practiced, and assessed. This variability, which is described in more detail in Chapter 2, has made it difficult to identify the critical elements of effective strategy instruction that should make their way into practice. When teachers and curriculum developers look to this literature for guidance, they are often left with more questions than answers: Which strategies should I teach? How much instruction is actually needed? Do all students benefit equally from strategy lessons? How should I group my students for these lessons? What sort of independent practice should I provide? Although comprehension strategy instruction is a heavily studied area, researchers simply have not provided enough guidance to help teachers resolve these and other fundamental questions.

The low performance of upper elementary and middle grades students in reading comprehension on national assessments indicates that students in these grades are not receiving effective comprehension instruction. In 2007, 33% of fourth-grade students scored at the below basic level in reading comprehension on the National Assessment of Educational Progress (Lee, Grigg, & Donahue, 2007). According to the performance level descriptions provided in the NAEP report, students who score at this level are

unable to "demonstrate an understanding of the overall meaning of what they read" (p. 20). The percentage of fourth graders who scored below the basic level was even higher for African American students (54%), Hispanic students (50%), students who qualified for free or reduced lunch (50%), and English language learners (70%) (National Center for Education Statistics, 2007). Among eighth-grade students, the data are similar. In the same year, 26% of eighth-graders scored below basic levels on the NAEP reading comprehension test (Lee, Grigg, & Donahue, 2007); this includes 45% of African-American students, 42% of Hispanic students, 42% of those who qualified for free or reduced lunch, and 70% of English language learners (National Center for Education Statistics, 2007). These data suggest that many students are entering fifth grade with inadequate reading comprehension ability, and for many of these students, their instructional needs are not met by the end of eighth grade. These findings underscore the need for improvements in reading comprehension instruction for students in the middle grades, especially for minority students, students from low-income families, and students learning English as a second language.

Comprehension strategy instruction alone is not sufficient for improving middle grades students' reading comprehension. Other reforms are also needed, including improved vocabulary, word recognition, and fluency instruction, as well as methods for building conceptual knowledge and increasing motivation for reading. In fact, strategy instruction has been criticized recently for failing to adequately focus students' attention on constructing coherent representations of text content (McKeown, Beck, & Blake, 2009). Nonetheless, a primary assumption of the present study is that strategy instruction is particularly suited for helping students learn to handle the increased instructional and

textual demands they experience as they transition from lower elementary to upper elementary and middle school.

Summary of the Problem

Despite a strong research base, comprehension strategy instruction has not become part of typical instructional practice in middle grades classrooms. Assessment data from students in these grades shows that the instruction they are receiving is not meeting their needs. Many in the field believe that strategy instruction is one of the elements that would improve middle grades literacy instruction. However, researchers have not provided sufficient guidance to teachers regarding the specific practices and procedures that constitute effective strategy instruction.

Overview of the Present Study

The successful translation of strategy research into pedagogical practice requires a clear understanding of what is known from the numerous studies that have been conducted in the past. To contribute to this understanding, this meta-analytic synthesis examines the relationship between multiple comprehension strategy instruction (the independent variable) and student outcomes (the dependent variables), as depicted with a solid arrow in Figure 1. The analysis also examines the moderating effects of several content and pedagogical attributes of MCSI (the moderator variables), depicted with a dashed arrow (see Chapter 3 for more details).

By identifying questions and conceptual issues that have not been adequately addressed in the past, this review provides strategic jumping-off points for researchers

interested in conducting future studies of text comprehension instruction. Also, this review provides more detailed understandings of the elements of strategic reading pedagogy associated with maximum impact and provides specific recommendations for improving middle school reading instruction.

The Benefits of a Meta-Analytic Review

Meta-analysis is a research synthesis tool used to summarize the relationship between two variables across a body of primary studies and to answer questions not addressed in the individual studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; Glass, 1976). Typically, the pooled data is used to calculate the average magnitude of the relationship between the two variables in question and to examine the extent to which conceptually relevant moderator variables strengthen or weaken this relationship (Lipsey & Wilson, 2001).

Meta-analytic techniques offer important advantages over traditional narrative (i.e., non-statistical) review techniques. In experimental and quasi-experimental literatures, many studies may produce gains in the expected direction as evidenced by positive and statistically significant differences between students in the treatment condition and those in the counterfactual condition. If this result is repeated often enough, a narrative reviewer may be inclined to conclude that the treatment in question produces positive outcomes. The tendency for narrative reviews to come out in favor of interventions with repeated positive evidence is often referred to as "vote counting" (Cook et al., 1992). The vote-counting method is an inadequate synthesis technique

because it places equal weight on the results from all studies and focuses on statistical significance rather than practical significance of the treatment effect estimates.

Meta-analytic reviews, on the other hand, synthesize information not only about the direction of treatment effects, but also their magnitude (Lipsey & Wilson, 2001). This is done by converting the results of individual studies into effect sizes that index the size of treatment impacts in contextually interpretable units (usually, standard deviation units). Furthermore, meta-analytic reviews use statistical procedures to ensure that larger studies, which produce more stable treatment effect estimates, carry more weight in the synthesis than smaller, less precise studies (Hedges & Olkin, 1985).

Another essential characteristic of meta-analytic reviews is methodological transparency, which is not always present in traditional narrative research syntheses (Borenstein, Hedges, Higgins, & Rothstein, 2009). Meta-analyses are governed by many of the same expectations for methodological quality as primary research studies, including the need for a transparent and systematic sampling scheme for collecting data (i.e., published studies), the need for reliable coding and data extraction procedures, and the requirement that interpretations based on quantitative data be consistent with the relevant methodological and statistical assumptions of the study design (Hedges & Olkin, 1985).

Perhaps the most important benefit of meta-analysis is its potential for generating explanatory knowledge of the "intervening mechanisms through which a treatment causes its effects" (Cook et al., 1992, p. 14). This benefit is particularly useful when synthesizing empirical literatures—like the MCSI literature—that are comprised of studies that test different content and pedagogical arrangements. While an individual

study can be set up to test a specific theory or question (e.g., does one experimentally manipulated variable induce student learning relative to a control), a meta-analysis can explore a wider range of theoretical variations and identify specific conditions that facilitate an effect. Knowledge of these conditions improves researchers' theories for explaining how their interventions work and can help predict the effects of similar interventions that share these conditions. In educational science, the explanatory knowledge gained from meta-analyses can produce immediate recommendations for instructional practice (Cook et al., 1992). Experimental studies in which these recommendations are explicitly manipulated will be needed to verify their hypothesized effectiveness, but in the interim, knowledge of these conditions can guide immediate changes to instructional practice.

Contributions and Limitations of Recent Narrative Reviews

Narrative reviews of empirical research are common in the reading comprehension literature. Since it would be impossible to summarize all of these, I have selected four exemplary reviews that provide insights related to comprehension strategy instruction.

The most comprehensive (and controversial) review of reading instruction in recent years is the report of the National Reading Panel, which was commissioned by U.S. Congress to synthesize and make recommendations based on the accumulated research related to reading instruction. The Panel's report (NICHD, 2000) summarizes the published experimental and quasi-experimental studies of reading and provides meta-analytic summaries in the form of numerical effect sizes for instruction related to

phonemic awareness (0.86 across all students and instruction types), phonics (0.44), and fluency (0.41 for guided repeated reading). Unfortunately, the Panel chose not to formally meta-analyze the studies they located on reading comprehension instruction because of the variability in research approaches and the small number of studies that met their methodological criteria.

Instead, the Panel identified six approaches to text comprehension instruction with strong positive empirical support. These include comprehension monitoring, cooperative learning, graphic and semantic organizers, answering questions, generating questions, and summarizing. Additionally, they recommended a seventh approach, multiple strategy instruction, which is the focus of this study. The NRP report provides a useful catalogue of the studies that fall into each of these categories, along with descriptions of the grade levels in which each intervention has been tested, the outcome constructs that have been measured, and a summary of study results using the vote-counting method described above. These extensive catalogues provide the starting point for the present study, which seeks to update and extend the Panel's summary of MCSI.

Another recent publication that typifies the contributions and limitations of the many reviews of strategy instruction is Block and Parris' (2008) edited book *Comprehension Instruction: Research-Based Best Practices*. In the chapter that focuses most heavily on strategy instruction (Block & Duffy, 2008), the authors provide a concise summary of what has been learned about comprehension instruction since the publication of the National Reading Panel report. They present a compelling argument for a direct explanation approach to strategy instruction that prioritizes the use of think-alouds to model the mental processes associated with skilled comprehension. The most prominent

argument in this chapter is that learning to be strategic is difficult, and it cannot be achieved through proceduralized strategy practice divorced of authentic purposes for reading.

The authors criticize commercial curricula for providing limited opportunities for students to participate in guided and independent strategy practice. They argue that these materials are oversaturated with more strategic content than can be adequately covered in a single school year. As an alternative, they present a list of nine consolidated strategies they argue are most effective. These include: predicting, monitoring, questioning, creating mental images, using repair strategies like re-reading, inferring, summarizing, evaluating, and synthesizing.

This review, like many others, provides clear and sensible theoretical insights to guide research and practice, but the methods by which these insights are gained are not made transparent. In particular, the authors do not present numerical information to support their claims of instructional effectiveness.

Another recently released edited volume, *Handbook of Research on Reading Comprehension* (Israel & Duffy, 2009), contains two chapters that directly focus on strategy instruction. In the first of these chapters (Dole, Nokes, & Drits, 2009), the authors critique the current state of strategy instruction, asserting that "In the transition from research to practice, strategy instruction has morphed into so many things that it no longer has a shared meaning" (p. 347). They describe a series of landmark studies of single strategy instruction from the 1970s and 1980s, including studies of mental imagery (Pressley, 1976), summarization (Brown & Day, 1983), text structure (Taylor & Beach, 1984), and story mapping (Idol & Croll, 1987). They also identify a set of seminal

multiple strategy interventions, including Collaborative Strategic Reading, Transactional Strategies Instruction, Peer-Assisted Learning Strategies, and Concept-Oriented Reading Instruction, all of which are described in more detail in Chapter 2.

The major contribution of this chapter is the categorization system the authors use to describe the various instructional delivery methods used in comprehension strategy instruction. The first of these, which the authors refer to as explicit instruction, emphasizes the importance of providing descriptions of what strategies are and how they work before providing extensive strategy practice. The second category of delivery systems, which is only subtly different from the first, is exemplified by Duffy's work in which teachers are taught to model the cognitive processes of comprehension through think-alouds instead of (or in addition to) didactic explanations (Duffy & Roehler, 1987). The authors describe a third category they call cognitive apprenticeship, exemplified by the original studies of Reciprocal Teaching (Palincsar & Brown, 1984). In this model, students learn strategies through scaffolded interactions with teachers without receiving explicit explanations or descriptions of strategies before these interactions occur. Finally, the authors characterize much of what passes for traditional comprehension instruction as implicit, or invisible, strategy instruction. In this delivery system, students primarily use strategies in response to teacher queries or prompts, but there is never the expectation that students will learn to strategically and consciously control their mental processes of comprehension.

In the other chapter of the *Handbook of Research on Reading Comprehension* that provides an overview of strategy instruction (Raphael, George, Weber, & Nies, 2009), the authors provide much of the same information found in the Dole, Nokes, and Drits

chapter described above. The authors of this chapter make two distinguishing contributions. First, they provide a more detailed historical overview of comprehension instruction, which they divide into three waves of research: the individual strategy research of the 1980s, the multiple frameworks of the 1990s, and the efforts to create cohesive school-wide reading programs in the past decade. In their description of multiple strategy instruction, they highlight many of the same instructional frameworks as the authors of the previous chapter (e.g., Reciprocal Teaching and Concept-Oriented Reading Instruction). Second, the authors acknowledge the importance of applying what is known about reading strategies in print media to digital media, and in particular, to online text.

Each of these narrative reviews provides important conceptual insights regarding historic and contemporary approaches to teaching students to use strategies for comprehension. However, as is common with reviews of this type, the procedures used to select and evaluate studies are not made transparent in these reviews, with the exception of the National Reading Panel Report. In most cases, the reviews describe only a small sample of instructional frameworks, drawing largely on work the authors themselves have been involved in. Also, these authors tended to adopt an implicit vote-counting rule, presenting an intervention as "research-based" as long as multiple studies report positive outcomes.

Meta-Analytic Reviews of Comprehension Instruction

One of the earliest meta-analyses to examine comprehension outcomes was conducted by Haller, Child, and Walberg (1988). The authors collected studies published

between 1975 and 1986 in which students were taught metacognitive skills to improve their text comprehension. As described by the authors, these skills included comprehension monitoring, self-questioning, paraphrasing, summarizing, utilizing prior knowledge, and repair strategies such as rereading (the word "strategy" to describe these mental operations was just coming into common use at that time). The unweighted mean treatment-control effect size for the 20 studies they located was 0.71. The authors found no systematic differences in treatment effects on researcher-designed or standardized measures of comprehension. They also found that the largest impacts were obtained with students in seventh and eighth grade and the lowest in grades 4-6. Unfortunately, the report does not provide the numerical effect size estimates for these subgroup analyses.

One of the more frequently cited meta-analyses of comprehension instruction (Mastropieri, Scruggs, Bakken, & Whedon, 1996) included 68 experimental or quasi-experimental studies published between 1976 and 1996 that reported reading comprehension outcomes for learning disabled (LD) students. The mean effect of reading comprehension instruction was 1.03 on criterion-referenced measures and 0.40 on norm-referenced measures. The authors also compared treatment effects across three broad instructional types. Interventions in which the text was modified or enhanced (i.e., adding facilitative text features such as illustrations, highlighting, or underlining) produced a mean effect of 0.92. Skill enhancement programs (e.g., those in which students received vocabulary or fluency instruction) produced a mean effect of 0.62. The largest impact was observed for studies in the self-questioning category, with a mean effect of 1.33. These studies tested interventions that in the current research climate would likely be

called strategy instruction, including studies of summarization and self-monitoring training.

In a similar meta-analysis, Swanson (1999) identified experimental studies published between 1963 and 1997 that reported word recognition and reading comprehension outcomes for LD students. From the 58 studies reporting comprehension outcomes, the average effect size was 0.45 for standardized measures and 0.81 for researcher-designed measures. The authors also report a mean ES of 0.98 for studies that included strategy instruction and 1.15 for studies that coupled strategy instruction with direct instruction. The authors identified instructional features associated with positive comprehension gains, including teacher modeling and the use of strategy cues.

More recently, Sencibaugh (2007) analyzed the impact of comprehension strategy instruction on the reading comprehension of students with learning disabilities in studies conducted between 1985 and 2005. The overall effect of interventions using what he calls "auditory-language dependent" strategies (e.g., summarizing, self-questioning, as well as frameworks such as Reciprocal Teaching and Collaborative Strategic Reading) was 1.18, compared to 0.94 for those using "visually dependent" strategies (e.g., illustrations and semantic organizers). However, the effect sizes reported in this study are difficult to interpret because it is not clear how (or if) the author separated pre-post and treatment-control effect sizes, which are not directly comparable.

Finally, the meta-analysis that is most akin to the present study is the influential Rosenshine and Meister (1994) review of research on Reciprocal Teaching (RT). They summarized the results of 16 studies of RT conducted between 1984 and 1992, including the original RT studies of Palincsar and Brown (Palincsar, 1987; Palincsar & Brown,

1984). The median effect size of the intervention was 0.32 on standardized measures of comprehension and 0.88 on researcher-designed measures. The authors found no systematic relationship between ES and instructional duration, size of instructional groups, or the personnel responsible for delivering the instruction. Also, the effects were found to be similar for studies with heterogeneous groups of students (0.32) and those with groups made up primarily of struggling comprehenders (0.29).

Taken together, these meta-analytic reviews provide statistical summaries of the overall effectiveness of selected branches of the strategy instruction literature for selected populations. Despite this important contribution, the meta-analytic reviews tend to fall short in comparison to the narrative syntheses in the way they conceptualize strategies instruction. In particular, the Mastropieri et al., Swanson et al., and Sencibaugh meta-analyses employ broad and ill-defined conceptual categories that are not commonly used in the comprehension literature. Furthermore, the meta-analytic studies have tended to either focus on LD students only or on a narrow range of frameworks for teaching comprehension strategies (e.g., Reciprocal Teaching in the Rosenshine and Meister review and "self-questioning" instruction in the Mastropieri et al. review). What is needed is a more comprehensive synthesis of the strategy literature that combines the theoretical richness of the narrative reviews with the methodological transparency of the statistical reviews.

How this Meta-Analysis Contributes to the Field

The present study will contribute to the knowledge base of MCSI in a number of ways, as described in the following sections.

Updating and Expanding the Findings of Previous Reviews

This study was designed to be more comprehensive in scope than previous meta-analytic reviews and more methodologically transparent than previous narrative reviews. This synthesis includes recent developments in strategic reading pedagogy that have gained popularity in the past decade, such as programs designed for English language learners (e.g., Klingner & Vaughn, 2000), programs providing strategy support in digital literacy environments (e.g., McNamara, O'Reilly, Best, & Ozuru, 2006), and instructional approaches that have not traditionally fallen under the heading of strategy instruction but clearly include instruction in multiple strategies (e.g., Think Aloud Instruction [Bereiter & Bird, 1985; Baumann, Seifert-Kessell, & Jones, 1992]). Furthermore, unlike previous studies, this analysis focuses on a specific age range (grades 4 – 8) where strategy instruction is thought to be particularly relevant. Within this age range, the review includes a wider range of student ability levels so that the relative effectiveness of MCSI for LD students, average and struggling readers, and English language learners can be directly compared. Finally, the present study includes outcome constructs that, in addition to reading comprehension, are frequently measured in the strategy instruction literature but are often left out of research syntheses. These include strategy knowledge and strategy use, which are described in more detail in Chapter 2.

Examining Explanatory Variables with an Emphasis on Practical Recommendations

As explained in a previous section, one of the major benefits of meta-analysis is the possibility of exploring conditions and factors that strengthen or weaken observed

treatment effects. Identifying these factors will help generate practical recommendations for the types of instructional decisions teachers should be making when planning and implementing MCSI in their classrooms. This parsing of practice into useable, learnable chunks is consistent with recent theoretical work in teacher education that holds that so-called "decompositions of practice" are important for helping teachers recognize, talk about, and develop expert ways of implementing complex pedagogies (Grossman et al., 2009; Singer-Gabella & Tiedemann, 2008).

This analysis focuses on three broad categories of explanatory conditions, described below. Specific details for how these characteristics are operationalized and coded are presented in Chapters 2 and 3.

Student characteristics. This meta-analysis examines the benefits of strategy instruction for different types of students by comparing the impact of instruction on students of different grade levels, linguistic heritages, and reading achievement levels.

Content characteristics. As explained in Chapter 2, various instructional frameworks have been developed for teaching MCSI to middle grades students. In the present analysis, the specific instructional content of these frameworks (i.e., what students are taught) is examined. For example, although most researchers in the field agree that students should learn to coordinate a flexible strategic repertoire instead of isolated strategies (Baker & Brown, 1984; Brown, Pressley, Van Meter, & Schuder, 1996), it is unclear which strategies should be included in this repertoire or if any repertoire will do (McKeown, Beck, & Blake, 2009). In this study, the different combinations of strategies that have been used in instructional studies are catalogued and compared to identify the most effective combinations.

Pedagogical characteristics. Frameworks for teaching strategic reading also vary in the way content is presented to students. For instance, some programs include dozens of lessons spaced over an entire year (e.g., Van Keer & Verhaeghe, 2005; Guthrie, Van Meter, McCann, Wigfield, et al., 1996) while others include only a few lessons (e.g., Johnson-Glenberg, 2005). This study examines the relationship between educational impact and the amount of instruction. Strategic reading frameworks also differ in the instructional contexts that are used to promote strategy learning and practice. Some studies place heavy emphasis on peer collaboration (e.g., Palincsar & Brown, 1984; Klingner & Vaughn, 2000) while others emphasize extended time in teacher-led activities (e.g., Englert & Mariage, 1991) or individually completed computer-prompted activities (e.g., Proctor, Dalton, & Grisham, 2007; Sung, Chang, & Huang, 2008). Some emphasize independent reading habits (e.g., Kim, 2006; Block, 1993), but most do not. Some frameworks are primarily organized around narrative text (e.g., Loranger, 1997) while others emphasize expository and content-area literacy (e.g., Klingner, Vaughn, Schumm, 1999; Radcliffe, Caverly, Peterson, & Emmons, 2004). A goal of the present study is to classify and evaluate the various instructional contexts employed in this diverse literature.

Creating a Blueprint for Future Research

The substantive information that is compiled during a meta-analytic study essentially falls into two categories: (1) descriptive information that indexes the distinguishing characteristics of the interventions being studied and (2) statistical information related to intervention impacts. When compiled, the information in the first category provides a global view of which types of intervention characteristics have come

to dominate the research base and which ones have not been adequately studied. In the present study, this descriptive overview will help identify areas that researchers should attend to in their future work and those areas where additional studies would most likely produce redundant and unfruitful findings.

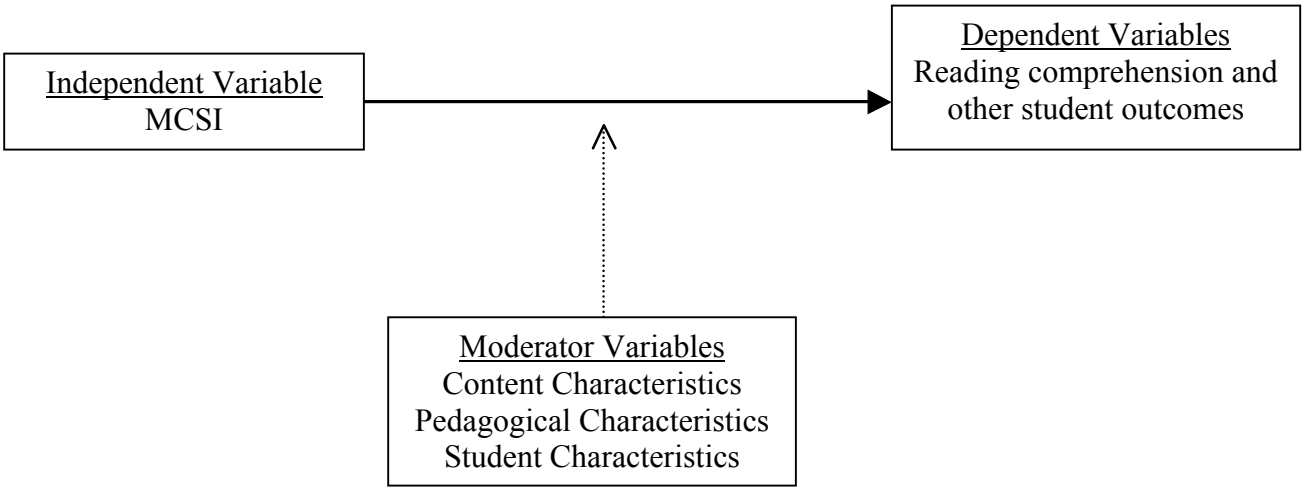

Figure 1: Diagram of Independent, Dependent, and Moderator Variables

CHAPTER II

REVIEW OF THE LITERATURE

A crucial initial step in the completion of a meta-analytic review is for the researcher to spend time developing a conceptual understanding of the relevant literature. In this chapter, I synthesize the published literature on multiple comprehension strategy instruction (MCSI) with the aim of developing a conceptual scheme to guide the work presented in later chapters. In this first section of this chapter, I clarify the definitions of key terms used in studies of MCSI in order to establish a precise vocabulary that will be used in the remainder of this study. In the remaining sections, I present a descriptive review of the intervention studies in the MCSI literature, which informs the subsequent analyses.

Clarifying Key Terms and Definitions

A key point of confusion in the MCSI research literature is the use of important terms related to strategy instruction. In this section I will briefly discuss some of these terms and describe how they will be used in this study, relying on the work of other authors who have attempted to clarify the use of these terms (e.g., Afflerbach, Pearson, & Paris, 2008; Alexander, Graham, & Harris, 1998).

Distinguishing Strategies and Skills

A reading skill is an ability or competence that a reader carries out automatically. A strategy, in contrast, is a mental tool a reader uses on purpose to monitor, repair, or bolster comprehension (Afflerbach & Cho, 2009). The key difference between strategies and skills is that strategy use is volitional and goal-oriented—a reader must choose to use a strategy to address a particular reading goal, while a skill will typically be carried out without conscious awareness (Afflerbach, Pearson, & Paris, 2008). In much the same way that a person who has learned to drive a car does not have to think about the various skills involved during his or her morning commute—such as maintaining the vehicle on the right side of the road, activating the turn signals at the right times, and using the brake to slow down suddenly when a dog runs out in the street—a skilled reader does not have to consciously coordinate the many skills involved in reading. These skills might include such cognitive behaviors as making inferences to connect information across different parts of the text and recognizing the referents for pronouns. Strategies, on the other hand, are called on when something suddenly gets in the way of automatic processing—for instance, if smoke begins to flow from under the hood of the car. When this happens, the driver will have to figure out the best way to pull over to the shoulder without endangering other drivers, who to call for repairs, and how to get to work on time. Analogously, a reader experiencing comprehension difficulties must take strategic action to repair these difficulties before proceeding through the text.

The difference between strategies and skills is very subtle for highly skilled readers, who automatically employ a lot of mental processes that for less advanced readers would require conscious awareness (Afflerbach & Cho, 2009; Paris, Lipson, &

Wixson, 1983). Consider, for example, the common strategy of attending to and restating main ideas while reading. To practice this cognitive move, a student who struggles with reading comprehension will need to purposefully stop reading every so often and try to identify the most important information and then paraphrase it in his or her own words. An advanced reader, on the other hand, will likely grasp the most important information of the text implicitly, without the need for raising this process to conscious awareness. This example illustrates another important distinction between strategies and skills: that they exist on different poles of a developmental continuum. Reading behaviors that are introduced to students as strategies can eventually become skills as their use becomes automatized (Anderson, Graham, & Harris, 1998).

The distinction between strategies and skills is also dependent on the context in which they are used. A strategy that has been internalized as a skill may not be used automatically in every reading situation. Whether a reader uses a reading behavior as a skill or a strategy depends on the demands of the text for that particular reader (Afflerbach, Pearson, & Paris, 2008; Paris, Lipson, & Wixson, 1983). If the advanced reader from the previous example is asked to read a passage from a college textbook on an unfamiliar topic, the comprehension difficulties he or she experiences will induce a "debugging"—or fix-up—state that he or she will have to resolve through strategy use (Palincsar & Brown, 1984). For example, this fictitious reader may need to strategically pause after each paragraph to paraphrase the most important information in order to form an adequate understanding of the passage.

Self-Regulated Comprehension

Consistent with broader theories of self-regulation (Zimmerman, 1990; Zimmerman & Schunk, 2001), self-regulated comprehension is a process made up of goal setting, comprehension monitoring, strategy use, and strategy monitoring (Horner & Shwery, 2002; Massey, 2009; Paris, Lipson, & Wixson, 1983). Goal setting, in this case, refers to the process of selecting particular comprehension goals. A reader may begin a reading episode with a variety of goals in mind, such as learning new content about a chosen topic or enjoying a story by a favorite author. During the reading episode, the reader monitors his or her progress toward the chosen goals. This process of comprehension monitoring requires a high degree of metacognition, which is a reader's awareness of his/her thinking processes (Collins, Dickson, Simmons, & Kame'enui, 1996; Jacobs & Paris, 1987). The term metacomprehension is sometimes used to describe the specific type of metacognition used during reading (Maki & Berry, 1984; Wong & Jones, 1982). When the reader recognizes that a goal is not being met, he or she must select and enact strategies to put the reading process back on track and continually monitor the usefulness of these strategies (Pressley, 2003). This strategy monitoring process requires what Paris, Lipson, and Wixson (1983) call conditional strategy knowledge, or knowledge of when and why to invoke a particular strategy.

Strategy Types and Purposes

A cognitive strategy can be any type of volitional tool designed to enhance mental activity in any area, such as mathematics, general problem-solving, or reading (Alexander, Graham, & Harris, 1998; Dole, Nokes, & Drits, 2009). Similarly, a

metacognitive strategy can be used to enhance mental processing in any discipline. There is an important difference between metacognitive and cognitive strategies, however. Cognitive strategies are used to improve cognitive performance (Alexander, Graham, & Harris, 1998). For instance, when putting together a puzzle, an individual may decide to separate the puzzle pieces into piles based on background color in order to make the process more efficient. Since metacognition is the process of thinking about one's own thinking (Jacobs & Paris, 1987), metacognitive strategies are used to monitor and evaluate thinking and problem solving (Hooutveen & van de Grift, 2007). A strategy of this type might be used to evaluate the suitability of a cognitive strategy for meeting a particular goal, like pausing to evaluate whether or not the sorting strategy was helpful for finishing the puzzle. In short, cognitive strategies aid the "doing" of a problem-solving process, and metacognitive strategies aid the monitoring of this "doing."

The term *reading strategy* can refer to any number of volitional strategies (metacognitive or cognitive) that aid the process of reading, including strategies for word recognition and comprehension, and those related to monitoring reading habits and practices, like book selection strategies and strategies for tracking what one has read and plans to read. *Comprehension strategies* are a subset of reading strategies that are used in the service of comprehension enhancement (Dole, Nokes, & Drits, 2009). These are the strategies discussed in this review.

Comprehension strategies can also be described as serving three basic purposes, as shown in Figure 2. First, they can serve a metacognitive purpose by helping a reader monitor comprehension—that is, to notice when a comprehension breakdown occurs and to set into motion a strategic plan for resolving the problem (Palincsar & Brown, 1984).

For example, a reader might put herself in a "watcher" state and carefully look for inconsistencies in the meaning that is drawn from the text (Bereiter & Bird, 1985). Once a comprehension breakdown has been identified, another set of cognitive strategies will be deployed to serve a repair—or fix-up—role (Casteel, Isom, & Jordan, 2000; Duffy & Roehler, 1987). Even when a specific comprehension breakdown has not been identified, the reader may seek to bolster his or her understanding of the text by using an enhancement strategy. Summarizing after reading to review what has been learned, making predictions while reading to increase textual interaction, and generating teacher-like questions to prepare for a class discussion can all serve an enhancement role.

Many strategies can serve more than one purpose, depending on how they are used. For instance, a student could use the question generation strategy to monitor comprehension by asking questions throughout the text and checking to make sure they can be answered correctly. Similarly, the main idea identification strategy can be used to monitor comprehension and also to repair comprehension when used to locate and practice the most important information in a troublesome section of the text.

Instructional Strategies Versus Comprehension Strategies

Another common confusion in the strategic reading literature is the distinction between instructional strategies used by teachers to help students improve their text comprehension and comprehension strategies students learn to use on their own to increase their textual understanding (Dole, Nokes, Drits, 2009). The former are instructional methods, including such well-known methods as guided reading (Fountas & Pinnell, 1996), KWL (Ogle, 1986), and the Directed Reading-Thinking Activity

(Haggard, 1988), as well as more general instructional arrangements like cooperative learning (e.g., Slavin, 1980) and reader's workshop (Atwell, 1987, 1998).

The confusion between instructional strategies and learning strategies arises in part because the word *strategy* is used to name both of them. This confusion can be clarified by considering the agent using the strategy. Teachers strategically select instructional methods to meet specific learning goals they have identified for their students, while students strategically select comprehension strategies to meet specific goals they have identified for their reading. A useful example of this distinction is provided by Conley (2008), who describes two hypothetical teachers using graphic organizers with adolescent students. In one class, the teacher guides students through the completion of the graphic organizer as a way of introducing and helping them organize new content, but there is no expectation that students will adopt the graphic organizer as a strategic tool they can use on their own when reading from their textbooks. In the second classroom, the teacher introduces the graphic organizer as a cognitive tool students can use to record and organize information as they read. In the first scenario, graphic organizing is an instructional strategy, and in the second, it is taught as a comprehension strategy.

Procedural, Orientational, and Discursive Views of Strategies

As this review will show, strategic reading is often conceptualized in procedural terms—that is, students are taught a particular routine or sequence of strategies to follow. In contrast, other studies conceive of strategic reading as an orientation or stance. Students must appropriate a conception of reading as a problem-solving activity that can

be approached using a body of strategies (Anderson & Roit, 1993; Duffy, 1993). In this view, strategies are not used in a prescribed sequence (although they may be practiced that way at first). Instead, strategies are selected in response to specific reading goals or problems that arise. This flexible, self-directed process has been described elsewhere as constructively responsive strategy use (Pressley, El-Dinary, & Afflerbach, 1995).

Consider again the example of the main idea identification strategy discussed above. When taught as a procedure, a student may be required to stop at specified intervals (every half page, for example) and restate the main idea, regardless of whether there was a need to clarify or monitor comprehension at that point in the text. When taught as an orientation, the student will be expected to decide—while participating in appropriately scaffolded strategy practice—where and how to use the main idea strategy depending on the goal of the reading activity and the comprehension difficulties he or she encounters. In some strategic reading interventions, researchers may take a procedural view at first, with the expectation that through structured practice students will adopt a strategic orientation.

Strategy learning can also be thought of from a discursive standpoint. An important part of the process of developing a strategic orientation is the adoption of a specific way of talking about comprehension (Englert, Tarrant, Mariage, & Oxer, 1994). The language of comprehension expected during strategy instruction is anything but natural. For instance, when discussing a narrative text during a strategic reading lesson, students are expected to not only comment on the literary elements and content, but also on the process of comprehending the material. This metacognitively intense language can

be thought of as a specific academic discourse, or social language, that identifies the speaker as a particular kind of reader (Gee, 2001).

Summary: A Precise Working Definition of Comprehension Strategies

The definition of comprehension strategies used in the remainder of this review is as follows: A comprehension strategy is a mental tool a reader uses volitionally to monitor, repair, and enhance his or her comprehension of a text. Strategies for monitoring are metacognitive, in that they are used to evaluate and regulate one's choice of cognitive strategies. Repair and enhancement strategies are cognitive because they improve the coherence of one's mental representation of a text. Comprehension strategies are different from instructional strategies because they are (eventually) initiated by a student in the service of particular comprehension or learning goals. Comprehension strategies may be learned and practiced in the context of instructional procedures, but learning a procedure is not the same as learning to use a strategy. Ultimately, effective strategy use requires the adoption of a conception of reading as a problem-solving process that is enhanced by strategic activity.

Descriptive Review Method

Three primary questions guided this portion of the review: (1) What are the various instructional frameworks that have been studied for teaching multiple comprehension strategies to middle-grades students?, (2) What are students taught in instructional studies employing these frameworks?, and (3) How is this content taught? Because this initial review was designed to provide a conceptual backdrop for the meta-

analytic review described in Chapter 3, these questions were designed to allow me to identify (1) the instructional attributes that are worth analyzing as moderators of instructional impact, (2) the major outcomes tracked in this literature, and (3) the student populations that are most frequently studied in this literature.

The three-step procedure used to conduct this systematic descriptive review is modeled after procedures used in meta-analytic syntheses (Lipsey & Wilson, 2001). First, I delineated a set of inclusion criteria that would guide the selection of studies representative of the knowledge base of multiple strategies instruction. Then, I used a series of search strategies to locate the studies that met all these criteria. Once the corpus of studies was located, I used a methodical notetaking protocol to extract the information from these studies needed to address the guiding questions listed above.

Study Inclusion Criteria

To be considered for inclusion, a study had to be published in a peer-reviewed journal between 1980 and May 2008. The decision to only include studies published in journals was made based on the belief that much of the knowledge base in a given discipline is conveyed to members of the discipline through journal articles. Furthermore, this field is so vast that it would have been unwieldy to include every conference paper, technical report, and dissertation.

Second, the review was limited to instructional (or intervention) studies reporting original research. This meant there had to be an explicit intent to teach comprehension strategies and track some type of student outcome. This included any study that tracked learning gains using any qualitative or quantitative research method, but excluded

practitioner articles that describe or recommend instructional methods in a "how to" format, reviews that summarize research reported elsewhere, and non-instructional studies in which groups of students were surveyed or tested at one point in time for their knowledge of strategies and comprehension abilities. Also, this criterion resulted in the elimination of studies that only reported teacher outcomes.

Third, this review focuses exclusively on studies in which students are taught two or more comprehension strategies. This decision reflects the movement within the field to teach students repertoires of strategies rather than isolated strategies (Brown, Pressley, Van Meter, & Schuder, 1996; NICHD, 2000). Additionally, this review focuses on strategy instruction as it occurs in upper elementary and middle school settings (grades 4-8), since students in these transitional grades stand to gain a lot from instructional frameworks that help them approach texts in sophisticated, self-directed ways (Biancarosa, 2005; Biancarosa & Snow, 2004).

Finally, studies selected for inclusion had to explicitly foreground strategy instruction in their instructional approaches. This resulted in the elimination of studies that simply mention strategy instruction as one of many instructional components of a larger program without placing special emphasis on tracking the results of strategy instruction.

Study Collection Procedures

Studies were located for this review using four search strategies. First, I collected all the studies identified in the National Reading Panel Report (NICHD, 2000) as multiple strategy interventions and the studies included in Rosenshine and Meister's

(1994) meta-analysis of Reciprocal Teaching that meet the inclusion criteria listed above. Then I conducted electronic searches of ERIC, Education Abstracts FTX, PsychInfo, and Social Science Citations Index, using the search terms strateg(y)/(ies)/(ic) AND reading AND comprehension. Third, collected studies (including conceptually useful but ineligible studies and syntheses) were mined for additional citations. Finally, a manual review of the Tables of Contents was conducted for three key journals: *Reading Research Quarterly*, *Journal of Literacy Research*, and *Journal of Educational Psychology*. Eighty-nine studies were found to meet all the inclusion criteria (see complete list in References section).

Data Extraction Procedures

The collected studies were read and summarized using a systematic notetaking protocol that was developed to address the specific questions guiding this review. For each study, I recorded information provided by the authors about the student participants, the number and duration of instructional sessions, and the reported outcomes. Additionally, I recorded information regarding the specific strategies taught, the types of texts used in the intervention, and the amount of emphasis placed on various instructional modes and arrangements.

A second reader who is familiar with research on comprehension and research methodology was taught to use the notetaking protocol, and she and I reviewed two studies together and discussed the information that was extracted. This process allowed me to refine the notetaking procedure and to identify areas of the protocol that required

extra caution. The data collected through the notetaking procedure were entered into a database that was used to identify trends across the body of empirical work.

Overview of Studies in the Descriptive Review

A total of 90 research reports are included in this review. These are listed in Table 1, organized by instructional framework. In the sections that follow, I provide a general overview of the characteristics of these studies.

Study Characteristics

Although the search criteria allowed for inclusion of studies published as early as 1980, the earliest located study was published in 1984. These studies were published in 44 different journals, including some of the respected journals in the fields of educational psychology and literacy research. Journals that yielded three or more articles include: *The Elementary School Journal* (7 articles), *Learning and Instruction* (5), *Reading Research Quarterly* (4), *Journal of Learning Disabilities* (4), *Cognition and Instruction* (3), *Exceptional Children* (3), *Journal of Adolescent and Adult Literacy* (3), *Journal of Educational Psychology* (3), *Reading Research and Instruction* (3), and *Remedial and Special Education* (3). The studies were conducted in 11 different countries: the United States, Canada, New Zealand, Netherlands, Finland, Spain, Belgium, Australia, Germany, Hong Kong, and Taiwan.

A variety of study types are included in this sample, including qualitative/interpretive studies, experimental and quasi-experimental group comparisons, repeated measures studies (e.g., pre/post comparisons and studies with multiple measures

across time), and mixed-method studies. Study size ranged from 3 to 2000 students receiving strategy instruction. The mean number of students in instructional conditions was 76 (sd=229), and the median was 24.

Student Characteristics

In total, the authors of these studies provided multiple strategy instruction to an estimated 6,180 students. The most commonly represented grade levels were fourth and fifth grade, which were included in 33 and 34 studies, respectively. Twenty-eight studies included sixth-grade students; 25 included seventh-grade students; and fifteen included eighth graders. Many studies provided instruction to multiple grades, and very few authors reported the exact number of students from each grade who participated. The race/ethnicity of the participating students was reported (or could be estimated) in only 53 of the studies. Of these, White students were the most highly represented ethnic group in 29 studies and Latino students were the most highly represented group in 15 studies. Only six studies included a predominantly African American student sample, and only three had a predominantly Asian sample.

In 15 studies, all or almost all of the participating students were officially diagnosed with learning or reading disabilities. In three studies, most (50-94%) of the participating students had an LD or RD diagnosis, while six studies included some students identified with this label (between 20 and 49%). Of the remaining studies, 25 included few or no students identified as LD or RD, and percentages of LD/RD could not be estimated for 32 studies.

In 44 of the reviewed studies, the authors describe the student sample as comprised mostly of struggling readers. In most of these studies, students were selected who scored below average on tests of reading comprehension but average in decoding ability. Two studies describe interventions designed specifically for students with below-average comprehension and decoding (Aarnoutse, van den Bos, and Brand-Gruwel, 1998; LeFevre, Moore, & Wilkinson, 2003). Thirty-four of the studies include mixed ability samples made up of average and struggling readers. Only three studies (Bereiter & Bird, 1985; Rauenbusch & Bereiter, 1991; Taylor & Frye, 1992) explicitly excluded struggling readers from their student samples.

Of the 47 studies that provide information about students' linguistic heritages, only six were conducted using student samples made up almost entirely of students learning to read in a second language. In five additional studies, more than half of the student participants were learning to read in another language; another three studies included some (20-49%) students learning strategies in their L2. Thirty-three studies included few or no students identified as second-language learners.

Instructional Frameworks of MCSI

Comprehension strategy instruction is not a single method, but rather a conglomeration of various instructional frameworks, each with its own preferred content, methods, and materials. In this section, I describe the major instructional frameworks that have been studied with upper elementary and middle school students.

The Seven Major Brands of MCSI

I refer to the instructional frameworks included in this first category as brands because, like brand-name products on a shelf at the store, their names are widely recognized and associated with specific expectations. The first four of these brands (Reciprocal Teaching, Informed Strategies for Learning, Peer-Assisted Learning Strategies, and Collaborative Strategic Reading) tend to be treated as packaged materials, as each replication includes a similar set of strategies and instructional methods. The last three (Think-Aloud Instruction, Transactional Strategies Instruction, and Concept-Oriented Reading Instruction) are more general approaches that provide researchers and teachers flexibility in what gets taught and how.

Reciprocal Teaching (RT). Reciprocal teaching is by far the most popular and influential type of multiple strategy instruction in the research literature for upper elementary and middle school students. Twenty-three percent (19 articles) of the articles reviewed here include RT instruction. The development of RT is described in a series of articles by Annemarie Palincsar and Anne Brown, who provide a theoretically rich and groundbreaking description of the mental processes involved in skilled comprehension (Brown & Palincsar, 1982; Palincsar & Brown, 1984; Palincsar, Brown, and Martin, 1987), One of the most influential ideas presented in this work is their distinction between automatic and "debugging" states in reading. The former is characterized by smooth and automatic construction of meaning, which is how reading typically occurs when text demands do not exceed a reader's skill level. The reader enters a debugging state when he or she realizes that a segment of text has not been adequately understood. To resolve these comprehension breakdowns so that automatic comprehension can

proceed, the reader must use "active strategies that take time and effort" (Palincsar & Brown, 1984, p. 118).

Another important theoretical contribution of this work is the identification of six mental processes inherent in text comprehension:

> (1) understanding the purposes of reading, both explicit and implicit; (2) activating relevant background knowledge; (3) allocating attention so that concentration can be focused on the major content at the expense of trivia; (4) critical evaluation of content for internal consistency, and compatibility with prior knowledge and common sense; (5) monitoring ongoing activities to see if comprehension is occurring, by engaging in such activities as periodic review and self-interrogation; and (6) drawing and testing inferences of many kinds, including interpretations, predictions, and conclusions. (Palincsar & Brown, 1984, p. 120)

Their reciprocal teaching procedure was designed to help students fine-tune these six processes through repeated use of four concrete strategic activities while reading expository text. These strategies are not presented as necessary mental processes themselves, but as concrete activities that allow students to practice the six mental operations they identified as important (Palincsar & Brown, 1984).

In the initial investigations of RT, seventh grade struggling readers practiced using the four strategic moves while working in small groups (dyads and triads) with an adult. Originally, the adult was a member of the research team, but in later versions, classroom teachers were trained in the use of RT (Palincsar & Brown, 1984). After the instructor modeled and explained the strategies, students took turns playing the teacher role. After the group read a segment of text, he or she would summarize the main points, ask a question, and make a prediction about what might happen in the next segment. If needed, the teacher-student would lead the group in clarifying a comprehension difficulty before moving on to the next segment, which would be led by another student. The

instructor's participation in these discussions was prominent at first and then gradually receded as students learned to gain more control over regulating their use of the strategies. In a later version of RT (Palincsar, Brown, & Martin, 1987), students were trained to serve as same-age tutors for their classmates. The tutoring students learned the strategies through direct instruction and guided practice, and then they taught the strategies to their tutees and practiced them using the procedure outline above.

As described in the 1984 report, the design of RT was influenced heavily by Vygotskian theories of learning and development, in particular, the notion that children learn by participating in scaffolded verbal interactions with adults and knowledgeable peers (Vygotsky, 1978). In all three of the studies conducted by Palincsar and Brown included in this review, the term *reciprocal* describes the practice of alternating teacher and student roles as students become more skilled in leading the strategy dialogue.

RT as a pedagogical procedure for improving text comprehension has been studied numerous times by other researchers. In their influential meta-analytic study, Rosenshine and Meister (1994) provide a summary of the effectiveness of these studies, which were found to have a median treatment effect size of 0.32 for standardized and 0.88 for researcher-designed measures of comprehension. The four published intervention studies included in the Rosenshine and Meister analysis are included in this review as well (Labercane & Battle, 1987; Lysynchuk, Pressley, & Vye, 1990; Palincsar & Brown, 1984 [two different studies]; Taylor & Frye, 1992), along with 15 additional RT studies conducted in upper elementary and middle school settings.

Some of these replication studies use scaffolded and highly interactive instructional procedures compatible with Palincsar and Brown's original vision (e.g.,

Kelly, Moore, & Tuck, 1994; Lederer, 2000). However, many of these studies are more accurately described as "spin-off" studies because the authors significantly modify the RT procedures. For example, Soriano, Vidal-Abarca, and Miranda (1996) excluded the self-questioning strategy from their version of RT, while Klingner and Vaughn (1996) added two strategies (brainstorming and highlighting main ideas). Other authors have attempted to improve upon RT, for example, by adding audiotaped texts to make the intervention suitable for struggling decoders (Le Fevre, Moore, & Wilkinson, 2003), using cross-age instead of same-age tutors (Klingner & Vaughn, 1996), adding independent strategy practice while students read on their own (Lysynchuk, Pressley, & Vye, 1990), and alternating lessons in students' first and second languages (Fung, Wilkinson, & Moore, 2003). One modification that is included in almost every spinoff study of RT is the infusion of introductory strategy lessons before the reciprocal teaching dialogues begin (e.g., Brand-Gruwel, Aarnoutse, & Van den Bos, 1998).

What is often missed by authors who have conducted RT spinoff studies is that Palincsar and Brown did not intend for the four strategies of predicting, clarifying, summarizing, and questioning to become the required list of strategic moves students should learn (Palincsar, 2007). The four strategies were intended to structure the way students and teachers collaborated with each other and with text so that the six mental processes they identified could be practiced within students' zones of proximal development (Palincsar & Brown, 1984). For example, imagine a student who has yet to develop the ability to monitor his/her understanding of a text (metacomprehension). Simply telling this student to think about whether or not she understands a text will not suffice because this is precisely the mental action that has not developed. A more

appropriate tack from a scaffolded learning perspective would be to guide the student through an activity that provides an entry point to the development of self-monitoring ability. This is precisely what the summarization procedure used in the original RT framework was designed to do. A failed summarization attempt allows a student to recognize that a comprehension breakdown has occurred. With the help of peers and teachers, the student can practice debugging understanding, and eventually, the monitoring and clarifying cycle will become an automatic process. Thus, summarization itself is not the goal of the instructional interaction. The goal is for the student to internalize a general procedure of active self-monitoring and problem solving (Rosenshine & Meister, 1994).

Also frequently missed by researchers who use RT in spin-off studies is the dialogic interaction at the core of the original RT framework (Palincscar, 2007). The four strategies provide students the opportunity to verbally practice their covert mental processes. The goal is for students to gain facility with a particular form of textual dialogue. The strategy labels not only index cognitive moves, but linguistic moves as well (Rosenshine & Meister, 1994). By discussing their strategic processes verbally, students can internalize a procedure for noticing and repairing comprehension breakdowns.

Informed Strategies for Learning (ISL). This instructional brand, first described for upper elementary and middle school students by Paris, Cross, and Lipson (1984), was developed about the same time as Reciprocal Teaching but has not been as influential. Two studies of ISL that were conducted with fifth grade students are included in this review.

ISL was designed specifically to help students become more metacognitively aware of their reading processes. One of the primary theoretical contributions of this line of work is that its authors helped distinguish three types of knowledge related to strategies. Students with declarative knowledge are able to list and describe strategies. Procedural knowledge refers to knowing how to use strategies. Conditional knowledge, which is heavily stressed in ISL, is knowledge for when to use a particular strategy (Jacobs & Paris, 1987; Paris, Cross, & Lipson, 1984; Paris, Lipson, & Wixson, 1983).

Unlike RT, which is a set of general instructional procedures, ISL is structured more like a curriculum. It includes a long list of strategies (including understanding purpose for reading, activating background knowledge, attending to main ideas, critical evaluation of text, comprehension monitoring, and making inferences) taught over the course of four months using specific lesson modules designed by the researchers. During these lessons, teachers introduce each strategy separately using decorative bulletin boards depicting metaphors for strategic reading. For example, students are introduced to active reading by discussing what it means to be a "text detective" who constantly looks for clues while reading. Also, traffic sign metaphors are used to illustrate the need for reading carefully and monitoring comprehension. These lessons are delivered using the typical faded support model, which includes direct instruction, guided practice, and independent practice (Paris & Oka, 1986).

Peer-Assisted Learning Strategies (PALS). Four studies investigating the effectiveness of Peer-Assisted Learning Strategies with upper elementary and middle school students were located for this review. The procedures used in this brand of strategy instruction are based on an earlier instructional program called classwide peer

tutoring in which students spent large chunks of time working on academic tasks in same-age dyads (Greenwood & Delquadri, 1995).

The strategic reading procedure used in PALS is similar to Reciprocal Teaching in that both interventions emphasize student collaboration and dialogue. However, the interactions in PALS are more highly structured than those in RT, with students working in pairs to complete three activities in sequence: partner reading with retelling, paragraph shrinking, and prediction relay. In initial training lessons, teachers explain and demonstrate the three activities. During the partner reading activity, each student reads aloud to his/her partner for five minutes then provides a verbal retelling of the text. While one student reads, the other corrects any reading errors, including omitted, inserted, and incorrectly read words. During the paragraph-shrinking phase, the students alternate reading one paragraph at a time, stopping to state the most important idea in each paragraph before proceeding. The prediction relay activity requires students to alternate reading half a page, make a prediction, then revise the prediction after the next half page has been read. These three activities are completed in the same order each time they are used, usually for 35 minutes a day, three days a week (Fuchs, Fuchs, Mathes, & Simmons, 1997).

The structured nature of these interactions is intended to increase the amount of time students spend actively engaged with text and to improve the quality of feedback they provide each other regarding the accuracy of their oral reading and strategy use (Sáenz, Fuchs, & Fuchs, 2005). In one study, the authors enhance the role of peer feedback in PALS by providing training that stressed the importance of "help[ing] classmates construct their own responses rather than providing peers with correct answers

or simply telling them that answers are wrong" (Fuchs, Fuchs, Kazdan, & Allen, 1999, p. 202),

Collaborative Strategic Reading (CSR). This brand of multiple strategy research was heavily influenced by both Reciprocal Teaching and Transactional Strategies Instruction (described below). Like its predecessors, CSR lessons emphasize the importance of collaborative strategy use. However, CSR focuses more explicitly on student-led cooperative learning, instead of teacher-led groups as in RT and TSI (Klingner, Vaughn, & Schumm, 1998). Also, studies of CSR have targeted classrooms with large numbers of English language learners (ELLs), given that cooperative learning may provide these students with extended exposure to the language of their English speaking peers and opportunities to receive help from their multilingual peers in their native language (Klingner & Vaughn, 2000).

In CSR, students learn to use four major strategic procedures while reading content area texts: a previewing procedure (skimming title and subheadings, making predictions, and recalling background knowledge), a strategy procedure known as "click and clunk" (identifying and clarifying difficult, or "clunky" words), a "get the gist" procedure (identifying and stating main ideas), and a wrap-up procedure (summarizing the text and asking teacher-like questions). These strategies are first modeled and explained by the teacher, then students practice them in small groups of four to six students. Although specific procedures vary slightly from study to study, students are typically assigned roles (e.g., clunk expert) to carry out while reading and practicing the four strategies in groups (Klingner & Vaughn, 2000). Three studies of CRS with upper elementary and middle school students are included in this review.

Think-Aloud Instruction. In this brand of MCSI, students are taught to verbalize the mental processes they are using while completing a task. By verbalizing their thought processes, students make their cognition publicly available for manipulation by teachers and themselves. The use of thinking aloud as an instructional approach was inspired in part by the work of psychologists (e.g., Meichenbaum & Goodman, 1971) who in the 1970s were using cognitive-behavior training to control impulsive behavior (Kucan & Beck, 1997). Their technique was based on a belief that behavioral self-regulation could be accomplished by training children to control their actions verbally. External verbal control would eventually take the form of inner speech that could guide behavior. The belief that verbalizing one's thinking might aid cognitive control is consistent with the view that children's self-regulatory thought processes develop when they re-voice and then internalize adult speech (Vygotsky, 1978).

Comprehension Think-Aloud Instruction can take many forms, but it typically includes a teacher-directed component in which the teacher introduces a comprehension strategy (e.g., summarizing) and then models the use of the strategy with a think-aloud. For example, the teacher may read a page of a text and then say, "Okay, now that I've read a whole page, I want to stop and summarize so I can remember what I've read." Then the teacher will verbalize his or her construction of the text summary so that the students can clearly distinguish the steps. The teacher will continue to model this and other strategies routinely during reading lessons. Also, Think-Aloud Instruction usually includes some type of guided or independent practice during which students work in groups or individually to verbalize their thinking while reading a text.

When teachers model their use of comprehension strategies using think-alouds, students are able to see (really, hear) the mental processes used by expert readers (Collins & Smith, 1982; Davey, 1983). When think-alouds are conducted by students, the teacher can assess how and if students are using strategies successfully, which allows the teacher to fine-tune his/her instructional approaches. Student think-alouds also allow students to hear and discuss the mental processes of their peers (Wilhelm, 2001). Various studies have demonstrated that through these experiences, students become more aware of and better able to control their reading (see Kucan & Beck, 1997 for a review).

Four investigations of Think-Aloud Instruction conducted with upper elementary and middle school students are included in this review. The most influential is a study by Bereiter and Bird (1985), who analyzed think-aloud data obtained from adult readers to identify a set of strategies and then taught these strategies to middle school readers using a think-aloud method. Their instruction consisted of two main phases; in the first, instructors explained the strategies and modeled them by thinking aloud, and in the second, students practiced the strategies by mumbling to themselves while reading independently.

Building on this instructional routine, Baumann, Seifert-Kessell, and Jones (1992) gave fourth-grade students explicit training in how to conduct think-alouds while using self-questioning, predicting, inferring, and retelling strategies, among others. This instruction took the form of ten pre-planned lessons using some of the same metaphors and procedures borrowed from Informed Strategies for Learning. Around the same time, Silvén and Vauras (1992) used think-alouds to model strategy use for sixth-grade students and then guided them in learning to verbalize their own cognitive and

metacognitive processes. More recently, McKeown and Gentilucci (2007) employed Think-Aloud Instruction to help middle school ELLs comprehend content area texts.

Transactional Strategies Instruction (TSI). Like Reciprocal Teaching, Transactional Strategies Instruction is a framework for teaching multiple comprehension strategies that centers on student-teacher dialogue. This particular brand of instruction is "transactional" because the activity of the class is determined by interactions among student, teacher, and the text (El-Dinary, 2002). During these interactions, the teacher's task is to model the coordination of multiple reading strategies and to give students the opportunity to explain their thinking while using the strategies independently (Pressley et al., 1992). In a typical TSI lesson, strategies are practiced during whole class and small group discussions. During these discussions, students and teachers identify comprehension difficulties and jointly resolve these difficulties using a variety of repair strategies. These interactions are not scripted, nor are they overly controlled by the teacher (El-Dinary, 1994; El-Dinary & Schuder, 1993; Pressley et al., 1992). TSI is a constructivist approach to strategy teaching enacted by teachers who "tailor and reshape instruction in light of particular student difficulties and do a great deal to make certain that students understand the nature of the task and the use and significance of the strategies they learn" (Harris and Pressley, 1991, p. 394-5).

The most rigorous and heavily cited study of TSI was conducted with second grade students, and thus is not included in this review. The procedures and outcomes of this study are, however, worth mentioning because of the influence they have had on subsequent versions of TSI. Brown, Pressley, Van Meter, and Schuder (1996) conducted a year-long quasi-experimental study of a teacher-developed version of Transactional

Strategies Instruction called *Students Achieving Independent Learning* (Schuder, 1993). Students in this study were taught when and how to use a variety of reading strategies through extensive teacher explanations and modeling while in their naturally occurring small reading groups and during whole-class activities. After a year of TSI, students showed improvements in their declarative knowledge of reading strategies, their use of reading strategies while reading independently, their use of inferential reasoning while reading, and in reading comprehension.

Two studies of TSI conducted with upper elementary and middle school students are included in this review. In the first, Anderson and Roit (1993) describe a professional development model for helping middle school special education teachers change their beliefs about reading instruction and their instructional practices. The program focused on helping teachers develop a theoretical and dispositional orientation toward teaching reading that includes the core components of TSI. Rather than teaching a series of specific lessons and activities, the researchers introduced a set of teaching *shifts*, each of which specified a tenet of traditional reading instruction and the behavior that replaces it in TSI. For example, instead of focusing solely on having students produce correct answers, the teachers in the study were encouraged to emphasize the process whereby these answers are generated. Videotaped lessons from each teacher's classroom were used during the professional training sessions to elicit discussions about how to implement these shifts. Also, the researchers introduced a list of ways to encourage active reading (e.g., by teaching kids to think aloud while reading). During the training sessions, the teachers discussed these principles and ways to implement them in their classes.

Finally, the teachers in the experimental group received peer support from teachers previously trained in the targeted teaching methods (see also Anderson, 1991).

During this process, teachers were taught to model and explain a number of research-based strategies to their students, including summarizing, questioning, visualizing, and using contextual cues for unfamiliar words. Additionally, they were taught to recognize and foster strategies that students were already using. This focus on emergent strategies makes TSI unique among strategic reading frameworks, which tend to deliver pre-set strategic repertoires with little emphasis on strategies students may have extrapolated from previous interactions with text. Using terms borrowed from Duffy (1992), the authors describe their instructional framework as designed to help students become "strategy generators" instead of "replicators" (Anderson & Roit, 1993).

As with RT, spin-off studies of TSI vary in their adherence to the core design principles of the brand, as exemplified in the other intervention study included in this review. Loranger (1997) taught the four reciprocal teaching strategies to fourth-grade students using a TSI approach. Strategies were introduced during whole-class lessons and then used in reading groups as students collaborated to read selected texts. As in other versions of TSI, the researcher stressed the importance of student collaboration and the naming of specific strategies during textual discussions. However, the lessons and discussions described in this study were more heavily controlled by the teacher than is typical of TSI. Furthermore, the exclusive use of pre-determined strategies separates this study from others in this category.

Concept-Oriented Reading Instruction (CORI). Three studies of Concept-Oriented Reading Instruction with upper elementary students (4^{th} and 5^{th} graders) are

included in this review. These studies are unique because they stress the use of comprehension strategies as tools in service of larger content-related goals.

CORI is an instructional model designed to improve students' literacy achievement by enhancing their reading engagement. Engaged reading in this case does not simply refer to frequent reading; it also refers to reading that is intrinsically motivated, strategic, geared toward conceptual learning, and socially interactive (Baker, Dreher, & Guthrie, 2000; Guthrie, 2001, 2004). Teachers using the CORI model provide direct explanation, modeling, and independent practice in the use of comprehension strategies during science lessons (e.g., how to activate background knowledge before reading a text, how to ask questions before and while reading a text, how to search for relevant information, how to produce concise summaries, how to graphically organize new conceptual knowledge, and how to identity the structure of narratives; Guthrie et al., 2004). Initially, each of these strategies is taught separately. Later, students are taught to integrate each strategy into a strategic repertoire that can be used during collaborative and independent reading (Guthrie, 2004).

Unlike other strategic reading approaches, CORI instruction is made up of various components aimed at improving literacy engagement. First, reading instruction in CORI classroom centers on the development of conceptual knowledge related to thematic science topics, for example, plant life cycles and geological cycles (Guthrie et al., 1998). Second, students are given opportunities for real-world interactions, such as hands-on activities and observational experiences that help them become engaged in the topics being examined. Additionally, students are provided opportunities to read a variety of authentic texts. During a typical 12-week CORI unit, students are expected to read 36

different books. Their selection of books includes fiction chapter books, poetry, legends, and informational books related to the thematic concept being studied (Guthrie, 2004). To ensure active participation by all students, these texts are carefully matched to students' reading abilities, and students are taught specific strategies for selecting books relevant to their personal interests (Guthrie & Cox, 2001). Autonomy support is another critical feature of CORI. When implementing this model, teachers are expected to provide options for their students that give them a sense of ownership for the activities of the class. For example, while investigating a thematic topic, students are given opportunities to generate their own questions, choose subtopics for independent study, and to choose books based on their own interests and needs (Guthrie, 2001; Guthrie, Anderson, Alao, & Rinehart, 1999). Finally, frequent opportunities for student collaboration and sharing of knowledge are also provided, including opportunities to share their areas of expertise with the class in oral presentations and through writing (Guthrie et al., 1996).

Acronym Procedures (POSSE, PLANS, and TWA)

Three different instructional procedures for teaching multiple comprehension strategies to students in grades 4–8 use acronyms to denote the strategy sequences students learn to use. Like the brand-name frameworks, the names of these approaches usually index a very particular set of instructional content and goals, but these approaches have been less widely replicated and studied.

The first of these, influenced by RT and research on expository text structure, is characterized by the acronym POSSE (Englert & Mariage, 1991). In this procedure, students are taught to predict the key concepts in a text based on their background

knowledge of the topic, organize the predictions in a graphic organizer using the text structure as a guide, search for the text structure, summarize main ideas while reading using the graphic organizer, and evaluate their comprehension of the text after reading. Students complete these strategies using preprinted graphic organizers and cue cards that contain sentence starters for each strategy. These materials serve as procedural facilitators (Scardamalia, Bereiter, & Steinbach, 1984) that scaffold students' strategy use until the strategies are fully internalized.

Two studies of POSSE are included in this review, one that describes the use of the acronym as a largely procedularized, teacher-directed activity (Englert & Mariage, 1991), and another that describes the social constructivist roots of the intervention (Englert, Tarrant, Mariage, & Oxer, 1994). The authors of this second study explain that by introducing the acronym procedure, "teachers provided access to the language tools governing comprehension, generated a common language that all members of the community could use to communicate their understandings, and modeled a comprehension language that could help students construct new knowledge as part of a process of learning" (Englert et al., 1994, p. 167). This perspective is unique in this literature. It is the only article that makes such an explicit link between the learning of cognitive strategies and the appropriation of a particular language of comprehension.

A similar instructional framework, represented by two studies in this review, uses the acronym PLAN. Students are reminded to predict the content of a text and make a conceptual map, locate information on the map that they already know from previous experience, add notes to the map as they read, and note a revised understanding of the text after reading by writing a summary. This procedure has been studied in middle

grades science classrooms as a way of helping students as they read their textbook chapters (Radcliffe, Caverly, Hande, & Franke, 2008; Radcliffe, Caverly, Peterson, & Emmons, 2004).

Finally, Mason (2004) describes an acronym program in which students are taught to think before reading, while reading, and after reading (TWA). In this approach to multiple strategy instruction, students use a mnemonic chart as a procedural facilitator that lists the specific strategies for each of the three phases of the procedure. Before reading, students are expected to identify the author's purpose, activate prior knowledge, and decide what they want to learn from the text. While reading, they are prompted to adjust their reading speed and link new content to information they already know. After reading, they restate main ideas, summarize the text, and review what they learned. The TWA procedure was developed based on the principles of self-regulated strategy development (SRSD), an approach directed at helping students with learning and behavioral problems successfully manage a set of strategies for improving their attention and performance in academic tasks (Harris & Graham, 1999). Thus, the goal of this instructional procedure is for students to begin self-regulatively monitoring their use of the mnemonicized strategic steps (Mason, Snyder, Sukhram, & Kedem, 2006; Rogevich & Perin, 2008).

Non-Branded Instructional Approaches

This category catches a wide range of instructional frameworks that go by relatively nondescript names, including cognitive strategy instruction (Lau & Chan, 2007), explicit reading strategies instruction (Van Keer & Verhaeghe, 2005),

metacognitive strategy instruction (Hooutveen & van de Grift, 2007), and comprehension monitoring training (Payne & Manning, 1992). These studies are described as non-branded because they investigate instructional procedures that have been included in only a few studies and have not gained the name recognition of programs like RT, ISL, and PALS. A large portion (40%) of the reviewed studies fall in this last group, many of which have been highly influential in the field (e.g., Block, 1993; Dole, Brown, & Trathen, 1996; Duffy et al., 1986; Jiménez, 1997).

There is tremendous variability in this group in terms of what is taught and how, but they all contain a few common principles. First, each study is motivated by a belief that students can improve their text comprehension by learning to use a multi-pronged strategic repertoire. The strategies included in these interventions are usually taken from previous studies that show positive relationships between individual strategies and improvements in text comprehension. For example, Tregaskes and Daines (1989) justify their use of visual imagery, summarizing, and self-interrogation by citing previous studies that investigated these strategies individually. Similarly, Van Keer and Verhaeghe (2005) cite previous work from a variety of strategic approaches to reading instruction, including ISL, RT, and PALS, to justify the six strategies they selected for their intervention. Many of these studies, especially the more recent ones, can be thought of as attempts to combine previously studied strategies and frameworks into innovative packages. Finally, these studies are fairly uniform in instructional design. Before instruction begins, the researchers select a group of strategies through a review of relevant literature. Students are then taught to use these strategies via a series of lessons consisting of direct instruction, guided practice, and group or independent practice.

A few of these studies stand out as particularly interesting and influential. For example, a study by Duffy et al. (1986) represents one of the first attempts in the field of reading research to convert the skills found in basal readers into volitional strategies. They investigated a professional development model in which teachers were taught to provide explicit verbal explanations of how reading skills could be used as "flexible plans for reasoning about how to remove blockages to meaning" (p. 239). For example, when teaching how to infer meaning of unknown words from a skill-based perspective, the teacher might have students practice this skill using a series of worksheets in which they mark the correct meaning of an underlined word presented in a sentence. From a strategy-based perspective, the teacher might model his/her use of context clues as a repair strategy while reading a story aloud to the class. Consistent with the three types of strategy knowledge included in programs such as ISL (Paris, Cross, & Lipson, 1984), this modeling would include an explanation of what the strategy is and how and when it can be used to disambiguate a confusing word.

In a very different kind of study, Rauenbusch and Bereiter (1991) provided opportunities for students to practice strategic reading using specially modified texts in which every third letter was blocked out. When reading one of these degraded texts, the reader is forced to consciously call on strategies that might otherwise remain hidden, thus providing "frequent occasions for recognizable comprehension breakdowns that could be remedied by applying the kinds of comprehension strategies that have been found to be used by good readers in normal text-reading environments" (p. 184). The researchers first observed seventh-grade readers as they read these degraded texts to identify the strategic moves that were elicited. They taught these strategies to another group of seventh graders

who then practiced these strategies while working in pairs to read additional degraded texts. Although the emphasis on inauthentic texts in this study might not appeal to some researchers, this procedure contributes to the field a potentially useful method for inducing the kind of debugging state that is needed to trigger strategic action and evaluation. This type of activity, used in conjunction with more authentic literacy activities, has the potential to help students recognize the importance of strategic reading and to provide practice in extending their strategic repertoires.

Several studies in this category employ a literature-based approach to strategic reading pedagogy. For instance, Jiménez and Gámez (1996) taught bilingual special education students to use knowledge of cognate relationships to figure out unknown English words, to ask questions, and to make inferences while reading children's books on the theme of Mexican cuisine. Baumann, Hooten, and White (1999) describe how a fifth-grade teacher collaborated with researchers to embed strategy instruction in the context of whole-class and independent reading of trade books. Block (1993) investigated a literature-based instructional framework in which students practiced strategies while reading self-selected books. Finally, Kim (2006) describes an intervention in which students were encouraged to practice strategies while reading independently from books that were mailed to them during the summer. These four studies are unique because they prioritize interaction with authentic narrative text and attempt "to strike an appropriate balance between comprehension strategy instruction and enjoyment of the books themselves" (Baumann, Hooten, & White, 1999, p. 42).

Digital Learning Environments

This last category is included here to highlight a new and promising direction in multiple strategy instruction. Scholars working within a New Literacies framework have urged researchers to foreground the role emerging technologies play in the development of literacy expertise, particularly for adolescent students (Freebody, 2001; Leu, Kinzer, Coiro, & Cammack, 2004; Street, 2003). Although most research on comprehension strategies continues to focus more on print literacy, a few researchers have begun taking advantage of digital media as an instructional tool for improving reading comprehension. Five studies are included in this review that use digital media for strategic reading instruction.

Kim et al. (2006) describe a computer-assisted version of Collaborative Strategic Reading (CACSR) that has been used with middle school students with learning disabilities. Students worked in pairs to use the four CSR strategies (preview, click and clunk, get the gist, and wrap up) while reading and discussing computer-presented passages. During the first phase of instruction (learning CSR), students were introduced to the four strategies, which were described on the screen along with information about how and when to use them. The program included video clips of a teacher using the strategies as a model. During the second phase of instruction (using CSR to learn), students read and discussed on-level passages using digital supports, including a "clunk expert" that provided definitions and pronunciations of difficult words, a dictionary, and a quick strategy review feature that listed the steps for each of the four strategies. As students read in pairs in the final phase (partner practice), they typed their responses to strategy prompts, which were stored in learning logs made accessible to the teacher for

assessment purposes. Additionally, the teacher provided comprehension strategy minilessons at the start of each class before students began working with their partners.

Another digital learning environment included in this review is iSTART, an interactive computerized strategy training program modeled after Self-Explanation Strategy Training (SERT). SERT is an instructional framework that has been used to teach high school and adult readers to verbalize what a text means in their own words as they read (similar to the think-aloud procedure described above) while using a variety of comprehension strategies (McNamara, 2004). The iSTART program teaches the same procedure to seventh and eighth-grade students. In the initial instructional and demonstration phases of the program, students watched an animated pedagogical agent teach the self-explanation procedure and a set of strategies to a group of animated students, who then provided examples of the procedures in action. During the practice phase, the students typed their self-explanations of individual sentences from a science text and received automated feedback.

Proctor, Dalton, and Grisham (2007) have developed a digital strategic learning environment that combines the four strategies of Reciprocal Teaching with vocabulary tools, text-to-speech supports, and bilingual resources. In this environment, designed to supplement regular classroom instruction for Spanish-speaking ELLs, students read four narrative folktales and four expository texts that contained information about the cultures from which the folktales originate. While reading, they were cued to type their responses to strategy prompts in text boxes (e.g., “Now would be a good time to stop and think about the story” [p. 79]). They could click on animated coaches who provided examples and think-alouds of the targeted strategies. One of the animated coaches provided this

support in Spanish as well. After reading a text, students generated written retellings aided by images that cued their memory of the major text content.

A study conducted in Taiwan investigated the effects of a Computer Assisted Strategy Teaching and Learning Environment (CASTLE) on fifth-grade students' text comprehension. In this version of strategy instruction, students practiced generating questions, detecting errors and inconsistencies in text, completing concept maps, highlighting important information, revising concept maps based on text material, making inferences (by filling in missing words), summarizing, and monitoring their use of strategies using a checklist presented on the screen. An animated agent explained each strategy, and then students practiced typing their strategy responses, for which they were given automated feedback.

Finally, Johnson-Glenberg (2005) describes a web-based program called 3D Reader that guided students as they read narrative texts with embedded science content. During reading, they practiced generating typed questions, which were automatically scored by the program. After reading, they practiced creating visual representations of the text content by building sentences from the text using pictographic icons presented on the screen.

The Outcomes of Strategy Instruction

Although the primary focus of this initial review is on the content and pedagogy of strategy instruction, in this section I provide a general overview of the constructs that have been frequently measured in this literature and the impacts that have been reported

by study authors. The five constructs that are measured in 10% or more of the reviewed studies are discussed here, followed by a brief summary of this evidence.

Reading Comprehension

As expected, the most common outcome that was tracked during and after strategic reading interventions is reading comprehension. About 90% of the studies included at least one comprehension outcome. The way comprehension was assessed varied from study to study. Forty-two percent of the reviewed studies used at least one standardized measure of reading comprehension. Some of these studies (about 12% of the total sample) used both standardized and researcher-designed measures of comprehension, while 48% of the studies exclusively used researcher-designed measures. The most common standardized measures included the comprehension subtests of the Gates-MacGinitie Reading Test, Metropolitan Achievement Test, Woodcock-Johnson Tests of Cognitive Ability, and the Stanford Diagnostic Reading Test. When researcher-designed measures were used, they usually included multiple-choice or short answer formats that students answered in written form.

Strategy Knowledge

Given that the studies included in this review were specifically designed to transfer strategy knowledge to students, it is surprising that only about one-fourth of the studies directly measured students' declarative knowledge of strategies. In these 21 studies, strategy knowledge was assessed in various ways, including researcher-designed strategy awareness questionnaires (Houtveen & Van den Grift, 2007: Walraven, 1993), a

multiple choice strategy knowledge test (Paris, Cross, & Lipson, 1984), interviews (Duffy, Roehler, & Meloth, 1986; Klingner & Vaughn, 1996), and qualitative analysis of verbal think-aloud data (Baumann, Hooten, & White, 1999; Jiménez, & Gamez, 1996).

Strategy Application

Forty-three percent of the studies measured students' ability to use strategies while reading. As with strategy knowledge, no standardized measures were available for this construct, so researchers relied primarily on locally developed measurement tools. The most common measurement methods included analysis of think-aloud data (Rauenbusch & Bereiter, 1991), strategy interviews conducted while students read and used strategies aloud (Carriedo & Alonso-Tapia, 1995; Klingner, Vaughn, Arguelles, 2004), analysis of video and audio of students working together during instructional lessons (Anderson & Roit, 1993; Klingner & Vaughn, 2000), and performance assessments during which students read texts and responded to strategy prompts (i.e., requests to write a summary, generate questions, etc.) in writing (De Corte, Verschaffel, Van de Den, 2001; Palincsar, Brown, & Martin, 1987).

Comprehension Monitoring

Twelve studies (15%) in this sample report outcomes for comprehension monitoring, defined here as a student's ability to detect breakdowns in understanding while reading (Wagoner, 1983). Eight of these studies measured this construct using an error detection test. In this type of assessment, students are presented with a passage that contains inconsistent information and are asked to mark that information while reading

(e.g., Sung, Chang, & Huang, 2008; Tregaskes & Daines, 1989). By successfully identifying the incongruencies in a text, students demonstrate their ability to notice comprehension breakdowns. In a few studies, this construct was measured using interviews or questionnaires (e.g., Carr & Borkowski, 1989; Lau & Chan, 2007), and in one study the researchers tracked eye movements to assess the amount of time students spent attending to textual inconsistencies (Kinnunen & Vauras, 1995).

Reading Attitude

In addition to the constructs described above, ten studies (12%) also investigated the gains made by students in their motivation or attitude toward reading. The authors of these studies relied on attitude questionnaires to measure this construct, including Wigfield and Guthrie's (1997) *Motivation for Reading Questionnaire*.

Summary of Reported Impact of MCSI

The variety of methodologies and measurement instruments employed in this literature makes it difficult to concisely summarize the instructional impact of these studies. That being said, a few general conclusions can be drawn:

(1) There is substantial evidence for the effectiveness of MCSI in improving middle grades students' comprehension. Positive results are reported for struggling and average comprehenders as well as for students with reading or learning disabilities. There is also evidence that strategy instruction improves the comprehension of students who are learning to read

in a second language, but more studies are needed that focus on multilingual students.

(2) Even though the reported impacts of strategy instruction on reading comprehension are mostly positive in these studies, a number of researchers found little or no change in reading comprehension ability, suggesting a need for closer and more methodical analysis of the differences between interventions that produced positive outcomes and those that did not.

(3) Improvements in students' declarative knowledge of strategies and application of this knowledge while reading have also been noted, although these constructs are less commonly measured than reading comprehension. Positive findings for these constructs have been reported for students with reading disabilities, struggling comprehenders, and average readers.

(4) Some studies have shown that struggling and average comprehenders improve their ability to monitor comprehension while reading as a result of strategy instruction.

(5) Improvements in reading motivation or attitude have also been reported, but more work is needed in this area to form a stronger case for this finding.

The Content of MCSI

Authors often describe the content of multiple strategy instruction in terms of the specific strategies that are taught to students. The goal of this section is to synthesize across the studies to identify which strategies are taught most frequently, which ones are typically left out, and of these, which ones might warrant further consideration. Because

each individual strategy can go by various names, it is difficult to compare strategy content across studies. To address this problem, I compiled a list of all the strategies included in each study (as named and described by the authors) and used these to produce a taxonomy of 31 distinct strategies. Table 2 summarizes the strategies that have been included in interventions for students in grades 4-8 in order from most to least common.

Commonly Taught Strategies: The Big Four

Not surprisingly, the four most commonly taught strategies for students of this age are those that are included in Reciprocal Teaching: summarizing (used in 77% of the reviewed studies), predicting (64%), generating questions (62%), and clarifying (57%). Although they have a common origin, these strategies are not uniformly applied in every study.

In some studies, students are taught to pause to summarize periodically (e.g., after every page or paragraph). In a study by Payne and Manning (1992), teachers led students through a series of strategies as they read their traditional basal materials. One of the main activities in these lessons was stopping to "summarize at various points." As explained in the sample lesson plan provided in the research report, students were given specific pages to read and then asked to review the main events from those pages afterwards by summarizing. In their study of Think-Aloud Instruction, Silvén and Vauras (1992) also taught students to summarize periodically while reading. Students were taught first to identify the most important messages in a segment of text and then to connect these messages across the text.

Interestingly, very few studies describe procedures for teaching students to decide on their own when to stop and summarize. Instead, these studies tend to favor instructional procedures whereby teachers control where and when students monitor their comprehension through summarization (see Anderson & Roit, 1993 and Rauenbusch & Bereiter, 1991 for exceptions).

Palincsar and Brown (1984) describe periodic summarization as a monitoring strategy: if a student is able to synthesize the most important information in a section of text, he or she should feel confident that adequate comprehension is occurring. If not, the student should use a repair strategy to ensure that an appropriate summary can be generated for the text. In most cases, the use of summarization as a monitoring strategy was done through collaboration with other peers who jointly evaluated each other's summaries for accuracy (e.g., in RT and PALS). In other studies, students produced summaries on their own or as part of teacher-directed lessons. In a study by Kitano and Lewis (2007), gifted students, many of whom were English language learners, learned a series of comprehension strategies from undergraduate students who were trained as tutors. In this case, students were taught to summarize while they read without being given opportunities to share their summaries with peers. Similarly, in an early study of Think-Aloud Instruction, students practiced summarizing while reading alone (Bereiter & Bird, 1985). They were taught to indicate their strategy use in writing by marking the text and were provided some opportunities to verbalize their use of strategies to the teacher as he/she moved around the room.

Summarization is also taught as a post-reading strategy for reviewing newly learned information at the end of a text. For instance, in a study of the Computer Assisted

Strategy Teaching and Learning Environment (CASTLE), students were taught to summarize digitally presented texts using a series of summarization macrorules (Sung, Chang, Huang, 2008). This procedure required students to identify important information, remove non-essential information, and to synthesize main ideas across the text to form a coherent summary (Brown & Day, 1983). Similarly, Collaborative Strategic Reading (e.g., Klingner, Vaughn, & Schumm, 1998) relies on a post-reading "wrap-up" procedure that includes summarization. When applied in this way, summarizing serves as a memory aid that helps students solidify what they have learned from a text rather than as a comprehension-monitoring tool.

The prediction strategy is also employed in different ways in the intervention studies. Often, prediction is taught as a pre-reading strategy. For example, in their study of POSSE, Englert and Mariage (1991) describe a prediction procedure in which students used their background knowledge to hypothesize the content that would appear in the text before they began reading and then arranged this content on a graphic organizer provided by the instructor. Radcliffe, Caverly, Peterson, and Emmons (2004) describe a similar prediction routine in their version of strategic reading using the PLAN acronym. In Collaborative Strategic Reading, predicting is taught as a previewing strategy (Klingner, Vaughn, Arguelles, Hughes, & Leftwich, 2004). When used in this way, the prediction strategy helps students develop an anticipatory frame for the text being read.

Prediction can also be taught as a procedure to use while reading. Sometimes, students are directed to stop at specified intervals (every paragraph or page) and generate a prediction, either on their own or with a partner. This is the procedure used in Reciprocal Teaching with peer tutors (Palincsar, Brown, & Martin, 1987). In Peer-

Assisted Learning Strategies, students participate in "prediction relays"—repeated cycles of predicting, reading half a page of text, and then evaluating their previous predictions and making new ones for the next segment of text (Sáenz, Fuchs, & Fuchs, 2005).

Question generation is also employed in different ways in this literature. There are studies in which students stop to ask and answer their own questions at specified intervals, often with a partner or small group, in order to monitor their comprehension. Typically, the questioning strategy is presented as a way of reviewing information learned in a text. An example of this type of self-questioning can be found in Carr and Borkowski's (1989) study of comprehension strategy training for underachieving students. These students were taught to generate questions after reading a paragraph that could be answered by stating the main idea of the paragraph. Similarly, Taylor and Frye (1992) taught students to produce questions after reading a section of the their social studies text. When used in this way, the questioning strategy is not intended to promote an inquisitive stance toward text content but rather to promote self-review of text information. Even in studies of RT, which tend to foreground student collaboration and dialogue, the questioning strategy is used primarily to help students generate "teacher like questions" or questions about important content (Palincsar & Brown, 1984). Thus, this strategy is more about attention allocation than curiosity.

Although less frequent, there are a few instructional frameworks in which students are taught to generate, discuss, and act on their self-selected questions. One example of a study that takes this orientation is Souvignier & Mokhlesgerami's (2006) program for teaching strategies using a self-regulation perspective. In their instructional routine, students were taught to become "text detectives," a process that involved

learning to check understanding by asking questions that were not answered explicitly on the page. They were taught to distinguish between questions that could be answered using surface-level facts and those that required in-depth understanding of the text (e.g., "why" questions). Although these questions were conceptually deeper, they were still used as a self-review mechanism rather than as genuine interrogation of text content. In contrast, Jiménez (1997) provides examples of how bilingual readers can learn to use a cognitive questioning strategy as a way of seeking meaning from textual interactions. When used like this, questioning helps students connect text information to background knowledge and sets students up to successfully infer unstated information.

The use of clarification as a strategy is particularly inconsistent across the studies in this review. In fact, the use of the term clarification to denote a single strategy is deceptive. Clarification implies a number of metacognitive and strategic moves, namely, the identification of comprehension breakdowns and the implementation of strategies to repair these breakdowns.

In the early studies of Reciprocal Teaching, the clarification strategy is described as a process of collaboratively resolving word-level and text-level breakdowns in understanding (Palincsar & Brown, 1984). In many of the RT replication studies, however, clarification is taught primarily as a strategy for inferring the meaning of unknown words or phrases. For instance, in one study (Takala, 2006), students were given a list of "rules of a good reader" that defined clarifying as follows: "If there are words you don't know, find out what they mean" (p. 563). During the implementation of this intervention, students worked in pairs to identify unknown terms and to clarify their meanings. Similarly, in studies of Collaborative Strategic Reading, students were taught a

process called “click and clunk.” This procedure, which was often labeled as a comprehension monitoring procedure, is described as a process of finding unknown words and using fix-up strategies to figure out what they mean (Klingner & Vaughn, 2000). In one study of CSR, the authors provided examples of students’ identifying “clunks”, all of which were vocabulary words unfamiliar to the students (Klingner, Vaughn, Schumm, 1998).

Clarification is used less commonly to repair comprehension breakdowns at the sentence or paragraph level. In Loranger’s (1997) study of Transactional Strategies Instruction, students were given strategy cards listing sentence starters they could use for the clarification strategy. These statements included: “This just doesn’t make sense” and “This is different from what I expected” (p. 36). Although this study stands out as one of the few in which the clarification strategy was explicitly defined in text-level terms, it is not made clear in the research report how often students used this strategy to clarify text-level difficulties. In fact, the one example of student discourse for the clarification strategy provided in the article depicts the clarification of a short phrase, not a larger concept. In a study of cognitive strategies for ELL students in secondary schools, Olson and Land (2007) also provide examples of sentence starters students used for discussing strategy use. For clarification, the authors list statements like: “To understand better, I need to know more about…” and “Something that is still not clear is…” (p. 280). As with the Loranger study, it is not clear how often students actually employed this dialogue; but this example does highlight one way in which students use of clarification at the text-level can be scaffolded.

Finally, in some studies, the clarification strategy is listed as a component of the instruction, but it is unclear whether students were taught to clarify words, larger chunks of meaning, or both (e.g., Labercane, 1987; Lederer, 2000; LeFevre, Moore, & Wilkinson, 2003).

Other Common Strategies

Other common strategies include using prior knowledge (42%), identifying main ideas (35%), rereading to repair comprehension (22.2%), and monitoring comprehension (21%). Like the most common strategies, these are operationalized in different ways in the research literature. Prior knowledge, for example, is used to aid prediction making during the pre-reading phase in some studies, including studies of the POSSE procedure (Englert & Mariage, 1991), Collaborative Strategic Reading (Klingner, Vaughn, Arguelles, Hughes, & Leftwich, 2004), and Mason's TWA approach (2004). In other studies, prior knowledge (or background knowledge, as it is frequently called) is used as part of a "making connections" strategy. For example, in a recent study of strategy instruction based on principles of self-regulation, Rogevich and Perin (2008) taught students to "link their background knowledge to information from the text that is new for them" (p. 137). Kitano and Lewis (2007), drawing on Harvey and Goudvis's (2000) *Strategies that Work* framework and Keene and Zimmermans' (1997) *Mosaic of Thought* framework, encouraged students to make "text-to-self connections" as a way of linking textual information with information they already know from personal experience.

In many studies, the main idea identification strategy is virtually indistinguishable from the summarization strategy described above in that students are expected to find and

verbalize the most important piece of information in a text section (usually in ten words or less). In some studies, this was used as a post-reading review activity. For example, in TWA instruction (Mason, Snyder, & Sukhram, 2006), thinking about main ideas was included as a strategy to be used after reading a paragraph. Students used the acronym RAP (from Ellis & Graves, 1990), which stands for: Read the paragraph, Ask yourself, "What is the main idea?", and Put the main idea in your own words. What distinguishes the main idea strategy from summarization in these studies is that summarization involves the synthesis or coordination of multiple main ideas across larger sections of text. In other studies, main idea identification is treated as an attention allocation strategy. Students are taught to distinguish between crucial and non-crucial text information so that their "memory processes will not be overstrained" (Silvén & Vauras, p. 74).

Although comprehension monitoring is often listed as a separate strategy, it is more accurate to think of it as a process inherent in all self-regulated strategy use. Comprehension monitoring is a process of self-appraisal readers use to evaluate whether or not a text has been understood (Jacobs & Paris, 1987; Wagoner, 1983). For instance, highly self-regulated readers might periodically ask themselves, "Do I understand what the text is saying?" This is the process by which students demonstrate their development of metacomprehension. Because this is a fundamental process in strategic reading, many studies that do not explicitly list comprehension monitoring as a discrete strategy still expect students to monitor whether or not text comprehension is proceeding as it should. In RT, for instance, students are expected to use summarization as a comprehension monitoring strategy.

Comprehension monitoring only results in improved understanding when it is accompanied by strategies for repairing comprehension; a reader who recognizes a comprehension breakdown must select appropriate strategies for repairing this breakdown. The most common repair strategy in the literature reviewed here is rereading, which is a process of referring back to previously read content to clarify a misunderstanding (Babbs, 1984). Other repair strategies include reading ahead (Taylor & Frye, 1992) and using social resources (i.e., asking others for help; Pickens & McNaughton, 1988). Clarification can also be thought of as a repair strategy, since it involves both recognizing and repairing comprehension difficulties.

In many studies, the authors combine the processes of metacomprehension and comprehension repair under the single label of comprehension monitoring. For example, Van Keer and Verhaeghe (2005) define comprehension monitoring as a process in which students "monitor comprehension and regulate understanding of difficult sentences or passages by rereading, adjusting reading speed, or tracing the meaning of unfamiliar words or expressions" (p. 303). In their study of Informed Strategies for Learning (ISL), Paris and Oka (1986) provide a nearly identical definition of comprehension monitoring. Described in this way, comprehension monitoring is virtually indistinguishable from the clarification strategy described earlier. Both are two-step processes that involve noticing and then repairing comprehension breakdowns.

Less Common but Potentially Important Strategies

There are some strategies that appear in this literature with surprising infrequency, given that strategic reading pedagogy has been heavily influenced by theories of

metacognition and self-regulated learning. The strategies described in this section warrant additional attention from researchers working in this field because they have the potential to help middle grades students as they make the transition from primary to secondary education.

Setting a purpose for reading. The purpose or goal a student sets when reading a text provides the standard by which all subsequent strategic moves are evaluated (Paris & Jacobs, 1984). A student whose purpose is to enjoy an entertaining narrative (when reading a mystery novel, for example) will likely focus primarily on major plot events and personal connections to characters and their actions. A strategic reader will evaluate whether or not the desired goals are being met while reading and make adjustments when necessary. When reading a science text, however, the student's purpose may be to learn the information required to score high on next week's test. This purpose will initiate a very different self-regulation cycle in which the student reviews important content periodically, takes written notes, and connects the new information to chapters read in previous weeks.

Setting a purpose for reading is listed as an instructed strategy in only 15% of the reviewed studies. In a study by Manset-Williamson and Nelson (2005), volunteer tutors taught middle school students to set process goals before reading a text. Informed Strategies for Learning (ISL) also emphasizes purpose setting as a strategy that helps structure a reader's subsequent strategy use. From a self-regulation perspective, one might expect setting a purpose for reading to play a more prominent role in this literature.

Monitoring strategy use. Monitoring is a common term in this literature. Sometimes students are taught to monitor their comprehension, and other times they learn

to monitor their knowledge of unfamiliar words and phrases or to monitor a text for inconsistent information. One type of monitoring that is especially promising as a way of improving the comprehension of middle-grades students is the monitoring of strategies that have been selected to determine if they are meeting the goals set for a particular reading episode (Souvignier & Mokhlesgerami, 2006). This metacognitive process is rarely included as a discrete strategy in this literature. Most instructional programs require strict adherence to a particular sequence of strategy use—for instance, students are often expected to follow the strategies in the order they are listed on a cue sheet (TWA, PLANS, POSSE) or to follow a teacher-directed routine (PALS).

In less than 9% of the reviewed studies, students were explicitly taught to evaluate their use of strategies and to make revisions to their strategic moves when desired learning goals were not being met. In Wise, Ring, and Olson's (2000) study of strategy instruction, fifth-grade students practiced "justifying how and why they used a strategy at a particular point in a story" (p. 207), a process that required students to think about the suitability of the strategies they chose to use. De Corte, Verschaffel, and Van De Ven (2001) also implemented a metacognitive strategy they called "regulating one's own reading process," which required students to regularly evaluate and modify their selected strategies. Anderson and Roit (1993) took a similar approach in their program for training teachers to use TSI. Teachers were encouraged to model reading as a problem-solving activity in which a reader attempts a set of strategic moves to enhance comprehension, evaluates the effectiveness of these moves, and then tries another set of strategies if needed. Throughout this program, students and teachers learned to view strategies as tools that could be used to address specific reading problems rather than as procedures to

be carried out in a prescribed sequence. These studies stand out as examples of how self-regulation can be incorporated into multiple strategy instruction.

Using social resources. The importance of social help-seeking as a sustainable and useful strategy is often overlooked in this literature. Although a large number of strategic reading frameworks include procedures for applying and discussing strategies in cooperative groups (e.g., RT and PALS), reading is primarily viewed as an individual activity in this literature. In most of these studies, the goal of collaborative interaction is stated in individualistic terms, for example, to help students internalize strategies so they can use them later on their own.

Very few studies (5%) include the use of social resources as a purposeful strategy for repairing or enhancing comprehension. The studies that include social resources provide only sparse explanations of how this strategy is taught or used. A typical example is Babbs' (1984) study in which students used monitoring cards that listed "ask someone" as a strategy for repairing comprehension breakdowns. Similarly, Pickens and McNaughton (1988) included help-seeking as a repair strategy in their study of peer strategy tutoring. But even these studies describe the use of social resources as a last resort. None of the reviewed studies acknowledges that even highly skilled readers, after demonstrating full mastery of a set of strategies, can benefit from strategically using social resources to enhance their understanding of a text.

Analyzing and evaluating. Much has been written about the increased demands of upper elementary instruction compared to primary grades instruction. As students transition from third to fourth grade, for example, they are expected to engage in a variety of self-regulatory academic behaviors as they navigate complex content-area texts (Chall,

Jacobs, and Baldwin, 1990; Pressley, Wharton-McDonald, Mistretta-Hampston, Echevarria, 1998). Comprehension alone is not adequate as students get older. They must also be able to gain new insights and knowledge from text that can be applied in other academic settings. Thus, there is a strong need for strategic reading frameworks to include strategies for deeper analysis and interpretation of text for older students.

Although many of the reviewed studies make a clear argument for the importance of learning from text, only 6% of the reviewed studies explicitly taught strategies for analyzing or evaluating text content. These include Block's (1993) literature-based strategy program, which, although focusing mostly on narrative text, strongly prioritized analytical strategies like character analysis, comparing and contrasting across texts, and identifying the point of view of a story. The brand of studies that focuses most heavily on analysis and evaluation of text content is Concept-Oriented Reading Instruction (CORI), which, as noted previously, is unique in its emphasis on conceptual learning goals (Guthrie et al., 1996). To be clear, the inclusion criteria used in this review limited the collected studies to those emphasizing comprehension strategies, so there may be other studies that focus more on analysis and evaluation that were not included. Still, this literature would benefit from the inclusion of analytical and evaluative goals alongside comprehension goals, given the expectation that older students will use texts to build background knowledge about a variety of thematic and conceptual topics.

The Role of Emergent Strategies

In 87% of the studies reviewed, instruction began with a pre-determined list of strategies with little or no room for additional strategies to emerge as lessons occurred. In

their study of TSI, Anderson and Roit (1993) describe the distinction between preplanned and emergent strategies as follows:

> Strategy instruction typically involves the passing on to students of one or more prescribed, experimenter-determined strategies, with little regard for whether learners have existing strategic knowledge. Adolescents, especially, cannot be viewed as strategic tabulae rasae, although they clearly have difficulty applying their strategic abilities to reading…. Although the teachers and students were receptive to strategies based on research, such as summarizing, question asking, imaging, and using context, the teachers made efforts to draw out existing strategies and to help students judge the efficacy of their strategies in light of the problems and texts at hand." (p. 126)

A few additional studies made room for students' pre-existing strategies to gain footing. For instance, Rich and Blake (1994) selected their strategy lessons from a checklist participating students created while watching videos of social studies lessons and annotating the cognitive and affective strategic behaviors they observed. In their study of a literature-based strategy method, Baumann, Hooten, and White (1999) describe two types of strategy lessons: planned lessons, during which specific strategies were introduced and practiced, and unplanned lessons, during which teachers took advantage of "teachable moments" to extend or clarify strategies based on how students were using them during class discussions.

In contrast, most of the other studies included in this review were less responsive to the strategic thinking abilities students brought with them at the start of the interventions. This curricularization (Fisher & Frey, 2008) of strategies goes against the theoretical assumptions underlying strategic reading. If strategic readers are viewed as active problem solvers who flexibly coordinate strategies in the service of specific comprehension goals, then strategy instruction at its best must be responsive to a variety of personalized ways of interacting strategically with text (Harris & Pressley, 1991).

The Role of Strategy Coordination

The most common form of teaching students to coordinate their use of multiple strategies is exemplified by Souvignier and Mokhlesgerami (2006). The last four lessons in their 20-lesson program were designed to help students "join the seven strategies together in an integrated routine that students can rely on" (p. 62). Other studies accomplished strategy coordination not by devoting individual lesson modules to the topic, but by focusing on flexible strategy selection throughout the instructional series. For instance, when reading degraded texts in Rauenbusch and Bereiter's (1991) study, students practiced selecting from a variety of strategies to help them understand the meaning of the text. This emphasis on flexible strategy selection from the very start of the intervention is also integral in studies of Transactional Strategies Instruction (Anderson & Roit, 1993) and Think-Aloud Instruction (Baumann, Seifert-Kessell, & Jones, 1992).

Although the field has recognized the importance of multiple strategy coordination for several decades now, nearly one-third of the studies included in this review do not address strategy coordination at all. The underlying assumption in these studies seems to be that once students internalize a canon of individual strategies, they will automatically begin using these strategies in coordination. It might seem reasonable for older studies to ignore strategy coordination, but many of the studies in this sample that do not address this important issue were conducted during the past decade (e.g., Kim, 2006; Little & Richards, 2000; McCrudden, Perkins, & Putney, 2005; Radcliffe, Caverly, & Peterson, 2004).

The role of strategy coordination in studies of RT is difficult to pin down. On the one hand, RT was one of the first instructional frameworks to prioritize the coordination of multiple strategies. The original RT publication (Palincsar & Brown, 1984) is most notable for its insistence that multiple strategies be learned through repeated use instead of through didactic instruction (Rosenshine & Meister, 1994). Most instructional programs that teach multiple strategies owe something of their development to this method. On the other hand, the original and spinoff studies of RT tend to describe summarization, questioning, clarifying, and predicting as if they are always used in sequence, with little discussion of flexible strategy coordination (i.e., the need to select different strategies depending on need and purpose). Recent work makes it clear that Palincsar and Brown did not intend for their RT activities to be implemented as one-strategy-at-a-time lock-step routines (Palincsar, 2007), but that is how they come across in the early publications. Consider for example the description of the procedure provided in the 1984 study:

> The adult instructor assigned a teacher for the first segment (usually a paragraph) and the group read the segment silently. Then the teacher for that segment proceeded first to ask a question, then to summarize, and to offer a prediction or ask for a clarification when appropriate. (p. 131)

When implemented in this way, RT provides few occasions for students to practice strategy coordination. To be fair, the authors go on to explain that as students internalize the mental processes over time, the interactions begin to take on a more conversational form. However, there are few examples of these conversational interactions in practice in this body of work.

The Pedagogy of MCSI

Whereas the previous section describes what upper elementary and middle school students are expected to learn in various versions of multiple strategy instruction, this section describes how that learning occurs. To address this question, I examine the instructional and textual contexts in which strategic reading pedagogy occurs and the roles of self-regulation and student collaboration in these contexts.

Instructional Context

In this review, instructional context was operationalized by considering the role of the instructor in the development and implementation of strategy instruction, the duration, location and grouping of the instructional activities, and the relative emphases on different instructional modes.

Location and grouping. Many of the researchers working in the field of reading research have heeded recent calls to locate more intervention studies in ecologically valid classroom contexts (Paris & Paris, 2001). Most of the studies reviewed here were conducted during regular classroom instruction (57%) or as part of in-school pull-out programs (38%). Among the studies that took place during typical classroom instruction, most took place during reading lessons, and a few during science or social studies content lessons. A few studies were conducted in afterschool or enrichment programs (5%). In about half of the studies (52%), students received most of their instruction during whole-class lessons, although many of these did provide additional practice in smaller groups or individually. Thirty-eight percent of the studies relied primarily on small group instruction, and in the remaining 10%, students mainly received one-on-one instruction.

Duration. There is tremendous variability in the amount of strategy instruction students received. Among the studies for which duration of instruction could be estimated, the amount of instruction ranged from five sessions to a full school year (estimated to be about 96 individual sessions), with a median of twenty sessions. The average instructional session was around 45 minutes long. In the shortest instructional programs, students received an estimated 100 minutes of instruction compared to over 4000 minutes in the longest programs (median = 810 minutes).

Examples of the shortest interventions include McNamara, O'Reilly, and Best's (2006) iSTART program, which provided students with two 50-minute sessions of computerized strategy practice, Johnson-Glenberg's (2005) web-based strategy program, which included one introductory lesson and four half-hour practice sessions, and an instructional framework described by Nolan (1991) in which students learned self-questioning and predicting strategies during three one-hour sessions. Some of the interventions of average duration include Dole, Brown, and Trathen's (1996) strategic reading intervention, which included 18 preplanned 50-minute lessons spread out over five weeks of instruction (900 minutes total). Of similar duration was the "text detective" program described by Souvignier and Mokhlesgerami (2006), which consisted of twenty 45-minute lessons that combined strategy instruction with motivational and self-regulation training. Two of the most intense instructional frameworks were CORI (Wigfield, Guthrie, & Perencevich, 2008), which included 90 minutes of instruction daily (in reading and science) for 12 weeks, and PALS, which in one study was implemented for about 100 minutes a week for 21 weeks (Calhoon, 2005)

Instructional personnel. Perhaps the most important component of any instructional context is the teacher who plans and monitors the instructional activities. In the reviewed studies, instruction was most often delivered by members of the research team (44 % of the studies) or classroom teachers (40 %). In other studies, the instruction was delivered jointly by teachers and researchers (4%), by tutors or volunteers (4%), and through interaction with computer programs or digital pedagogical agents (7%).

When classroom teachers were responsible for teaching strategy lessons, they typically had little agency in the design or implementation of the instruction. In about half of these studies (18 out of 35), the researcher began the intervention by providing training to the teachers on a set of lessons and activities that pre-dated the teachers' involvement. The studies of generic instructional frameworks conducted by Taylor and Frye (1992) and Payne and Manning (1992) typify this trend. Teachers in studies of Peer-Assisted Learning Strategies attended full-day workshops where they learned specific procedures for implementing PALS, including how to assign students to dyads and how to enact the scripted lessons provided in the training manual (Fuchs, Fuchs, & Kazdan, 1999; Fuchs, Fuchs, & Mathes, 1997). In studies of Informed Strategies for Learning, a similar procedure was followed. Teachers received lesson plans and bulletin board materials from the researchers, which they then implemented as directed (Paris, Cross, Lipson, 1984).

There are a few examples of more collaborative endeavors between teachers and researchers. These can be aligned along a continuum of increasing teacher agency. At one end of this continuum are studies in which the researcher specifies the strategies and the general instructional approach, and the teacher is then given flexibility for how and when

to teach each strategy (e.g., Tregaskes & Daines, 1989). At the higher end of the continuum are studies in which teachers and researchers fully collaborate to develop the instructional materials and procedures, such as in Rich and Blake's (1994) study of a generic instructional framework, which was developed in collaboration with special education teachers, a speech therapist, and classroom teachers at the school site. Also in this category are interventions in which researchers develop their programs with input from classroom teachers along the way (e.g., Anderson and Roit's [1993] collaborative development of a TSI program). In the middle of this continuum are instructional approaches like CORI in which teachers attended summer workshops and then integrated the instructional practices in their reading and content-area instruction (Guthrie, Anderson, Alao, & Rinehart, 1999). The distinguishing feature of the more collaborative studies is that instead of learning to replicate specific instructional routines, teachers were scaffolded in their development of "conceptual models for what it means to be strategic" (Duffy, 1993, p. 244).

Modes of instruction. Most studies included in this review used some version of a faded or scaffolded instructional sequence, based loosely on Pearson and Gallagher's (1983) often-cited gradual release of responsibility model. This model depicts an instructional cycle containing an instructor-directed phase, a joint responsibility phase, and a student practice phase.

The instructor-directed phase took two major forms: direct explanation of strategies and process modeling. Almost every intervention included direct explanation to some extent as part of the instructional program, and a few studies seemed to emphasize this mode of instruction over all others. For example, in studies of Informed Strategies for

Learning, instructional time was dominated by descriptions of individual strategies and how and when they should be used (Paris, Cross, Lipson, 1984). Similarly, when teachers implemented the acronym strategy lessons (e.g., TWA), they spent a large portion of the introductory lessons explaining the individual pieces of the strategy routine (Mason, Snyder, Sukhram, & Kedem, 2006).

Although almost all of the studies also used teacher modeling to some extent, it is possible to identify a few studies that clearly prioritized teacher modeling as an important avenue for strategy learning. Think-Aloud Instruction clearly falls in this category since it was designed specifically to help teachers make the covert mental processes of comprehension visible for students through teacher modeling (Bereiter & Bird, 1985; Kucan & Beck, 1997). Also, three of the digital learning environments provided extensive strategy modeling in the form of written and animated think aloud demonstrations (Kim 2006; McNamara, O'Reilly, & Best, 2006; Proctor & Dalton, 2007).

Joint responsibility, or assisted practice, took a number of forms in the reviewed studies. Sometimes, assisted practice took the form of teacher-guided small group dialogues, such as those that occurred during Reciprocal Teaching (Palincsar & Brown, 1984). Other times, it took the form of whole-class conversations as described in studies of Transactional Strategies Instruction (Loranger, 1997) and Informed Strategies for Learning (Paris, Cross, & Lipson, 1984). Also, several of the generic instructional models describe teacher-guided whole-class activities. For instance, Dole, Brown, & Trathen (1996) provide a sample lesson plan that depicts a teacher providing directions and strategy prompts as students read short sections of text on their own and then discuss

them with the whole class. In their study of metacognitive strategy instruction, Houtveen and van de Grift (2007) characterize assisted practice as a process in which "the teacher gives careful feedback to the students on how to improve implementation of the strategy" (p. 177). In most studies, assisted practice activities are described very generally, using phrases like "decreasing teacher responsibility" and "increased student control" to describe the shifts in teacher and student roles.

Finally, the "all student," or student practice, phase was enacted in several formats, including strategy practice during independent reading and during small-group reading. A few studies heavily emphasized student experience with text as a necessary component of strategic reading development. Most notably, studies of RT, PALS, and CSR included extensive practice in reading and using strategies in structured collaborative learning groups. Several of the generic instructional frameworks also employed peer tutoring or collaborative reading activities (e.g., Pickens & McNaughton, 1988; Van Keer, 2004; Van Keer & Verhaeghe, 2005). Other examples include Proctor and Dalton's (2007) digital literacy environment, which gave students extended strategy practice while reading eight digital texts independently, and Kim's (2006) summer strategy intervention, which sought to increase students' strategic engagement outside of school. Also, Concept-Oriented Reading Instruction is based on a model of reading development that heavily prioritizes extended engagement with text (Guthrie, Van Meter, & McCann, 1996). These studies stand out as those in which students spent as much or more time reading and using strategies on their own (or with a partner) than they did listening to teachers describe and model strategies.

Self-Regulative Context

A common learning goal stated in the intervention studies reviewed here is that students should learn to be self-regulated strategy users. This goal is described in various ways. For several of the strategy instructional brands (e.g., ISL and TWA), the authors explicitly draw on previous research in self-regulated learning as a theoretical justification for their intervention procedures. Other studies that do not explicitly draw on the self-regulation literature nonetheless assume that students should eventually internalize—i.e., come to use independently and automatically—a set of strategic procedures that are first practiced in interaction with others (Vygotsky, 1978).

Despite this theoretical basis in self-regulation and related literatures, few of the interventions in this review provide students with opportunities to use strategies in self-initiated, independent ways. Only a handful (about 19%) of the instructional programs reviewed here emphasize students' abilities to use the strategies on their own in situations where they are not prompted by teachers or prompt cards. Even less common are studies that treat strategies as mental tools that have to be used flexibly depending on the specific reading goals and contexts (17% of the reviewed studies). Clear examples of flexible strategy use can be found in descriptions of Think-Aloud Instruction (Baumann, Seifert-Kessell, & Jones, 1992) and Transactional Strategy Instruction (Anderson & Roit, 1993), both of which are notable in that students are expected to adopt a strategic stance rather than a particular strategy sequence or routine.

The majority of the studies primarily provide strategy practice in the form of instructional procedures that are heavily regulated by the teacher. For instance, when students participate in the PALS procedure, they are expected to use the strategies in the

specified order at stopping points chosen by the instructor. As noted earlier, many of the RT spinoff studies follow this same pattern. In this type of instructional intervention, students learn a strategy routine regulated by the instructor, but it is unclear how much self-regulation occurs. Externally regulated routines and procedural facilitators (strategy prompts, for example) might be necessary as an introduction to strategy use, but what is missing in this literature is a critical mass of studies that seek to extend beyond the introductory phase to promote genuine self-regulation in settings and activities other than those directly overseen by the instructor.

The Collaborative Context

Students not only interact with teachers during instructional encounters but with their peers as well. Thus, the dialogic encounters students have with each other form another important component of the pedagogical context. This contextual element was assessed by examining the role of student collaboration and dialogue in the collected studies.

Both Reciprocal Teaching and Transactional Strategies Instruction are particularly conducive for promoting student collaboration and strategic dialogue. As described previously, RT was designed to provide students an entry point into literature-based discussions that foster deep and critical text processing (Palincsar, 2007; Rosenshine & Meister, 1994). TSI is also dialogue centered, designed to engage students and teachers in discussions where meaning is constructed through collaborative negotiations (Pressley et al., 1992). Unfortunately, not all of the TSI and RT studies reviewed here effectively communicate this conceptual orientation. For example, many of the RT spinoff studies

focus more intently on strategy routines than on the development of strategy dialogue. Similarly, the TSI study conducted by Loranger (1997) relies more on direct instruction of strategies than on the text-based discussions envisioned by the developers. As a result of these types of "lethal mutations" (Haertle, 1986, as cited in Palincsar, 2007), the field is left with a body of literature that fails to recognize the role of genuine student collaboration and dialogue in the development of strategic reading expertise.

The majority of studies included at least some emphasis on student collaboration, which usually took the form of structured partner work. As noted in previous sections, studies of Collaborative Strategic Reading (Klingner & Vaughn, 2000), Peer-Assisted Learning Strategies (Sáenz, Fuchs, & Fuchs, 2005), and most of the Reciprocal Teaching spinoffs (e.g., Soriano, Vidal-Abarca, & Miranda, 1996) provided students with repeated practice in the use of a sequential strategy routine. Other studies provided structured collaborative practice in the form of peer-tutoring, including a study by Van Keer and Verhaeghe (2005) in which fifth-grade students tutored other fifth-grade and second-grade students, and a study by Pickens and McNaughton (1988) in which middle school students tutored same-age peers. Student collaboration is mentioned in other studies as part of the scaffolded instructional sequence. For example, studies using the PLAN acronym as a strategy cueing system included a guided practice portion in which students completed the strategy sequence together (Radcliffe, Caverly, & Peterson, 2004).

What is less common across this body of studies are investigations of instructional frameworks that provide opportunities for student-guided collaborative activities, like book clubs or group projects (CORI is a notable exception). These would provide opportunities for students to extend their strategic reading expertise by designing their

own learning goals, strategically selecting texts to meet those goals, and drawing on social resources to clarify misunderstandings and evaluate their progress.

Textual Context

The way strategies are taught, practiced, and learned also depends on the textual environment in which the instruction occurs. In this section, I describe the types of texts middle grades students worked with in the reviewed studies. I also describe the textual environment by examining the role of independent reading in this literature.

Text types. Expository texts are the focus of more intervention studies than narrative texts, presumably because of work that has stressed the importance and prevalence of expository text for older students who are expected to "read to learn" after having already "learned to read " (Moss, 2005). In half of the studies reviewed here, students learned and practiced strategies using expository texts only. These include most of the RT studies and those using acronym procedures (TWA, PLAN, and POSSE). A few studies (about 20%) included expository and narrative texts (e.g., studies of CORI and Proctor, Dalton, and Grisham's digital strategy tutor), while others (16%) focused exclusively on narrative texts (e.g., Baumann, Hooten, & White, 1999; Loranger, 1997; Payne & Manning, 1992). A number of studies (15%) failed to provide information about the types of texts used, which could be symptomatic of a lack of emphasis in this literature on the influence textual experience has on the way students practice and internalize comprehension strategies.

Hypermedia, online, and other digital texts are underrepresented in this sample of studies. Besides the five studies of digital literacy environments, only two additional

instructional frameworks provided strategy practice with digital text. These include two studies of CORI in which the authors briefly mention that students used the Internet as a textual resource (Wigfield, Guthrie, & Perencevich, 2008; Guthrie, Van Meter, & McCann, 1996) and one generic instructional framework (Wise, Ring, & Olson, 2000) in which students practiced strategies by reading computerized text. Missing from this literature are published studies of interventions designed to teach strategies specific to hypermedia and online comprehension, although studies of this type are forthcoming (Leu, Kinzer, Coiro, & Cammack, 2004).

Basal readers and textbooks are the most common textual source, having been used in about half of the studies. These include content area textbook chapters used in studies of the POSSE procedure (Englert & Mariage, 1991), reading textbooks used in Duffy et al.'s (1986) study in which skills are recast as strategies, and passages pulled from reading books used in the original RT studies (Palincsar and Brown, 1984). Authentic texts, such as trade books, magazines, and websites, are used in only 23% of the studies, including novels and children's books used in some of the literature-based interventions (Block, 1993; Jiménez & Gámez, 1996; Kim, 2006) and the narrative and expository books on science topics used in CORI (Guthrie, Anderson, & Alao, 1999). Inauthentic texts (i.e., worksheets and passages created specifically for strategy practice in the intervention) are used in 39% of the studies. These include the degraded passages created by Rauenbusch and Bereiter (1991) to induce strategic activity from average readers, and the specially created passages in Bruce and Chan's (1991) RT spinoff study.

The role of independent reading. If the goal of strategy instruction is for students to become self-regulated comprehenders, one might expect a heavy instructional

emphasis on the transfer of strategy use from teacher-directed to student-controlled independent reading activities. Surprisingly, only about 15% of the reviewed studies directly address independent reading in their descriptions of the instructional procedures. In the remainder of the studies, there is little discussion of students selecting and reading books on their own outside of the teacher-prompted reading that occurs during direct explanation lessons and guided practice.

Although studies of RT provide students with a wealth of experience in collaborative practice using passages that are provided by the teacher, the instructional procedures do not typically emphasize the need for students to use strategies on their own in independent settings (e.g., reading a novel for a book report or a social studies text at home to review for a test). In their 1984 publication, Palincsar and Brown only mention independent strategic reading once:

> Throughout the interventions, the students were explicitly told that these activities were general strategies to help them understand better as they read, and that they should try to do something like this when they read silently. It was pointed out that being able to say in your own words what one has just read, and being able to guess what the questions will be on a test, are sure ways of testing oneself to see if one has understood." (p. 131)

Although "pointing out" that strategies can be used elsewhere is a good first step, this alone is unlikely to result in effective independent strategy use. The tendency for authors to devote very little space to independent reading in their instructional procedures (and their research reports) does not necessarily mean that students never used strategies on their own or that independent reading never occurred as part of the instructional program. It does, however, suggest that in most of the studies reviewed here, it was not prioritized as a critical attribute of the intervention.

Notable exceptions to this trend include two of the literature-based instructional frameworks discussed previously (Block, 1993; Kim, 2006) and CORI (Guthrie, Van Meter, & McCann, 1996), which is based on a model of reading that prioritizes extended exposure with text while exploring student-selected topics. Also, several studies describe the use of practices and procedures that facilitate independent reading, such as Loranger's (1997) description of reading response journals and bins of books sorted for classroom use. Still, in most of these studies, independent reading is described as an add-on—something encouraged but not explicitly modeled or assessed. For example, in a study of peer tutoring by Pickens and McNaughton (1988), students were given time to read independently during the last five sessions of the intervention without explicit instruction in how to transfer their strategy knowledge to this new setting.

Among the studies that prioritized independent reading, about half included opportunities for students to select their own texts. These include studies of CORI (Guthrie, Van Meter, & McCann, 1996), Loranger's (1997) study of TSI, and Block's (1993) study of literature-based strategy instruction. A few additional studies provided limited opportunities for students to select texts even though they did not heavily emphasize independent reading habits. For instance, in their literature-based version of strategy instruction, Baumann, Hooten, and White (1999) allowed students to choose one book (from a menu of four) to read in a book club format as the culminating event in their instructional sequence.

Summary and Implications for the Meta-Analysis

This review has demonstrated that instructional studies of multiple strategy instruction for upper elementary and middle school students cover a vast terrain that is anything but uniform. At least ten different instructional types exist, ranging from Palincsar and Brown's influential Reciprocal Teaching to other relatively obscure frameworks. The individual strategies that are included most frequently in this body of work are summarizing, predicting, generating questions, clarifying, using prior knowledge, identifying main ideas, rereading, and monitoring comprehension. Some strategies that seem particularly important for middle grades students appear infrequently in this body of literature: namely, setting a purpose for reading, monitoring one's strategy use, using social resources, and analyzing and evaluating text. The studies included in this review rarely built on students' previous strategic repertoires, and in only about two-thirds of the studies were students directly taught how to coordinate multiple strategies. Finally, self-regulation was stated as a goal in many of the studies, but few of them propelled students beyond externally regulated strategic routines.

It is difficult to concisely describe the prototypical pedagogical arrangements used in multiple strategy instruction because there is so much variability across this body of work. Some instructional frameworks took a year to complete, others only a few days. In some studies, content was taught by classroom teachers participating in professional development, and in others it was taught by university faculty. Most frameworks used a three-part instructional sequence comprised of direct explanation, modeling, and student practice, but there was substantial variability in the way student practice was emphasized. Although expository print texts were used most frequently, there were a number of

studies that included narrative and digital texts as well. Some studies relied most heavily on basal readers and textbooks, others on worksheets and passages developed specifically for targeted strategy practice, and a few on authentic texts.

A few instructional characteristics were less variable. For instance, these studies rarely emphasized the development of independent reading habits or provided opportunities for autonomous text selection. Also, when student collaboration was emphasized, it usually took the form of structured small-group activities or teacher-directed whole-class discussions.

The most important contribution of this descriptive review for the meta-analytic study is that it provides a categorical scheme for describing the variability in content and pedagogy across the MCSI literature. The top half of Figure 3 shows five areas of variability found in the instructional studies of MCSI with middle grades students. Multiple comprehension strategy instruction can vary in terms of the instructional framework being used, the specific content being taught, the way instructional content is delivered, the types of opportunities students are given to practice strategic reading, and the textual environment students encounter. A major goal of the meta-analytic review in the remaining chapters is to examine how these features moderate (i.e., strengthen or weaken) the effectiveness of strategy instruction.

This initial descriptive review also allows me to identify the constructs that are most commonly tracked as outcomes of multiple comprehension strategy instruction. The bottom half of Figure 3 lists the three constructs that are frequently tracked in this literature. A goal of the meta-analysis is to determine the impact of MCSI on each of these constructs. The bottom half of Figure 3 also shows the five major student groups

who are frequently included in studies of MCSI. A goal of the meta-analysis is to examine the effectiveness of MCSI for each of these groups to determine who benefits most.

Table 1: MCSI Intervention Studies Collected for the Descriptive Review (By Framework Type)

Reciprocal Teaching (RT)
Aarnoutse, Van den Bos & Brand-Gruwel (1998)
Brand-Gruwel, Aarnoutse & Van den Bos (1998)
Bruce & Chan (1991)
Fung, Wilkinson & Moore (2003)
Gilroy & Moore (1988)
Johnson-Glenberg (2000)
Kelly, Moore & Tuck (1994)
Klingner & Vaughn (1996)
Labercane & Battle (1987)
Lederer (2000)
LeFevre, Moore & Wilkinson (2003)
Little & Richards (2000)
Lovett, Borden & Warren-Chaplin (1996)
Lysynchuk, Pressley & Vye (1990)
Palincsar & Brown (1984)
Palincsar, Brown & Martin (1987)
Soriano, Vidal-Abarca & Miranda (1996)
Takala (2006)
Van den Bos, Brand-Gruwel & Aarnoutse (1998)

Think-Aloud Instruction
Baumann, Seifert-Kessell & Jones (1992)
Bereiter & Bird (1985)
McKeown & Gentilucci (2007)
Silven & Vauras (1992)

Concept-Oriented Reading Instruction (CORI)
Guthrie, Anderson & Alao (1999)
Guthrie, Van Meter & Hancock (1998)
Guthrie, Van Meter & McCann (1996)
Guthrie, Wigfield & Barbosa (2004)
Guthrie, Wigfield & VonSecker (2000)
Wigfield, Guthrie & Perencevich (2008)

Informed Strategies for Learning (ISL)
Paris, Cross & Lipson (1984)
Paris & Jacobs (1984)
Paris & Oka (1986)

Peer-Assisted Learning Strategies (PALS)
Calhoon (2005)
Fuchs, Fuchs & Mathes (1997)
Fuchs, Fuchs & Kazdan (1999)
Saenz, Fuchs & Fuchs (2005)

Transactional Strategies Instruction (TSI)
Anderson & Roit (1993)
Loranger (1997)

Collaborative Strategic Reading (CSR)
Klingner & Vaughn (2000)
Klingner, Vaughn & Schumm (1998)
Klingner, Vaughn & Arguelles (2004)

Acronym Procedures
Englert & Mariage (1991)
Englert, Tarrant & Mariage (1994)
Mason, L.H. (2004)
Mason, Snyder & Sukhram (2006)
Radcliffe, Caverly & Hand (2008)
Radcliffe, Caverly & Peterson (2004)
Rogevich & Perin (2008)

Non-Branded Approaches	**Digital Approaches**
Babbs (1984)	Johnson-Glenberg (2005)
Baumann, Hooten & White (1999)	Kim, Vaughn & Klingner (2006)
Block (1993)	Kramarski & Feldman (2000)
Carr & Borkowski (1989)	McNamara, O'Reilly & Best (2006)
Carriedo & Alonso-Tapia (1995)	Proctor, Dalton & Grisham (2007)
De Corte, Verschaffel & Van de Den (2001)	Salomon, Globerson & Guterman (1989)
Dole, Brown & Trathen (1996)	Sung, Chuang & Huang (2008)
Duffy, Roehler & Meloth (1986)	
Franklin, Roach & Clary (1992)	
Houtveen & Van de Grift (2007)	
Jiménez & Gámez (1996)	
Jiménez (1997)	
Kim (2006)	
Kinnunen & Vauras (1995)	
Kitano & Lewis (2007)	
Lau (2006)	
Lau & Chan (2007)	
Manset-Williamson & Nelson (2005)	
McCrudden, Perkins & Putney (2005)	
McKeown, Beck & Blake (2009)	
Miranda, Villaescusa & Vidal-Abarca (1997)	
Nolan (1991)	
Olson & Land (2007)	
Payne & Manning (1992)	
Pickens & McNaughton (1988)	
Rauenbusch & Bereiter (1991)	
Rich & Blake (1994)	
Souvignier & Mokhlesgerami (2006)	
Taylor & Frye (1992)	
Tregaskes & Daines (1989)	
Van Keer (2004)	
Van Keer & Verhaeghe (2005a)	
Van Keer & Verhaeghe (2005b)	
Walraven & Reitsma (1993)	
Wise, Ring & Olson (2000)	

For studies with multiple authors, only the first three authors are listed in this table; see References for full citations.

Table 2: Comprehension Strategies Taught to Students in Grades 4-8

Strategy	Description	Frequency	Percent
Summarizing	Identifying and synthesizing the most important information in a text, either by stopping periodically while reading or by reviewing text content after reading	62	76.5
Predicting	Making predictions regarding what will happen or be explained in an upcoming segment of text; also, evaluating and revising these predictions as new information is encountered	52	64.2
Generating Questions	Asking questions (of oneself or peers) while reading	50	61.7
Using Prior Knowledge	Activating background knowledge before reading a text; also, linking background knowledge to text content while reading	34	42.0
Clarifying Word Meaning	Identifying potentially confusing words or phrases in a text and using contextual information or other resources to figure out what they mean	29	35.8
Main Idea	Attending to the most important information in the text; distinguishing important text content from trivial content	28	34.6
Rereading	Going back to read a confusing segment of text, or reading closer to gain a more detailed understanding	18	22.2
Monitoring Comprehension	Identifying breakdowns in text-level understanding while reading	17	21.0
Graphical Organizers	Creating graphic representations of text content; e.g., story maps, concept maps, and flow charts	16	19.8
Paraphrasing/Retelling	Restating segments of text using one's own words	13	16.0
Setting a Reading Goal or Purpose	Establishing a specific goal for a reading episode; e.g., to learn new information about a topic, to enjoy a story by a favorite author, to review key concepts for an upcoming test	12	14.8
Text Structure	Identifying the structural and organizational flow of a text and using this information to guide one's reading (e.g., using physical markers like headings and captions and rhetorical markers like transitional devices and genre-specific organizational schemes)	11	13.6
Clarifying (General)	Identifying confusing parts of text and using a variety of strategies to clear them up; this designation is used for studies that do not specify whether students learn to clarify word-level or text-level confusions	10	12.3
Inferring Unstated Information	Inferring unstated information or drawing conclusions from stated information	10	12.3

Strategy	Description	Frequency	Percent
Monitoring understanding of words/phrases	Identifying unfamiliar words or phrases in text	9	11.1
Visualizing	Creating mental images of text content	8	9.9
Clarifying Understanding	Identifying confusing segments of text and using a variety of strategies to clear them up; this designation is used for studies in which students learn to clarify text-level comprehension difficulties (e.g., confusing events or concepts)	7	8.6
Monitoring Strategy Use	Evaluating the effectiveness of selected strategies for meeting specific comprehension goals; explaining or justifying why certain strategies are suitable for a particular goal	7	8.6
Previewing	Previewing or skimming a text to form a preliminary understanding of the content that will be encountered	6	7.4
Reading Ahead	Continuing reading after noticing a comprehension breakdown or inconsistency in the text to see if the confusion persists or to look for clues in subsequent content that might resolve the confusion	5	6.2
Reflecting and Relating	Using a variety of strategies to respond to text beyond surface-level comprehension, such as viewing a text from different perspectives and expressing agreement or disagreement with text content	5	6.2
Regulating Reading Speed	Modifying one's reading rate as needed depending on text demands and the reading purpose (e.g., slowing down when reading to remember details)	5	6.2
Analyzing and Evaluating	Using a variety of analytical and evaluative strategies, such as analyzing characters, critiquing the author's writing style, comparing and contrasting across texts, and critically evaluating the messages presented in text	5	6.2
Monitoring Text	Noticing inconsistencies or contradictory statements in a text	4	4.9
Taking Notes	Taking notes while reading to record important information	4	4.9
Using Social Resources	Asking peers and teachers for help when comprehension breakdowns cannot be resolved on one's own	4	4.9
Author's Purpose	Identifying the author's reason for writing a text (e.g., to inform, to persuade, or to entertain) and using this purpose to frame one's thinking	3	3.7
Drawing	Creating drawings while reading as a form of notetaking	2	2.5

Strategy	Description	Frequency	Percent
Searching for Information	Searching through a text to locate specific information	2	2.5
Using Graphic Sources	Paying attention to pictures and diagrams presented in a text to enhance understanding or to resolve comprehension breakdowns	2	2.5
Genre	Identifying the type/genre of the text (e.g., narrative or expository) and using this information to frame one's thinking	1	1.2

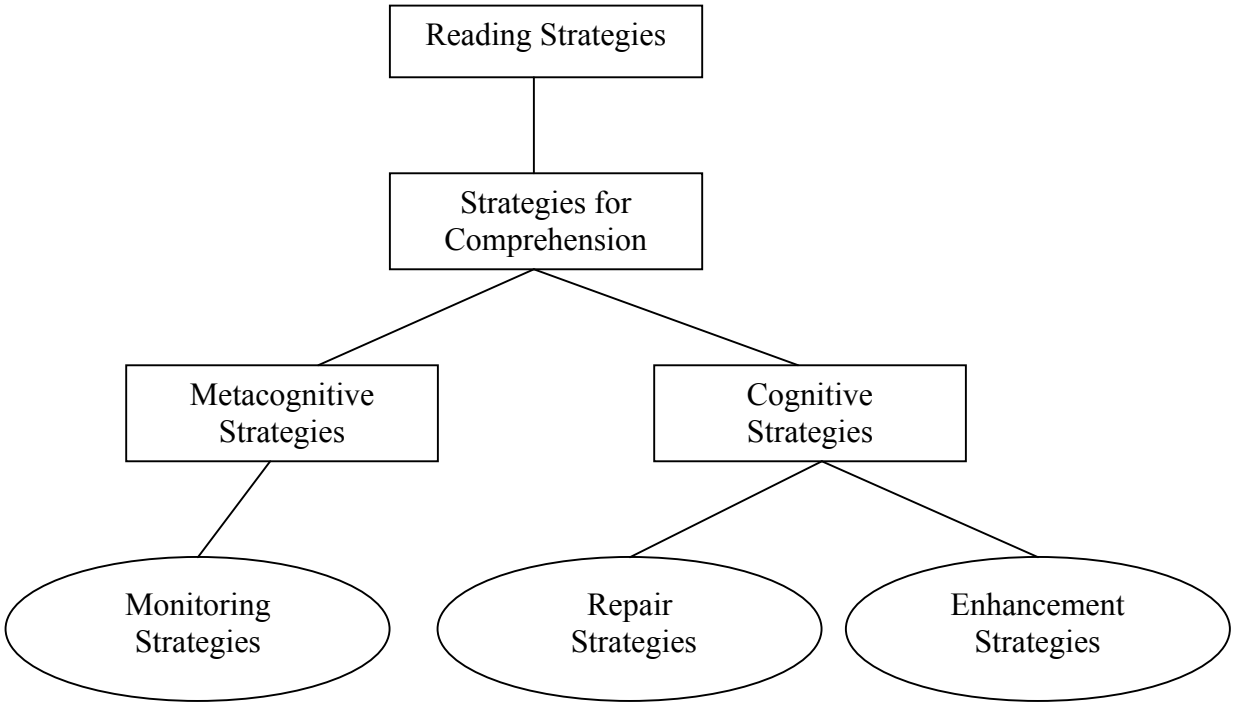

Figure 2: Diagram of Strategy Types and Purposes

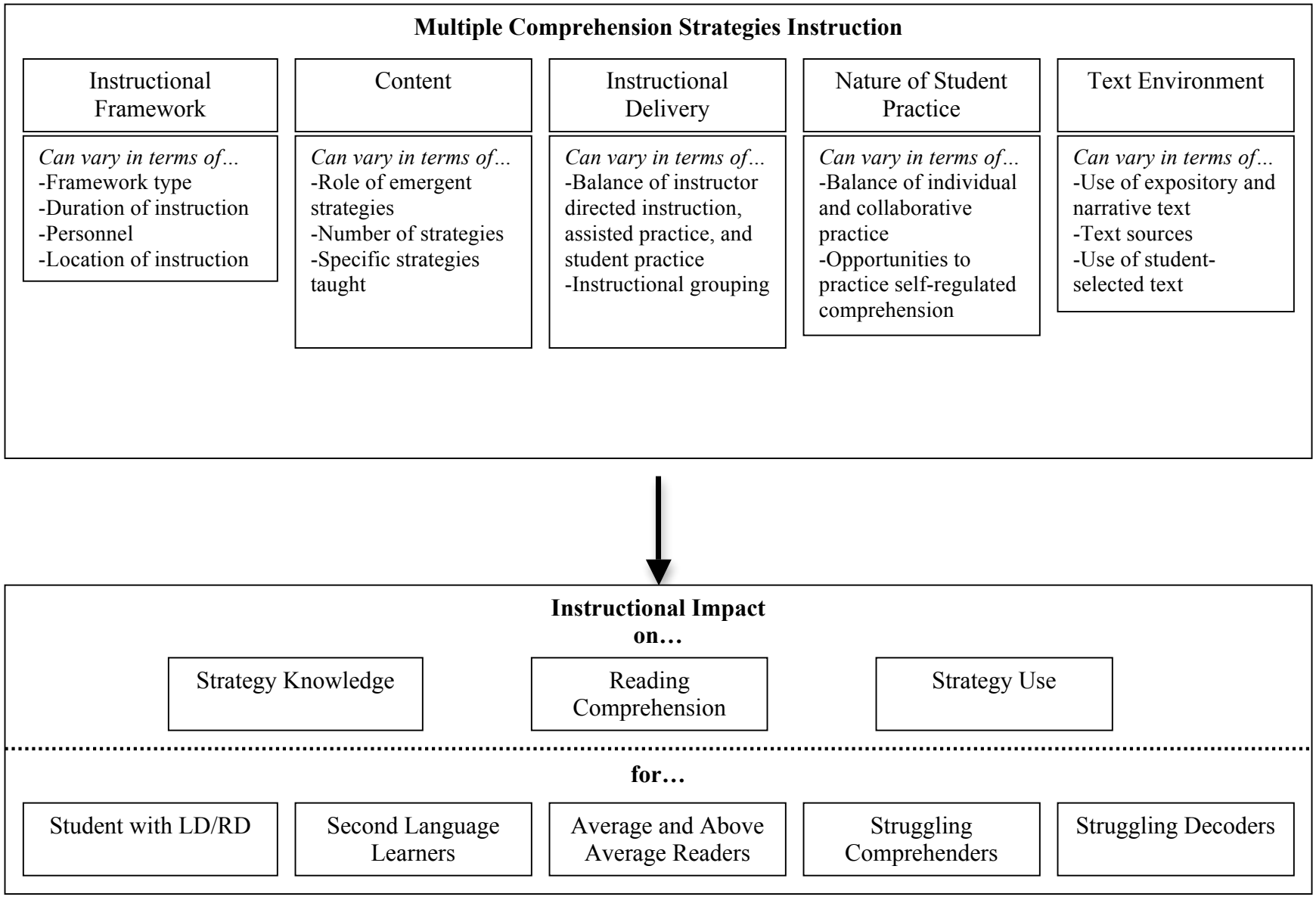

Figure 3: Categorization Scheme for Instructional Characteristics and Impact

CHAPTER III

RESEARCH METHODS

This meta-analytic synthesis addresses the following questions:

1. What is the average impact of MCSI on middle-grades student achievement?
2. Does the impact of MCSI vary across studies?
3. Of the various MCSI instructional frameworks studied with middle-grades students, which ones are most effective?
4. For which student populations is MCSI most effective?
5. Which content and pedagogical characteristics are associated with maximum impact?

Defining the Variables

A meta-analysis shows the cumulative relationship between an independent variable and a set of dependent variables across a body of studies and allows for the examination of how other variables strengthen or weaken this relationship (Lipsey & Wilson, 2001). The three sets of variables used in this analysis are defined as follows.

The Independent Variable

Multiple comprehension strategy instruction (MCSI) is an approach in which students are taught two or more strategies they can use to improve their text comprehension. These include strategies for monitoring comprehension (e.g., pausing to

summarize while reading), repairing comprehension breakdowns (e.g., re-reading and using context to figure out the meanings of unknown words), and enhancing or extending comprehension (e.g., making inferences). Numerous instructional frameworks for MCSI have been developed and studied, such as Reciprocal Teaching (Palincsar & Brown, 1984), Collaborative Strategic Reading (Klingner, Vaughn, & Schumm, 1999), and Transactional Strategies Instruction (Schuder, 1993). Examples of strategies that are commonly taught in this literature and instructional frameworks for teaching these strategies to middle-grades students are provided in Chapter 2. The meta-analysis does not include studies in which students are taught single strategies.

The Dependent Variables

The four most commonly tracked dependent variables in the MCSI literature, as documented in Chapter 2, are (1) standardized measures of reading comprehension, (2) non-standardized measures of reading comprehension, (3) strategy knowledge, and (4) strategy use. These four constructs served as the dependent variables in this analysis.

A reading comprehension measure is any assessment that indexes a student's ability to form or retain a mental representation of text content. Reading comprehension is typically measured by having students read a passage or series of passages and then respond to questions in written form. Although less common, other assessment formats are sometimes used to assess comprehension, such as cloze and maze tests, sentence verification tests, and oral and written retellings. Standardized measures of comprehension are frequently used in the MCSI literature, including the comprehension subtests of the Gates-MacGinitie Reading Test and the Stanford Achievement Test. Non-

standardized measures, typically designed by the researchers themselves or borrowed from previous studies, are also used regularly. Both assessment types were included in this study, but they were used in separate analyses because they have different measurement characteristics and have been shown in the past to differ in sensitivity to instructional interventions (Rosenshine & Meister, 1994).

The remaining two constructs are typically measured using questionnaires, observational protocols, and performance assessments. These measures are typically designed by the researchers themselves or adapted from measures used in previous studies. Standardized measures that have been normed with large, nationally representative student samples are not available for these constructs, although some of the more common measures have been shown to have acceptable reliability during field testing with convenience samples.

Measures of strategy knowledge assess students' declarative knowledge of comprehension strategies (their ability to name and/or describe relevant strategies). The Metacomprehension Strategy Index (MSI; Schmitt, 1990) and the Index of Reading Awareness (IRA; Paris & Jacobs, 1984) are questionnaires that are commonly used in the MCSI literature to measure this construct.

Measures of strategy use assess students' ability to apply their strategy knowledge while reading a text. This construct is typically measured in the MCSI literature with performance assessments that ask students to read a text and write (or state) their responses to one or more strategy prompts (e.g., provide a question that comes to mind when reading this section; write a prediction about what will happen next).

The Hypothesized Moderating Variables

Moderator analyses played a major role in this study, which was designed not only to summarize the effectiveness of MCSI but also to identify content and pedagogical factors associated with increased effectiveness. As discussed in Chapter 2, five different categories of moderator variables were hypothesized: characteristics related to (1) instructional frameworks, (2) instructional content, (3) instructional delivery, (4) the nature of student practice, and (5) the text environment. The specific questions examined in these moderator analyses are as follows:

- Is there a relationship between instructional duration and treatment effectiveness?
- How effective is MCSI when taught by classroom teachers compared to interventions delivered by researchers or computerized agents?
- Does MCSI effectiveness differ when implemented in different locations (i.e., reading classes vs. content classes; regular classroom activities vs. pull-out format)?
- Does the addition of student practice to the instructional sequence produce noticeable changes in effectiveness?
- How does instructional grouping affect the impact of MCSI?
- How does student collaboration affect the impact of MCSI?
- How does emphasis on self-regulated comprehension affect the impact of MCSI?
- How does the genre of instructional materials affect the impact of MCSI?

- Is there a relationship between the number of strategies included in the intervention and impact?
- Is there evidence that some strategies have more impact than others?

Study Collection Procedures

Criteria for Inclusion/Exclusion of Studies

A total of 101 research reports were included in this meta-analysis. Journal articles, book chapters, dissertations, conference papers, and technical reports were eligible for inclusion. Seven inclusion criteria were used to identify eligible studies.

Publication date. Only studies published/released between January 1980 and October 2009 were eligible for inclusion. The start date of 1980 was selected because prior to that, most strategy interventions included only isolated strategies.

Grade level. Studies had to include middle grades students (grades 4-8). If students in lower elementary or high school were instructed as well, the study had to either (a) report instructional procedures and outcome data in such a way that the instructional impact for middle grades students could be isolated from the impact on students in other grades, or (b) use a student sample of which at least 30% of the students were in grades 4-8.

Student characteristics. Studies were excluded that specifically targeted students with low incidence developmental disabilities (e.g., autism, visual, and hearing impairments). Studies that included students with learning disabilities (LD) and reading

disabilities (RD) were included. Studies that targeted students labeled as English Language Learners (ELLs), "at risk," or "disadvantaged" were also included.

Instructional emphasis. Only instructional studies reporting original research were included. A study was considered instructional if (a) there was an explicit intent to teach comprehension strategies and (b) at least one student outcome was tracked. These criteria exclude practitioner articles that describe or recommend instructional methods in a "how to" format, reviews that summarize research reported elsewhere, and non-instructional studies in which groups of students are surveyed or tested at one point in time for their knowledge of strategies and comprehension abilities. Also, this criterion resulted in the elimination of studies that only report teacher outcomes.

Qualifying MCSI condition. This review only included studies in which two or more comprehension strategies were taught. Single strategy studies were not eligible. As explained in Chapter 2, a common confusion in the comprehension literature is the use of the word *strategy* to refer to either the mental tools students use to bolster their comprehension or the pedagogical tools used by teachers to scaffold students' interactions with texts. This review focuses on the former. Studies that examined the impact of pedagogical strategies (e.g., teacher questioning, reader's workshop) without explicitly manipulating the strategies students were taught to use were not eligible.

Prominence of strategy instruction. Studies selected for inclusion had to explicitly foreground MCSI. This eliminated studies that mention strategy instruction as one of many instructional components of a larger program without placing special emphasis on tracking the results of strategy instruction itself.

Qualifying non-MCSI comparison condition. Eligible studies had to include a comparison group made up of students who were not taught using an eligible MCSI program. This condition could be: 1) a no treatment control, 2) a business as usual control (i.e., typical classroom practice without an explicit focus on strategies), or 3) an alternative intervention (e.g., fluency or vocabulary instruction). This criterion eliminated studies that compared two MCSI programs (e.g., technology-mediated strategy instruction versus print-based strategy instruction) if they did not include an additional non-strategy control. This also eliminated studies that only reported pre- and post-intervention data for a single MCSI group without the use of a comparison group.

Quantitative information. Finally, only studies that reported sufficient quantitative information for the calculation of numerical effect sizes were included in the meta-analytic portion of this review. To avoid minimizing the importance of seminal qualitative literature in this field, published qualitative studies were included in the descriptive synthesis portion of the review (Chapter 2).

Study Collection Strategies

Studies were located for this review using four search strategies.

1. Studies identified in previous research syntheses were collected first. The following syntheses were consulted:
    a. The National Reading Panel Report (NICHD, 2000), in particular, the studies identified as multiple strategy interventions (Section 4, pp. 61-63);
    b. Gersten, Fuchs, Williams, and Baker's (2001) review of comprehension strategy instruction for learning disabled students;

c. Block and Duffy's (2008) review of text comprehension research;
d. the Dole, Nokes, and Drits (2009) chapter on comprehension instruction;
e. the Raphael, George, Weber, and Nies (2009) chapter on comprehension instruction;
f. Mastropieri et al.'s (1996) meta-analysis of comprehension instruction for learning disabled students;
g. Swanson et al.'s (1999) meta-analysis of reading comprehension for learning disabled students;
h. Sencibaugh's (2007) meta-analysis of comprehension strategy instruction;
i. Rosenshine and Meister's (1994) meta-analysis of Reciprocal Teaching; and
j. the narrative review presented in Chapter 2.

2. Potentially eligible studies were also collected using electronic searches of ERIC, Education Abstracts FTX, PsychInfo, Social Science Citations Index, and ProQuest Digital Dissertations using the search terms strateg(y)/(ies)/(ic) AND reading AND comprehension.
3. Collected studies, including conceptually useful but ineligible studies and syntheses, were mined for additional citations.
4. Finally, a manual review of the Tables of Contents was conducted for three key journals: *Reading Research Quarterly* (the primary research journal of the International Reading Association), *Journal of Literacy Research* (the primary journal of the National Reading Conference), and *Journal of Educational*

Psychology (the primary educational research journal of the American Psychological Association).

For each search strategy, irrelevant studies were eliminated based on titles and abstracts. Then a full text version of each potentially eligible study was located and evaluated using the inclusion criteria. Table 3 shows the search yields for each step in the study collection process. A total of 507 full studies were collected for final review, and from these, the 101 that met all eligibility criteria were identified.

Treatment of Studies Reported in Multiple Sources

When a study was reported in multiple sources, the most recent and comprehensive source was used as the primary document. For example, if the same study was described in a technical report from 1999 and a journal article from 2000, the journal article was used as the primary source for coding the intervention characteristics and numerical impact data. The secondary source (in this case, the technical report) was used for coding if it reported a relevant piece of information that was excluded from the primary document. When conflicting information was presented in the two reports, information from the primary source was considered authoritative.

Reliability of Inclusion/Exclusion Decisions

A second rater was asked to make inclusion/exclusion decisions for a randomly selected sample of 12 dissertation abstracts and 12 full studies after being trained on the study selection criteria. Percentage agreement with the primary coder was 92% for the dissertation abstracts and 92% for full studies (Cohen's kappa = 0.83 for both).

Data Extraction Procedures

The coding protocol developed for this study is included in Appendix A. This protocol was used to extract information from each study in the following categories: bibliographic information (Section A of coding document), general information about the MCSI instructional framework being studied (Section B), information related to the comparison condition (Section C), content and pedagogical information for the MCSI treatment (Sections D and E), characteristics of participating students (Section F), research design characteristics (Section G), and numerical information needed to calculate effect sizes for the four eligible outcome constructs (Section H).

The coding protocol used in the meta-analysis was a revised version of the one used in the narrative review described in Chapter 2. Revisions were made to the protocol in several iterations. For each iteration, the protocol was used to code two studies. Unclear codes were identified and repaired before attempting to code additional studies. The penultimate version of the protocol was presented to a set of independent raters who helped clarify several codes to produce the final version. The coding manual developed for this study is provided in Appendix C.

All coding was completed using hard copies of the coding sheet. Then the data were entered in a PASW (formerly SPSS) data file by trained data entry personnel who regularly checked the data files for accuracy.

For tracking purposes, each study was assigned a unique report ID number. In keeping with meta-analytic conventions, a study was defined as "a set of data collected under a single research plan from a designated sample of respondents" (Lipsey & Wilson,

2001, p. 76). It is possible for multiple studies to be reported in a single research report and for a single study to be reported across several articles. When data were reported separately for relevant subgroups of students (i.e., grade level, second language learner status, and reading achievement status), these were treated as separate studies. The reports that were collected for this review provide effect size data for 127 independent study samples.

Procedures for Ensuring Reliable Coding

The researcher was the primary coder for this analysis and coded all 127 studies. Reliability of the coding procedures was assessed by comparing the primary coding with that of independent raters. Two master's-level students and one experienced teacher were recruited to serve as independent raters for the descriptive (non-statistical) sections of the coding protocol. Two doctoral students with meta-analytic experience were recruited to serve as independent raters for the portions of the coding protocol requiring statistical expertise. The independent coders were compensated for their time.

The training provided for the descriptive coding team consisted of the following:

1. An introductory session to review the general procedures of the study and each coding item (sections A-E), with examples (4 hours).
2. One session during which the raters and the trainer completed the coding protocol together for three studies (1.5 hours x 3 studies = 4.5 hours total)
3. One session during which the raters worked together to code three studies, compared their responses to those of the trainer, and then worked with the trainer to resolve any discrepancies (1.5 hours x 3 studies = 4.5 hours total)

4. Two sessions during which each rater independently coded a common study, compared their responses with each other, and then worked with the trainer to resolve any discrepancies (2 hours x 2 sessions = 4 hours total).

Upon successful completion of the 17-hour training, each rater was given 6-7 studies (n=20 in total), selected at random from the studies that had been collected to date. These studies were distributed to the coders two per week. Their ratings for these studies were compared to the researcher's ratings to calculate reliability coefficients. Percent agreement and Cohen's kappa were used as reliability indices for categorical items on the coding protocol, and bivariate correlations were used for continuous items.

The researcher met with the raters to resolve coding discrepancies as needed. The coding team also met with the researcher three additional sessions to code by committee. In total, 36 reports (32%) were coded and discussed by multiple raters.

The training provided for the doctoral-level statistical coders consisted of the following:

1. One session during which each item in sections F through H were explained, with examples (1.5 hours)
2. One session during which the rater and the trainer completed the relevant sections of the protocol together for three studies (1 hour x 3 studies = 3 hours total)
3. One session during which the raters coded three studies and then worked with the trainer to resolve any discrepancies (1 hour x 3 studies = 3 hours total)

Then the independent raters were given a randomly selected sample of studies. The responses were compared with those of the trainer to generate the reliability indices

described above. These coders also met with the researcher two additional times for joint coding.

The inter-rater reliability indices for each coding item are listed in Appendix B. Percentage agreement was moderate to high for most items, ranging from 55 to 100%. Cohen's kappa, which provides a conservative chance-adjusted estimate of inter-rater reliability, ranged from 0 to 1.0, with a mean of 0.65 across all items. For continuous items, correlations ranged from 0.41 to 1.0, with a mean of 0.94. The inter-rater reliability for items coded by the second team of coders (the numerical and methodological sections) was much higher than for those coded by the descriptive coders. The statistical coders began this study with more expertise in the area, and were able to learn the coding system with less training and support.

All items with a percent agreement below 75% or a kappa below 0.60 were re-evaluated. The researcher and coding assistants reviewed the coding definitions and examples in the coding manual to identify areas of concern. In some cases, items were eliminated from subsequent analyses because acceptable definitions and examples could not be reached in conversations with the independent coders. For most items, however, the low inter-rater reliability likely resulted from inadequate training of the additional coders; these items were retained after coding discrepancies were resolved for studies that had been coded jointly.

Procedures for Calculating Effect Sizes

The primary effect size statistic used in this analysis is the standardized mean difference (SMD) effect size. For this study, the SMD is defined as the difference

between the MCSI treatment group and a non-MCSI comparison group on an eligible outcome measure divided by the between-student pooled standard deviation (Lipsey & Wilson, 2001). This effect size statistic provides an estimate of the achievement gain attributable to an MCSI intervention expressed in standard deviation units.

When possible, both immediate and delayed SMDs were calculated. The first post-test measure recorded after the treatment ended was used to calculate the immediate post-test SMD for each study. The final post-test measure reported by the authors was used to calculate the delayed post-test SMD.

For each study in this analysis, the standardized mean difference effect size for each eligible outcome was calculated using the formula:

$$ES_{smd} = \frac{\bar{X}_{tx} - \bar{X}_{c}}{s_{p}},$$

where $\bar{X}_{tx}$ and $\bar{X}_{c}$ are the mean posttest scores for the treatment and comparison groups, respectively, and s_{p} is the pooled standard deviation across students. The pooled standard deviation is defined as the following:

$$s_{p} = \sqrt{\frac{(n_{tx}-1)s^{2}_{tx} + (n_{c}-1)s^{2}_{c}}{(n_{tx}-1)+(n_{c}-1)}},$$

where s_{tx} and s_{c} are the standard deviations for the treatment and comparison groups, respectively.

Effect sizes were calculated by applying these formulas in the PASW database for each separate outcome. An Excel effect size calculator (Wilson, 1996) and Comprehensive Meta Analysis software (Borenstein, Hedges, Higgins, & Rothstein, 2005) assisted in the calculation of effect sizes. When means and standard deviations

were not available, SMDs were calculated using procedures recommended by Lipsey and Wilson (2001) for computing effect sizes from p-values, t-values, and F-ratios.

Effect Size Corrections and Adjustments

A number of adjustments were made to the calculated effect sizes to increase their precision.

Pretest adjustments. When study authors reported post-test values that were already adjusted for pre-test differences, these values were used to calculate SMD effect sizes. When pre-test adjusted values were not reported, post-test data were adjusted for pre-test differences by subtracting the pre-test means for the treatment and comparison groups from their respective post-test means. Thus, the SMD effect size was calculated as follows:

$$ES_{smd} = \frac{(\bar{X}_{tPOST} - \bar{X}_{tPRE}) - (\bar{X}_{cPOST} - \bar{X}_{cPRE})}{s_{pPOST}},$$

where s_{pPOST} is the post-test standard deviation pooled across the treatment and comparison groups.

When sufficient pretest information was not reported to make a pre-test adjustment, the SMD was calculated using the post-test data only. Pretest adjustment was tested as a moderator variable to assess the need to control for systematic differences between ESs that were corrected for baseline differences and those that were not.

Cluster adjustments. In many of the eligible studies, intact classes or schools were assigned to condition instead of individuals. When assignment is made at the aggregate level, the student-level sample size is not appropriate for calculating the variance and weighting function because it exceeds the number of independent

observations. SMDs and standard errors from these studies were adjusted using an Excel cluster adjustment calculator (McHugh, 2004). The algorithms used by this calculator are based on calculations recommended by Hedges (2004a, 2004b, 2007) for determining cluster-adjusted SMD effect sizes, standard errors, and the effective sample size that would need to be observed to produce the adjusted standard error. The cluster adjustments have minimal effect on the SMDs themselves while attenuating the inverse variance weights used in the main effects and moderator analyses.

To make these cluster adjustments, an estimate of the proportion of variance shared by students in the same classroom or school cluster was needed, referred to as the intraclass correlation coefficient (ICC). School-level ICC estimates have been found to range between 0.20 and 0.26 for upper-elementary and middle school students on global measures of reading ability (Hedges & Hedberg, 2007). To verify that ICCs in this range were applicable in the current study for the four outcome constructs of interest, other studies of comprehension instruction that report class and school-level outcomes were consulted. Unfortunately, studies that report these data for middle school samples are relatively rare in literacy research. The studies that were located report class-level ICCs of 0.05 (De Corte, Verschaffel, & Van De Den, 2001), 0.13 (Van Keer & Verhaeghe, 2005), and 0.27 (Reis et al., 2008) for standardized measures of reading comprehension; school-level ICCs of between 0.18 and 0.23 (Torgesen et al., 2008) for standardized measures of reading comprehension; a class-level ICC of 0.10 on a strategy knowledge survey (Guthrie, Wigfield, & VonSecker, 2000); and class-level ICCs between 0.001 and 0.09 for measures of strategy use (De Corte et al., 2001).

These estimates, coupled with the Hedges and Hedberg estimates, provide the empirical basis for the ICCs used in this study. An ICC of 0.15 was assumed for classroom-level clustering and 0.20 for school-level clustering when reading comprehension was the outcome. Strategy knowledge and strategy use were expected to be less influenced by class or school membership than reading comprehension. For these outcomes, ICCs of 0.15 and 0.10 were assumed for school and classroom-level clustering, respectively. Because the empirical basis for these estimates was thin, sensitivity analyses (discussed in a later section) were conducted to examine the impact of these adjustments on the treatment estimates obtained in this review.

Small sample size corrections. The correction recommended by Hedges (1981) was applied to effect sizes from studies with small sample sizes ($n_{cluster\text{-}adj}<30$). The unbiased Hedges' g was calculated as follows:

$$g = [1 - \frac{3}{4N - 9}] * ES .$$

Procedures for Ensuring Independence of Observations for Each Analysis

An important requirement for a meta-analysis is that the effect sizes be statistically independent (Borenstein, Hedges, Higgins, & Rothstein, 2009). Dependence becomes an issue when multiple conditions are contrasted with a common comparison group or when multiple outcomes are reported for a single sample. In each case, procedures were in place to select independent observations for the present analysis.

Treatment of Studies Reporting Multiple Comparisons

Some studies in this literature compare two different types or levels of strategy instruction to a single comparison condition. Both contrasts cannot be used in a single analysis because the outcomes for the singe comparison group would appear twice in the dataset (in the effect size for each contrast). To avoid this problem, the most saturated MCSI condition (i.e., highest duration or intensity, if multiple levels were provided) was selected while avoiding conditions that included non-strategy supplements (e.g., extra collaboration or attribution training). Conditions with non-strategy supplements were selected as treatments only when the comparison condition included the same supplement without MCSI—for example, when a study compared strategy instruction plus attribution training to attribution training alone.

A similar procedure was in place for studies that compared a single MCSI condition to multiple non-MCSI comparison conditions. When this occurred, the comparison condition that most resembled typical basal reading instruction was selected for calculating effect sizes. This condition was chosen because it was the most typical comparison condition across the body of eligible studies.

The reliability of these selection rules was tested as part of the inter-rater reliability checks described previously. In a sample of randomly selected studies, exact agreement was 100% when selecting the best MCSI condition and 90% when selecting the best comparison condition.

Treatment of Studies Reporting Multiple Outcomes

Because the four outcomes included in this study were handled in separate analyses, each study could conceivably be associated with up to four unique effect size estimates without the threat of dependency. When multiple measures of a single outcome were reported in a study, a set of decision-making rules was employed to choose the most relevant outcome for inclusion in the analysis. Out of a total 227 calculated effect sizes, 30 were redundant and had to be eliminated. For measures of comprehension, the following decision-making rules were used:

1. Measures that most closely resembled a typical passage comprehension measure (read a text and answer questions) were selected.
2. Written measures were chosen over oral measures, since written measures are more typical in this sample of studies.
3. Effect sizes computed for disaggregated test items (e.g., when text-explicit questions and text-implicit questions) were combined to form a single estimate.
4. Multiple measures of the same format and genre were combined to produce a single treatment estimate (e.g., when science and social studies comprehension tests were reported separately).
5. In rare cases where multiple standardized measures were reported, the measure that occurred most commonly in this dataset was selected (i.e., Gates-MacGinitie comprehension subtest or other common test).

When multiple measures of strategy knowledge were reported, the measure that was most like the Metacomprehension Strategy Index or Index of Reading Awareness was selected, since these are the most typical questionnaires for this construct in this set

of studies. For strategy use, it was common for researchers to report the results for each strategy separately. When this occurred, the effect sizes were aggregated for each independent sample.

Data Analysis Procedures

A database was created that contains the descriptive and numerical data extracted from each eligible primary study. This database was used for the three layers of analysis conducted in this study, as described in the sections that follow.

Layer 1: Summarizing the Impact of MCSI Through Analyses of Main Effects

Identifying and treating outliers. Before the main effects analysis was conducted, the distributions of effect sizes and sample sizes were examined to identify potential outliers. An outer fence was calculated for each distribution using the procedure described by Tukey (1977). ESs and effective sample sizes that lay outside the outer fence were checked for coding and data entry errors and then "Winsorized" to fall just inside the maximum acceptable value (Lipsey & Wilson, 2001). This procedure preserves the relative order of effect sizes while preventing exceptionally high values from skewing the mean.

Main effects analysis. In this phase of the analysis, the average effect of strategy instruction on each of the targeted outcomes was calculated using a weighted averaging procedure and a random effects statistical model (described below). These calculations were primarily conducted using a mean effect size macro written for SPSS by Wilson (2005). These weighted values provide a summary of the effectiveness of strategy

instruction across the body of reviewed studies. The average standardized mean difference effect size (adjusted for pretest differences, clustering, and small sample size bias) was calculated for:

1) standardized measures of reading comprehension;
2) non-standardized measures of reading comprehension;
3) non-standardized measures of strategy knowledge;
4) non-standardized measures of strategy use.

Summary effects were also calculated for delayed measures of reading comprehension using these same procedures, although delayed data were infrequently reported in the collected studies.

Weighted averaging procedure. As is typical in meta-analytic syntheses, a weighting procedure was used to increase the precision of the main effects estimates. Instead of calculating main effects using raw ES data, each individual ES statistic was weighted by its inverse variance weight (defined below). This procedure allows larger studies (i.e., those likely to produce more precise treatment effects) to carry more weight in the main effect estimate (Lipsey & Wilson, 2001).

Random effects statistical model. The analyses used a random effects statistical model because significant variability in effectiveness across studies was expected (Borenstein, Hedges, Higgins, & Rothstein, 2009). Under the random effects model, the standard error and inverse variance weights for an ES statistic are calculated using these formulas:

$$SE = \sqrt{v_i + v_0} \quad \text{and} \quad w = \frac{1}{v_i + v_0},$$

where υ_i is the within-studies variance component for each study and υ_0 is the between-studies variance component (a constant). The value υ_i is calculated as the standard error of the effect size statistic under the fixed effects model. For SMD effect sizes, this is calculated as follows:

$$SE_{smd,fixed} = \sqrt{\frac{n_{tx} + n_c}{n_{tx} n_c} + \frac{ES^2_{smd}}{2(n_{tx} + n_c)}} \quad ,$$

where n_{tx} and n_c are the sample sizes for the treatment and comparison groups, respectively. The value υ_0 was calculated using the noniterative method of moments within the SPSS mean effect size macro (Wilson, 2005).

Homogeneity analysis. To determine if treatment effects vary across studies, a test of homogeneity was conducted for each of the four weighted mean effect sizes. This test helps determine if there is more variability across effect sizes than expected from student-level sampling error. As described by Lipsey and Wilson (2001), this test of homogeneity is conducted using the Q-statistic, calculated as:

$$Q = \sum_{i=1}^{k} w_i (ES_i - \overline{ES})^2 ,$$

where w_i is the inverse variance weight for the i^{th} effect size, ES_i is the i^{th} effect size, and $\overline{ES}$ is the weighted mean of *k* effect sizes. A mean ES with a Q-statistic that exceeds the critical chi-square value (with df = k - 1) is assumed to be heterogeneous across studies (or more accurately, not homogeneous).

In addition to the Q-statistic, the I^2 value for each mean effect was examined, which is itself a function of Q. This statistic estimates the percentage of total observed variance (between and within studies) that is attributable to between-study treatment

differences. A large I^2 value indicates the need for moderator analyses to explain these between-study differences (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Sensitivity analyses. Analyses were conducted to examine the robustness of the treatment estimates to bias related to coding decisions. First, main effects analyses were repeated using standard errors and weights that were not adjusted for clustering to assess the impact of the cluster adjustments on the treatment estimates. Second, because some treatment estimates (about 13%) had to be eliminated to avoid statistical dependencies in the dataset, main effects analyses were conducted using ESs that were removed (in place of those that were retained) to examine the impact of these elimination decisions. Additional details about these sensitivity analyses and the findings are described in the Results.

Analysis of publication bias. Publication bias is the tendency for effect sizes reported in published studies to be larger than those in the population of effect sizes produced by all eligible studies (Lipsey & Wilson, 2001). Larger studies are more likely to produce statistically significant results for a given effect size and are therefore more likely to be made public by researchers. This makes it likely that systematic reviews drawing from publicly available work report treatment effects that are unrepresentative of the full universe of public and non-public data (Rothstein, Sutton, & Borenstein, 2005).

This problem was partially dealt with in the current study by allowing the inclusion of unpublished dissertations, conference papers, and technical reports. Even with the inclusion of these studies, the possibility of publication bias does not disappear. Several analytical procedures were employed to assess whether the data are consistent with the pattern one would expect from publication bias.

First, funnel plots showing the relationship between ES and standard error for each outcome were created using CMA software. Visual inspection of these plots can provide information about the possibility of publication bias. Larger studies are expected to produce treatment estimates near the mean, while smaller studies are expected to produce a wider range of treatment estimates that spread across the lower band of the funnel plot. If the smaller studies tend to cluster on the right-hand side of the plot, this alerts the researcher to the possibility that a set of small studies with negative or null results would be needed to produce a symmetrical distribution (Borenstein, 2005). The trim and fill method, also available in the CMA software, was used to adjust the observed effects for possible asymmetry due to publication bias (Duval, 2005). Finally, to provide an empirical basis for interpreting the funnel plot, rank correlation and regression intercept tests were conducted (Sterne & Edgar, 2005). The findings of these analyses are detailed in the Results chapter.

Layer 2: Characterizing the MCSI Literature and Identifying Useable Moderators

Response distributions for each item in the coding protocol were examined to identify the variables that could be used as moderators in layer 3. Several instructional characteristics were eliminated as moderators because they occurred in too many or too few studies. For example, one moderator that was originally hypothesized to positively influence treatment effectiveness was the use of a balance of teacher-directed instruction, assisted practice, and student practice. Virtually every study included some type of instructor directed instruction and assisted practice in their treatment specification, so the presence of these phases was not a meaningful analytical category. However, the

presence of student practice was retained as a moderating variable because there were a substantial number of studies that did not include this component.

Continuous variables with highly skewed distributions were converted to ordinal variables. For example, duration of instruction, which was highly skewed to the left, was recoded into a three-point scale (low, moderate, and lengthy). Categorical variables were dummy coded so they could be included as moderators in the weighted regression analyses.

Correlations among moderating variables were also examined. Four coding items (monitoring strategy effectiveness, setting goals for reading during student practice, monitoring comprehension during student practice, and selecting strategies flexibly) were found to be positively correlated and were combined to form a composite variable (Cronbach's alpha = 0.62). When these variables were entered in a factor analysis, a single component was identified with an eigenvalue of 1.895, providing additional evidence that these items could be aggregated in a single scale. The items were summed to create a score (ranging from 0-4) that indexes each study's emphasis on self-regulated comprehension.

Correlations among the individual strategy variables (coded in item 18 in Appendix A) were also examined to identify strategies that tended to be taught together. Because of the prevalence of Reciprocal Teaching in the MCSI literature, the four RT strategies (summarizing, predicting, clarifying, and questioning) were frequently found together. Among studies that did not include these four strategies, no discernible categories were evident. Therefore, studies were coded as belonging to one of two categories: those that included the RT strategies ("big four") and those that did not. A

two-step cluster analysis (with the presence or absence of each strategy entered as a categorical variable) produced two main clusters of studies that agreed 86% with the original categorization scheme. Big four category membership was used in the moderator analyses to determine if these four strategies produced a systematic treatment benefit over studies that included other strategies.

In addition to these empirically derived categories, an attempt was made to combine individual strategies into conceptually similar categories. Studies were coded for the presence or absence of at least one strategy in the following clusters: strategies for (1) using prior knowledge (making inferences and activating prior knowledge); (2) notetaking (drawing and taking notes); (3) identifying and restating important information (paraphrasing, summarizing, and finding main ideas); (4) anticipating text content (predicting, previewing, identifying genre and author's purpose, specifying a purpose for reading); (5) monitoring and repairing comprehension (comprehension monitoring, clarifying, using general fix-up strategies); (6) seeking external assistance (using social resources, using other resources like dictionaries, etc.); (7) general comprehension enhancement (graphic organizers, questioning, and visualizing); and (8) textual analysis (analyzing, evaluating, and reflecting). As moderators, these clusters did not produce any meaningful findings and may have masked some of the impact of individual strategies. Therefore, this scheme was abandoned in favor of testing the impact of the most common strategies (those included in 10% or more of studies) individually.

Sample characteristics, which were originally coded as percentages, were recoded as dichotomous categorical variables because the percentage distributions were highly asymmetric. In many cases, these variables produced bimodal distributions because most

studies that reported sample information either included many students with a particular characteristic or very few. Coded percentages were used to create a dichotomous variable for each characteristic that indicated whether or not most of the sample (75% or more) had that characteristic. Dichotomous variables were created for the following characteristics hypothesized to be related to treatment effectiveness: percentage of students at each grade level, percentage of second language learners (SLL), percentage of students diagnosed with learning or reading disabilities (LD/RD), percentage of struggling readers, percentage of average or above average readers, and percentage of students who come from low socioeconomic status (SES) backgrounds.

Unfortunately, detailed descriptions of student samples are rarely reported in the MCSI literature. This produced a large amount of missing data for many of these categories. When possible, sample characteristics were estimated from school or district characteristics, which were more commonly reported. Some sample characteristics that were coded were not included in the analyses because they were reported too infrequently (e.g., gender composition and percentage of struggling decoders and struggling comprehenders).

In addition to providing a basis for the moderator analyses in layer 3, the results of the layer 2 analyses were compared to the findings presented in the narrative review (Chapter 2) to help characterize the MCSI literature.

Layer 3: Using Moderator Analyses to Identify the Critical Attributes of Effective MCSI

Because the homogeneity analyses indicated that substantial variability in treatment effects exists between studies for all four outcomes, moderator analyses were

warranted to explain this variability. The specific questions addressed in these analyses are shown in Table 4. Moderator analyses were not conducted for delayed measures because these were only reported in a handful of studies.

Moderator analyses were conducted in the following stages for each of the four outcome measures.

Exploratory breakouts by framework and sample characteristics. To compare the impact of various MCSI frameworks, the weighted mean ES was calculated for each of the treatment types. Because there were only a few studies in most of the categories, a formal comparison across framework types was not possible. However, these breakouts provide useful information about the potential for each of these frameworks to impact student performance. Summary effects were also calculated for various student groups to examine the differential impact of MCSI by sample characteristics. Because of limited data, it was not possible to formally compare across these groups.

Examining the relationship between method and impact. This meta-analysis was inclusive of studies of varying methodological quality, ranging from carefully designed experimental manipulations to less tightly controlled quasi-experiments. Rather than excluding studies with known methodological flaws, design characteristics were tested as moderators to examine the possibility that effect size estimates would vary depending on the methodological characteristics of the studies. Two categories of method variables were used.

The first group of method variables was designated as indicators of study quality. These included: type of assignment to condition (random versus nonrandom), use of measures to track implementation fidelity, use of a comparison condition that only differs

on one dimension (content similarity of treatment and control), demonstrating that groups were equivalent at baseline on relevant measures, and providing the necessary information to adjust post-test data for potential differences at pre-test. Percentage of attrition was also desired as a quality indicator but could not be used because of inadequate reporting in this literature. These variables were used to identify differences in treatment effects for higher and lower quality experimental/quasi-experimental designs and to control for these differences in the moderator analyses.

Another group of method variables was also tested that included: assignment level (student vs. group), use of a multi-school sampling procedure, and type of comparison condition. These variables were used to examine additional methodological characteristics that are not necessarily indicators of quality but could be related to effect size.

Zero-order correlations were computed between these variables and treatment effectiveness for each of the four outcomes. These correlations were obtained by entering each method variable individually as the sole predictor of ES in a random effects weighted regression analysis (using the SPSS metaregression macro described previously). Method variables were identified that produced standardized regression coefficients approaching 0.20 or higher with a p-value around 0.30 or lower. These liberal criteria for selection (i.e., "breaking" the alpha = 0.05 convention) were warranted given the modest sample sizes of these analyses and the need to pinpoint any factor associated with treatment that would need to be controlled for in subsequent analyses. The method variables that passed the first screening were then entered simultaneously in a weighted regression analysis, again with effect size as the outcome variable. The

resulting regression models were used to examine the relationship between method and treatment effectiveness. The variables in these models were identified on an exploratory rather than confirmatory basis, so the relationships that are found should not be generalized beyond this set of reviewed studies.

Building the control models. A set of control variables was identified for each outcome so the hypothesized content and pedagogical characteristics could be tested after accounting for non-substantive factors associated with effect size estimates. Research method characteristics (described in the previous section), publication status (journal article vs. unpublished report/dissertation), measurement characteristics (e.g., test genre and format), and sample characteristics were candidates for inclusion in the control models.

Control variables were identified using the same exploratory screening process described above. Zero-order correlations were computed between these variables and treatment effectiveness for each of the four outcomes. Any variable with a marginally sizeable standardized regression coefficient (0.20) that was marginally statistically significant ($p<0.30$) was selected as a potential control variable. The control variables that passed the first screening were then entered simultaneously in a weighted regression analysis. Variables that continued to have sizeable associations with effect size (using the same liberal criteria above) were retained as covariates in subsequent analyses to account for differences in effectiveness attributable to factors other than the content and pedagogical characteristics of interest. A separate control model was developed for each of the four outcomes because the control variables correlated differently with each outcome.

Testing the hypothesized moderator variables. This portion of the moderator analyses was conducted on a confirmatory basis in that the moderating variables were hypothesized in advance (see research questions in Table 4). Each hypothesized moderator variable was added individually to the control model to examine their degree of association with effect size after accounting for the influence of the selected control variables. The moderators with sizeable positive or negative associations with treatment effectiveness were added to the regression model simultaneously to examine the unique influence of each one. These models were used to identify instructional characteristics that potentially promote or impede the impact of MCSI on middle grades student achievement.

Table 3: Primary Search Sources and Yields

Source	Number of Titles Located	Number of Articles Collected for Additional Review
National Reading Panel Report (section entitled *Multiple Strategies Instruction*)	40	38[1]
Education Abstracts Electronic Database (*strateg** AND *reading* AND *comprehension* in abstract)	482	130
Social Sciences Citation Index Database (*strateg** AND *reading* AND *comprehension* in topic; published in English; articles and reviews only)	811	147
PsychInfo Electronic Database (*strateg** AND *reading* AND *comprehension* in abstract; published in English; all journals)	765	85
ERIC Electronic Database ("reading strategies AND "reading comprehension" as descriptors; all ages except adults and postsecondary students)	493	32
Other previous reviews	19	7
ProQuest Dissertation Abstracts (*strateg** AND *reading* AND *comprehension* in abstract)	1,611 (122 abstracts selected for review)	65
Mining and Manual Searches	9	3
Totals	4,230	507

[1]Excludes 2 articles that were incorrectly cited and could not be found

Table 4: Summary of Planned Moderator Analyses

Construct	Question(s)
Student characteristics	Does the effectiveness of MCSI vary for different student populations?
Instructional framework type	Are some of the most popular instructional frameworks for teaching MCSI more effective than others?
Instructional duration	Is there a systematic relationship between instructional duration and effectiveness? How much MCSI is needed to produce a substantial effect, and at what point does added instruction stop making a difference?
Instructional personnel	How effective is MCSI when taught by classroom teachers, and how does this compare to the impact of MCSI taught by researchers and delivered by computer?
Location of instruction	What is the average effect of MCSI when conducted in classroom settings (reading and content classes), and how does this compare to instruction in non-classroom settings (e.g., pull-out or enrichment programs)?
Number of strategies	Is there a systematic relationship between the number of instructed strategies and instructional impact? Is there an upper limit on the number of strategies that should be taught?
Student practice	Are interventions that include student practice (without direct teacher or computer prompts) more effective than those that do not?
Instructional grouping	Is MCSI more effective when delivered to large groups, small groups, or individually?
Individual and collaborative strategy practice	Are interventions that include collaborative and individual strategy practice more effective than those that only include one or the other?
Opportunities to practice self-regulated comprehension	Are opportunities that provide extensive opportunities for students to practice self-regulated comprehension more effective than those that do not?
Text type	Are interventions that focus on both narrative and expository texts more effective than those that focus on only one text type?

CHAPTER IV

RESULTS

In this chapter, the numerical results from this study are presented. The implications of these results for practice and research are discussed in Chapter 5.

A total of 197 effect sizes were computed for this analysis. The total includes 136 treatment effects for reading comprehension (60 for standardized and 76 for non-standardized measures), 28 for strategy knowledge, and 33 for strategy use. The standardized mean difference effect sizes for each study are presented in Table 5. A visual depiction of these estimates is provided in the forest plots (Figures 4 - 7). Many of the point estimates for individual studies failed to reach statistical significance as shown in the forest plots by confidence bars that cross zero.

Layer 1: Main Effects Analyses

Although most of the individual point estimates fail to reach statistical significance, all main effects were positive and statistically significant at the 0.01 level. The results of the main effects and homogeneity analyses are presented in Table 6. For standardized measures of reading comprehension, the weighted mean effect size was 0.36, with a confidence interval of 0.21 to 0.51. For non-standardized measures, the mean was 0.55, with a confidence interval of 0.36 to 0.73. The summary effects for strategy knowledge and strategy use were 0.73 (CI=0.53 – 0.92) and 0.77 (CI=0.54 – 0.99), respectively.

Main effects were also calculated for studies that reported delayed post-test information. Delayed post-test administration ranged from 5 to 60 days after treatment ended. Only three effect sizes could be calculated for delayed effect of MCSI on standardized measures of reading comprehension; together these produced a weighted mean ES of 0.27 (SE = 0.16; 95% CI = -0.04 – 0.58). Thirteen effect sizes were available for non-standardized measures of comprehension, with a mean of 1.06 (SE=0.21; 95% CI = 0.65 – 1.46). Delayed effect sizes were not computed for the other two outcomes because delayed outcome data for these were rarely reported. No systematic relationship was observed between length of delay (in days) and size of delayed treatment effects.

Homogeneity Analyses

Tests of homogeneity suggest that treatment effectiveness varies significantly across studies for each of the four outcomes (Q = 127.08 for standardized comprehension, 233.58 for non-standardized comprehension, 55.57 for strategy knowledge, and 74.49 for strategy use; $p<0.01$ in all cases). For standardized comprehension outcomes, nearly 54% of the total variability was between studies. For non-standardized measures, this percentage was around 68%. For strategy knowledge and strategy use, percentages of between-study variability were 51% and 60%, respectively. These large I^2 values provided an empirical justification for the moderator analyses described in a later section.

Sensitivity Analyses

As explained in the Method section, some treatment estimates (about 13%) had to be eliminated to avoid statistical dependencies in the dataset. To assess the impact of these deletions, main effects analyses were repeated using the ESs that were removed in place of the ones chosen for retention. For standardized measures of comprehension, the weighted mean effect size was 0.39 (SE = 0.08; 95% CI = 0.23 – 0.54) compared to 0.36 (SE = 0.08; 95% CI = 0.21 – 0.51) reported above. For non-standardized measures, it was 0.57 (SE = 0.10; 95% CI = 0.37 – 0.78) compared to 0.55 (SE = 0.09; 95% CI = 0.36 – 0.73). The revised point estimates and confidence intervals overlap almost completely with the results obtained with the original elimination decisions, indicating that these decisions do not introduce bias. This sensitivity analysis was not conducted for the remaining two outcomes (strategy knowledge and use) because only two non-independent ESs had to be eliminated for these outcomes.

A second set of sensitivity analyses was carried out to assess the impact of the cluster adjustments on the treatment estimates obtained in this study. The mean weighted effect size for each outcome was recalculated using raw sample sizes and weights (unadjusted for classroom or school-level clustering). These sensitivity analyses produced point estimates and confidence intervals that did not vary substantially from the cluster-adjusted estimates (with a random effects model and pre-test and small sample size adjustments applied). Without cluster adjustments, the main effect of MCSI on reading comprehension was 0.42 (SE = 0.09; 95% CI = 0.24 – 0.59) for standardized measures and 0.53 (SE = 0.09; 95% CI = 0.36 – 0.70) for non-standardized measures. Recall that with cluster adjustments, the treatment estimates were found to be 0.36 (SE = 0.08; 95%

CI = 0.21 – 0.51) and 0.55 (SE = 0.09; 95% CI = 0.36 – 0.73). The summary effect without cluster adjustments was 0.73 (SE = 0.09; 95% CI = 0.55 – 0.92) for strategy knowledge and 0.72 (SE = 0.10; 95% CI = 0.54 – 0.91) for strategy use, compared to the adjusted values of 0.73 (SE = 0.10; 95% CI = 0.53 – 0.92) and 0.77 (SE = 0.54 – 0.99), respectively.

It appears that the cluster adjustments had minimal impact on the treatment estimates obtained in this study. To ensure a conservative analysis, the cluster-adjusted weights were retained. The similarity of standard errors and confidence intervals for the adjusted and unadjusted results is unexpected. The cluster adjustments decrease the sample size for individual studies, which should increase the standard errors and widen the confidence intervals. This does not appear to be the case in this dataset, suggesting that the random effects variance component is so large that it overwhelms the within-study variance.

Assessing the Risk of Publication Bias

Several analyses were used to assess the possibility that treatment effects obtained in this analysis were inflated due to publication bias. The funnel plots depicted in Figure 8 appear relatively symmetrical around the treatment means, which suggests that the threat of publication bias in these analyses is fairly low. Results of the trim and fill procedure confirm the visual interpretation of the funnel plots. For standardized comprehension, strategy knowledge, and strategy use, no studies were needed to improve the symmetry of the plot. For non-standardized measures of comprehension, two additional studies were needed on the left side of the plot to improve the symmetry. The

revised treatment estimate with the imputed studies added was virtually identical to the original treatment estimate of 0.55. Again, this suggests that the possibility of an inflated main effect due to publication bias is low.

It should be noted, however, that the results of the rank-correlation and regression intercept analyses provide some evidence of funnel plot asymmetry. These techniques provide more sensitive tests of publication bias and can reveal some tendency for bias that is not apparent using the trim and fill procedure. For all four outcomes, the rank correlations (Kendall's tau) between ES and variance were positive and statistically significant (Table 7). This suggests that small studies in this literature tend to be associated with higher treatment effects, a pattern consistent with publication bias. For three of the outcomes, positive and statistically significant intercepts were obtained when ES was regressed on study precision (a function of sample size). This also indicates the presence of asymmetry consistent with publication bias.

Publication bias is one explanation for the observed relationship between sample size and treatment effects. Another possibility is that smaller studies tend to have certain substantive characteristics that make them more effective (Borenstein, 2005). To test this possibility, zero-order correlations were examined between sample size (effective N) and the instructional characteristics that were found to be associated with ES. The one characteristic found to correlate with effective sample size was instructional duration, with smaller studies tending to be longer in duration than larger studies ($r = -0.241$, $p = 0.036$). As described later, duration was found to have a small positive relationship with effect size for non-standardized comprehension measures but was not associated with effect size for the other measures. Therefore, it is possible that the small amount of

asymmetry observed in the funnel plot for non-standardized comprehension outcomes is the result of something other than publication bias.

Cumulatively, these analyses indicate that if publication bias is a threat in this study, it is not severe enough to invalidate the results. At worst, the treatment estimates reported here are slightly inflated due to publication bias, but they are not likely to be zero.

Layer 2: Characterizing the MCSI Literature Using the Collected Studies

This portion of the meta-analytic synthesis is meant to confirm and extend the findings presented in the narrative review in Chapter 2. While Chapter 2 describes a wide body of published studies of MCSI (quantitative and qualitative), this part of the analysis includes published and unpublished intervention studies for which numerical effect sizes could be computed. Descriptive statistics for the collected studies are presented in Tables 8 and 9.

Methodological Characteristics

The studies collected for this analysis vary greatly in terms of study design and quality. Most of the collected reports were found in journal articles (58%). A large number of eligible dissertations were also available, making up about 39% of the eligible reports. Other reporting formats were also eligible, including conference papers, book chapters, and technical reports. However, fewer of these were located that met all the eligibility criteria. For many of those that were located, newer and more complete

versions were reported in dissertations or journal articles. Only two non-redundant conference reports were included here.

Nearly all studies in this literature use a sampling scheme based on convenience. Schools or teachers are usually recruited based on availability and willingness to participate. A systematic selection process to produce samples that are representative of a target population is rarely seen. Therefore, sampling was not included as a methodological quality variable. Type of assignment, on the other hand, does vary in this literature. About 40% of the independent samples in the present study were randomly assigned to condition. Simple non-random assignment with no effort to equate groups was used about 40% of the time, while non-random assignment with blocking or matching was used to assign about 18% of the time. Level of assignment also varied across studies in this literature, with only about 30% of the reports assigning at the student level. In the remaining reports, intact groups (e.g., pre-existing reading groups), classes, or schools were assigned to condition.

Approximately half of the samples in this analysis were drawn from multiple schools. Effect sizes from these studies are standardized on a somewhat different variance (student and school-level) that might make these effect sizes different from those reported in single-school studies (student-level variance only). Ideally, the effect sizes across studies using different sampling levels would be standardized on a common variance structure. A crude way to handle this situation is to test the relationship between this characteristic and effect size estimates to determine if these studies produce differential results (M. Lipsey, personal communication). Therefore, the sample level (multiple vs. single school) was considered as a potential control variable in this analysis.

An effort was made in this analysis to determine how closely aligned the treatment and comparison groups were on relevant measures prior to the beginning of the intervention. Using the information presented in each study, baseline equivalence was determined for reading comprehension, strategy knowledge, strategy use, other measures that were reported (e.g., IQ, self-efficacy, and academic performance in other subjects), and demographic characteristics (gender, ethnicity, and socioeconomic status). As shown in Table 8, equivalence could not be deemed on any of these measures for about 25% of these studies, usually due to lack of pretest reporting. About 36% of the samples were equivalent on one of these measures, 25% on two measures, 33% on three, and 7% on four.

An attempt was also made to determine the percentage of each sample that left the study between assignment and post-test administration. Because sample size varied slightly for different post-test measures, attrition was calculated separately for each effect size in the dataset. For about 29% of the computed effect sizes, zero attrition was reported. It should be noted, however, that attrition was impossible to determine for a large number of effect size estimates due to inadequate reporting – it was common for researchers to report the number of individuals who completed all post-tests without reporting the number who were originally assigned to each condition. This also made it difficult to distinguish between studies with zero attrition and those with unreported attrition. As a result of these difficulties, it was not tenable to test attrition as a potential control variable.

Studies in this literature use a variety of comparison conditions, which can greatly impact the size and interpretation of the standardized mean difference effect sizes

computed for this study. In a few cases (11%), no treatment controls were compared to MCSI. More common, however, were business as usual conditions (e.g., typical classroom practice) or alternative interventions. The similarity of content between treatment and control conditions also varied somewhat, with about 30% of studies comparing MCSI to a condition that varied on characteristics other than the multiple strategy focus. For instance, some lower quality studies compared MCSI to a social studies or science condition that did not explicitly focus on reading comprehension.

The final methodological characteristic examined was the presence of systematic procedures for measuring implementation fidelity. Originally, an attempt was made to evaluate the extent to which each intervention was implemented as intended, but this was not possible because researchers rarely provided enough information about actual implementation to make a reliable assessment. Instead, a dichotomous variable was created that indicated whether or not researchers paid explicit attention to fidelity. This decision was based on the assumption that there may be a difference in implementation quality between studies conducted by researchers who recognize the need for fidelity measurement and those conducted by researchers who ignore this important issue. Fidelity measures were in place in about 34% of the research reports.

Sample Characteristics

The studies included in this analysis report outcomes for a total of 10,765 students. Study size ranged from very small to large (6 to more than 1500), with a median study size of 45 students. Many of the samples included in the present analysis were drawn from multiple grades (about 40%), and for many of these, the exact distribution of

students across grades was impossible to estimate. About 20% of the samples were predominantly made up of 4th grade students, and about 13% were mostly made up of 5th graders with another 13% comprised mostly of 6th grade students. Samples that included mostly 7th or 8th grade students were a little more rare (about 6% for each).

About 14% of the study samples were made up predominantly of students with LD/RD; about 8% mostly included second language learners; and about 44% mostly included struggling readers (i.e., students below grade-level expectations, as identified by the researchers). As explained in the Methods, additional sample characteristics were desired but could not be examined due to poor reporting in this literature.

General Characteristics of the Instructional Framework

Table 8 shows the breakdown of collected studies by instructional framework. The most represented name-brand framework was Reciprocal Teaching, which was tested in over 20% of the reports. Non-branded frameworks were also very common, making up around 45% of the eligible studies. Digital versions of MCSI were studied in around 6% of the collected reports. Few qualifying studies of other name-brand and acronym frameworks were located for this review.

The versions of MCSI included in the study varied in terms of specificity. Some versions were presented to teachers as a general approach to providing strategy instruction without a specific curriculum (about 22%), while others were presented in curricularized form, either general (33%) or highly specified (46%). Instructional duration also varied, ranging from 100 minutes of instruction to year-long interventions.

In about half of the studies, instruction was provided by the researcher or members of the research team, while 42% of the interventions were provided by classroom teachers. Although computerized strategy instruction is becoming more common, only about 5% of the identified interventions were delivered mainly by computerized agents. Similarly, while the hybridization of comprehension and content instruction is becoming more common, the majority of reports identified for this review describe interventions that took place in reading classes and other non-content focused settings (85%). Finally, the interventions were more likely to be conducted during regular classroom activities than in supplemental pull-out programs or out-of-school settings.

Instructional Content

The 26 strategies included in the MCSI programs reviewed in this study are listed in Table 8, along with their frequencies of inclusion. As in the narrative review in Chapter 2, the most commonly taught strategies were summarizing, generating questions, predicting, and clarifying. Nearly 40% of the interventions included all four of these strategies in their instructional repertoires. Other commonly taught strategies included identifying main ideas, monitoring comprehension, and using prior knowledge. In over 90% of the reports, all strategies were pre-planned before instruction began; because of the lack of variability for this characteristic, the possibility of emergent strategies was not included as a moderating variable in this study.

Instructional Delivery

In addition to the content of MCSI, this study was concerned with the way the content was delivered to students. Instructional delivery was operationalized as the size of instructional groups during instructor directed and assisted practice and the heterogeneity of these groups. As expected, instruction was most commonly provided to whole classes (44%). In other studies, instruction was provided to students in large (15%) and small groups (36%). Students were grouped homogenously with respect to reading ability in about half of the eligible interventions.

Nature of Student Practice

Virtually every version of MCSI located for this study includes instructor directed instruction and assisted practice. The majority of interventions (74%) also provided students with opportunities to practice strategies on their own or in small groups without direct intervention by the teacher or computerized agent. In the remaining studies, students primarily practiced strategies while working collaboratively with the teacher or digital agents, who provided guidance for how and when to apply each strategy. In about 25% of the interventions, student practice was completed individually; in about 20%, there was a mix of collaborative and individual work; and in 25%, students exclusively practiced strategies while working with other students.

A major goal of this study was to investigate the relationship between treatment effectiveness and emphasis on self-regulated comprehension. As explained in Chapter 2, self-regulated comprehension is comprised of several factors, including: setting specific goals or purposes for a reading activity, monitoring one's progress toward those goals,

using strategies flexibly as needed to aid in accomplishing these goals, and monitoring the relative efficacy of chosen strategies and selecting new ones when needed. As shown in Table 8, these components were explicitly included in only a minority of eligible interventions. Students set their own reading goals in about 30% of the interventions; monitored comprehension during student practice in about 64%; selected strategies flexibly for different purposes in about 38%; and monitored strategy effectiveness in only 9%. These four components were used to create a composite variable—emphasis on self-regulated comprehension—that had a mean of 1.40 (SD=1.17) across all collected reports. An additional feature of student practice is whether or not students were taught to monitor their own strategy behavior, which occurred about 17% of the time. This was often accomplished by teaching students to keep a list or make checkmarks on a list provided by the teacher to keep track of each strategy they used. Although this component of student practice has links to self-regulation, it was not included in the self-regulation scale because it tended to correlate negatively with other factors in that scale.

Text Environment

The final category of pedagogical characteristics listed in Table 8 shows how the instructional materials varied across the eligible interventions. Nearly half of the interventions focused exclusively on expository or informational text, about 24% included a mix of expository and narrative materials, and 19% only included narrative texts. These texts were drawn from several sources, but most commonly, they came from basal or basal-like materials and photocopied passages (70%). Authentic materials (trade books, magazines, publicly available websites) were used less commonly (11%).

Students were rarely given opportunities to select their own reading materials in these versions of MCSI; in 86% of the reports, teachers selected most or all of the texts students read.

The instructional level of the texts used during MCSI was rarely reported. When text level could be estimated from other information in the report, it was found that in about half of the reports (47%), students were given reading materials on their assigned grade level regardless of ability. In about 17%, students read primarily below their assigned grade. An attempt was also made to code the readability of instructional texts relative to students' ability levels, but sufficient information to code this item was almost never provided.

Layer 3: Moderator Analyses

This portion of the analysis was intended to identify the characteristics of MCSI associated with maximum effectiveness, given the substantial between-study variability observed in layer 1.

Exploratory Breakouts by Framework and Sample Characteristics

The first question of interest was to summarize the effectiveness of the various MCSI frameworks. Because only a few studies were located for many of the frameworks, a formal comparison across treatment types (using the Q statistic) was not possible. However, for descriptive purposes, the mean effect size for each outcome was calculated for each of the ten framework types. These summary effects are reported in Table 10.

The most commonly represented framework, Reciprocal Teaching, had a mean ES of 0.31 for standardized measures of reading comprehension, while studies in the non-branded category had a mean of 0.37. Slightly larger impacts on standardized comprehension performance were obtained for Informed Strategies for Learning (0.81), Peer-Assisted Learning Strategies (0.43), and Concept-Oriented Reading Instruction (0.62), but these summary effects do not reach statistical significance because of the small number of studies employing these frameworks.

For performance on non-standardized comprehension measures, Reciprocal Teaching (0.46), acronym frameworks (0.85), and non-branded frameworks (0.55) were associated with positive and statistically significant mean ESs. Informed Strategies for Learning and Transactional Strategies Instruction were also associated with statistically significant effect sizes, but these estimates must be interpreted with caution because only a few studies were included in these categories. For strategy knowledge and strategy use, non-branded frameworks had a mean effect size of 0.77 and 0.78. Reciprocal Teaching had a mean impact of 0.58 on strategy use, compared to 0.95 for Concept-Oriented Reading and 1.03 for digital frameworks.

To examine the differential impact of MCSI by student population, summary effects were computed for studies with different sample characteristics. Again, formal subgroup analyses were not warranted given the small number of studies in each category. For standardized comprehension measures, the impact of MCSI did not differ significantly from 0 for studies that included mostly White students, mostly African American or Hispanic students, or mostly second language learners. Positive and significant effects were found for studies that included mostly average or above readers

(0.51), mostly struggling readers (0.40), and mostly students identified as having LD/RD (0.43). For non-standardized measures of comprehension, MCSI had a positive effect on White students (1.00), minority students (0.62), students with LD/RD (1.03), and struggling readers (0.64). A similar pattern was seen when strategy knowledge was the outcome; positive effects were obtained for White students (0.62), minority students (1.70), students with LD/RD (1.14), and struggling readers (0.78). Positive treatment effects were also observed for minority students (1.97), students with LD/RD (1.12), average and above average readers (0.81), and struggling readers (0.95) on measures of strategy use.

Finally, to examine the impact of MCSI on students across the upper-elementary and middle school grades, average treatment effects were calculated for each grade. This analysis was difficult because most studies in this literature are conducted in multiple grades and researchers rarely report separate data for different grade levels. Although formal comparisons were not possible, there was some evidence that 4th and 5th graders experienced lower impacts on standardized measures of comprehension, especially when compared to 7th graders. On non-standardized measures, 4th, 6th, and 7th graders appear to benefit more compared to 5th and 8th graders. For strategy knowledge, 6th graders had the lowest treatment effects, while 5th and 6th graders had the smallest gains on strategy use.

The Relationship of Method and Effect Size

As explained in the Methods, five characteristics of study design were treated as indices of methodological quality: use of random assignment to condition, use of systematic procedures to measure fidelity of implementation, use of a comparison

condition that only varies on one dimension, empirical demonstrations of baseline equivalence, and the inclusion of pre-test information so that post-test data could be adjusted for potential group differences. In addition to these, several design characteristics that are not direct indices of study quality were also tested.

For standardized measures of comprehension, two quality indicators were found to have marginally significant ($p<0.30$) relationships with treatment effectiveness, as shown in the top portion of Table 12. Studies in which fidelity was measured tended to produce lower effect size estimates. Effect sizes tended to be lower when they were adjusted for pre-test differences. Using this model, the predicted impact of a study using fidelity procedures and providing pretest adjusted post-test information is 0.16, which is a substantial reduction from the weighted mean effect across all studies.

The top portion of Table 13 shows the relationship between method variables and effect size for non-standardized measures of comprehension. Again, two quality variables—assignment type and use of fidelity measures—were found to be negatively related to effect size. The predicted impact of a high quality study with random assignment and fidelity procedures in place is 0.31. This value is substantially lower than the mean effect of 0.55, suggesting a strong relationship between quality and treatment effectiveness.

For measures of strategy knowledge, effect size estimates also tend to be lower for studies in which fidelity is measured and for treatment estimates that are adjusted for post-test differences, while they tend to be higher for studies that used random assignment (Table 14). The predicted impact of a study using randomly assigned groups with fidelity procedures and pretest adjustments in place is 0.62. This represents a slight

reduction in impact compared to the mean effect across all studies (0.73) for this outcome. Effect sizes were also higher when researchers provided an alternative intervention to the comparison group.

Finally, for measures of strategy use, a marginally significant positive relationship was observed between effect size and assignment type and a negative relationship for use of fidelity measures (see top portion of Table 15). The predicted effect size from studies with these two characteristics is 0.77, which is identical to the weighted mean effect size across all studies. Reductions in ES were seen for student samples that were drawn from multiple schools. Higher ESs were obtained from studies that used alternative interventions in the comparison condition.

The Control Models

Methodological characteristics (described above), publication status, measurement characteristics, and sample characteristics were all considered possible control variables. Table 11 shows the zero-order correlations of these characteristics and effect size (for each outcome separately). The results for each outcome are described in the sections that follow.

Standardized reading comprehension. When standardized reading comprehension was the outcome, slightly larger effect sizes were observed for studies that used student-level assignment compared to those that assigned intact groups (β = 0.21); for studies published in journals compared to unpublished studies (β = 0.25); and for studies for which sample sizes had to be estimated in order to compute effect sizes (β = 0.21). Treatment effects were slightly lower when treatment and control groups were

highly equivalent at baseline (β = -0.18); when implementation was systematically tracked by the researchers (β = -0.21); when post-test effect sizes were adjusted for pre-test differences (β = -0.32); and when the majority of students were 4th or 5th graders (β = -0.31).

When these seven variables were entered simultaneously, four maintained a sizeable relationship with treatment effectiveness, as shown in Table 12. Fidelity, pretest adjustment, and presence of 4th and 5th grade students continued to be associated with small decreases in treatment effectiveness, while journal publication was associated with a small increase.

Non-standardized reading comprehension. When reading comprehension was measured using non-standardized instruments, several characteristics were associated with small decreases in treatment effectiveness. Studies that used random assignment (β = -0.15), measured implementation fidelity (β = -0.16), and included a majority of students from 5th or 8th grade (β = -0.27) had slightly lower effect sizes. Studies that used student-level assignment (β = 0.20) and studies with high percentages of students identified with learning or reading disabilities (β = 0.25) tended to have higher treatment effectiveness, as well as studies that measured comprehension using a constructed response format (β = 0.18). When these factors were entered together, attention to fidelity, presence of 5th and 8th graders, presence of students with LD/RD, and use of a constructed response measure continued to have small but nontrivial associations with effect size (see Table 13).

Strategy knowledge. For this outcome, small positive relationships were observed for studies that used random (β = 0.26) and student-level (β = 0.32) assignment;

studies that compared MCSI to an alternative intervention (β = 0.23); studies that drew their samples mostly from 7th grade (β = 0.25); and studies in which most of the sample was identified as LD or RD (β = 0.35). Small negative associations were found between treatment effectiveness and the use of samples drawn from multiple schools (β = -0.36), attention to fidelity (β = -0.18), the use of pretest adjustments (β = -0.20), and having a sample that is mostly made up of 6th grade students (β = -0.31). Six of these characteristics maintain their association with treatment effectiveness when modeled together, as shown in Table 14.

Strategy use. Various control variables were also found to be correlated with the impact of MCSI on strategy use. Studies with samples drawn from multiple schools produced lower treatment estimates than single-school studies (β = -0.52), as did studies that tracked fidelity (β = -0.21), studies reported in journal articles (β = -0.47), and studies conducted with 5th and 6th grade students (β = -0.45). Studies that used random (β = 0.20) and student-level assignment (β = 0.23) and provided alternative interventions for students in the comparison condition (β = 0.54) tended to report higher effect sizes. Also, positive trends were observed for studies conducted with students identified as learning/reading disabled (β = 0.23) and studies that measured strategy use with single-genre expository tests (β = 0.23). Five of these characteristics that continued to have sizeable relationships with effect size when included together were retained in the final control model (see Table 15).

Influence of the Hypothesized Moderator Variables

The next step in the moderator analyses was to add the pedagogical and content features of MCSI individually to their respective control models. Moderators were selected for the multiple regression analysis based on their zero-order correlations with effect size (Table 16). The standardized regression coefficients obtained from the metaregression are semi-partial correlations that index the unique contribution of each characteristic after controlling for the contributions of the other independent variables in the model. The full models are useful not only for examining the contribution of the hypothesized moderators, but also for examining the relationship of methodological and sample characteristics to treatment effectiveness. These relationships are shown in Table 17 for the four outcomes examined in this study.

Standardized reading comprehension. Several pedagogical features were associated with decreased treatment effectiveness for this outcome (see correlations in Table 16). Studies that included a student practice portion in addition to instructor directed or assisted practice had lower impact ($\beta = -0.15$), along with studies that provided instruction in mixed ability groups ($\beta = -0.20$) and studies in which students were taught to monitor their own strategy behavior ($\beta = -0.27$). Slightly higher effect sizes were observed for interventions provided by teachers compared to those provided by researchers or computers ($\beta = 0.18$). Also, higher impact was observed for versions of MCSI that included instruction in narrative and expository text ($\beta = 0.25$). Several features that were hypothesized to improve the effectiveness of MCSI were not systematically associated with treatment impact for this outcome, including: amount of peer collaboration ($\beta = 0.00$) and emphasis on self-regulated comprehension ($\beta = 0.00$).

Note also that duration of instruction did not have a systematic relationship with effectiveness. To examine the possibility that duration would have a nonlinear relationship with effectiveness, a scatterplot of effect size organized by total minutes of instruction (Figure 9) was examined. This plot reveals a relatively flat relationship.

In addition to pedagogical features, the content of MCSI was hypothesized to influence treatment effectiveness. Interventions that taught students to identify main ideas ($\beta = 0.29$), use prior knowledge ($\beta = 0.16$), and analyze and reflect on their reading ($\beta = 0.25$) tended to produce higher results. Interventions that included all four of the RT strategies tended to be less effective than those that included fewer of these "big four" strategies ($\beta = -0.44$). The number of strategies included in the instructional repertoire did not appear to have a systematic relationship with treatment effectiveness.

When these characteristics were modeled together (as shown in Table 17), only analysis/reflection retained a sizeable relationship with treatment effectiveness in the full model ($\beta = 0.33$). As noted in the Table, the big four strategy variable was tested separately from the other strategy variables. This content characteristic does not appear to have a strong impact on MCSI effectiveness after accounting for the instructional characteristics in the model.

Consistent with the exploratory breakouts described above, this model shows that gains on standardized measures of comprehension were reduced for studies conducted with 4th and 5th grade students ($\beta = -0.28$), although there is no theoretically plausible explanation for this. The final model also helps clarify the relationship between method and effect size. Even with the inclusion of pedagogical and content characteristics, indicators of design quality (fidelity and use of pretest adjustments) continue to be related

to effect size (β = -0.24 and -0.38, respectively). There is also evidence that published studies tend to report higher effect sizes than unpublished studies (β = 0.23).

Non-standardized reading comprehension. Table 16 shows how each pedagogical characteristic was associated with treatment effectiveness for this outcome. The inclusion of student practice was associated with a small increase in effectiveness after accounting for the control variables (β = 0.30). Also, increased effectiveness was observed for interventions conducted in a pull-out format compared to those that occurred during regular classroom activities (β = 0.27) and for interventions that included instruction in narrative and expository text (β = 0.14). Studies that prioritized self-regulated comprehension also tended to have slightly greater impact (β = 0.16). Also, this was the only outcome that had a systematic relationship with treatment duration, with longer interventions outperforming shorter ones (β = 0.21).

Of the various strategies that were taught in the MCSI interventions, the use of graphic organizers (β = 0.20), previewing (β = 0.23), and analysis/reflection (β = 0.27) were associated with increased effectiveness, while the use of the big four strategies was associated with decreased effectiveness (β = -0.22). Again, no systematic relationship was observed between treatment effectiveness and strategy concentration (β = 0.10).

When the nine pedagogical and content characteristics were included together in a weighted regression (Table 17), the inclusion of student practice (β = 0.18) and use of a pull-out setting (β = 0.28) maintained small but statistically significant positive relationships with treatment effectiveness. Also, the inclusion of strategic content related to graphic organizers (β = 0.24), previewing (β = 0.25), and analysis/reflection (β = 0.29) tended to increase treatment impact.

As with standardized measures, the relationship of method to effectiveness does not fully disappear with the inclusion of content and pedagogical variables. Studies in which fidelity measures were used tended to produce lower treatment estimates (β = -0.22). As described previously, studies conducted with 5th and 8th grade samples reported lower gains (β = -0.30). There was a small tendency for studies with students identified with learning or reading disabilities to produce larger gains (β = 0.11), although this result was only marginally significant.

Strategy knowledge. A more conservative model had to be developed for this measure because the number of cases was relatively small. Therefore, pedagogical characteristics were considered as moderators but content characteristics (individual strategies) were not. This decision was also warranted because it did not make sense to test the impact of instruction in an individual strategy on a measure that assesses general knowledge of a set of strategies.

The pedagogical characteristics associated with treatment impact, as shown in Table 16, were computerized instruction (β = 0.20), amount of collaboration (β = -0.33), and exclusive focus on expository text (β = -0.47). When entered together (Table 17), all three maintain the same direction and strength of association, but they are only marginally statistically significant.

This model also shows an added benefit of MCSI for students identified with learning or reading disabilities (β = 0.55). None of the quality indicators were related to effect size in the full model, but sampling students from multiple schools was found to have a negative relationship with treatment effectiveness (β = -0.54).

Strategy use. This final model was constructed using the same procedures as the strategy knowledge model above. Pedagogical characteristics associated with increased effectiveness for this outcome included: providing instruction in a pull-out setting (β = 0.24) and teaching students to monitor and control their own strategy behavior (β = 0.18). Negative associations were found for studies that included an emphasis on student practice (β = -0.22), studies that provided instruction in a content area setting (β = -0.18), interventions provided to whole classes of students (β = -0.22), and those that prioritized peer collaboration over individual strategy practice (β = -0.25). In the full model (Table 17), there was some evidence that amount of collaboration was negatively associated with treatment effectiveness (β = -0.19), while none of the other hypothesized moderators appeared to exert unique influence.

After controlling for the other factors in the model, one methodological quality variable (random assignment) maintained a marginally significant negative relationship to effect size (β = -0.20). Published articles were also associated with lower effect size estimates (β = -0.40) as well as the use of samples drawn from multiple schools (β = -0.22). There was also a small marginally significant reduction in MCSI effectiveness for samples made up mostly of 5^{th} or 6^{th} graders (β = -0.26).

Table 5: Effect Size Estimates for Each Study Included in the Meta-Analysis

		Effect Size Estimates*			
Report	**MCSI condition**	*1*	*2*	*3*	*4*
Aarnoutse, Van den Bos & Brand-Gruwel, 1998	Listening comprehension training	-0.21			0.40
Anderberg, 1996	RT (Limited English Proficient)	-0.04			
	(Not limited English Proficient)	-0.03			
Babbs, 1984	Strategy instruction with monitoring cards		0.77		
Baumann, Seifert-Kessell & Jones, 1992	Think Aloud Instruction	0.19		1.10	0.71
Bereiter & Bird, 1985	Think Aloud Instruction (modeling + explanation)	1.34	0.80		0.86
Berkeley, 2007	Strategy instruction		-0.06	0.67	
Block, 1993	Strategy instruction	1.77[w]			
Bottomley, 1993	RT		2.25		
Bottomley, 1993	TSI		2.13		
Brady, 1990	Semantic mapping + RT	0.54	0.60		
Brimmer, 2004	Think Aloud Instruction	0.11		0.20	
Burke, 1992	Strategy training		-0.14		
	Strategy + attribution training		0.14		
Calhoon, 2005	PALS + phonological training	0.70			
Carr & Borkowski, 1989	Strategy instruction + attribution training		0.17	0.13	0.19
Carriedo & Alonso-Tapia, 1995a	Comprehension Strategy Training	0.13	0.71	0.59	0.44
Carriedo & Alonso-Tapia, 1995	Comprehension Strategy Training				
	(6th grade)		0.23	0.56	0.37
	(7th grade)		-0.50	0.73	0.85
	(8th grade)		-0.17	0.99	0.74
Chapman, 1997	RT		0.80	1.03	
Dao, 1993	RT	1.48			
Dehn, 1992	Informed strategy training + computer practice	-0.28			
Dermody & Speaker, 1995	RT	-0.29			

Dole, Brown & Trathen, 1996	Strategy Instruction		1.31		
Duffy, Roehler & Meloth, 1986	Explicit verbal explanations	0.09			
Ehlinger, 1988	Fullrange cognitive modeling		-0.04		
Englert & Mariage, 1991	P.O.S.S.E.		0.68		1.18
Englert, Tarrant & Mariage, 1994	P.O.S.S.E.			0.37	
Esser, 2001	Metacognitive strategy training + attributional training		2.02	2.87[w]	2.80
	Metacognitive strategy training		1.16	1.10	2.80
Etsey, 2004	Comprehension strategy instruction		0.60		
Fuchs, Fuchs & Kazdan, 1999	PALS (At-risk students) (Average/high achieving)	 0.59 0.00			
Fuchs, Fuchs & Mathes, 1997	PALS (Learning disabled) (Low achieving) (Average achieving)		 0.58 0.37 0.11		
Galbert, 1989	RT (4th grade)	 0.35			
	(5th grade)	0.10			
Gee, 1997	Strategic knowledge condition		0.05		
	Strategies + content		-0.21		
Gilroy & Moore, 1988	RT (4th grade)		 2.78[w]		
	(6th grade)		2.78[w]		
Glaeser, 1998	Inclusive Strategies Instruction Model	0.14	-0.12		1.49
Gordan,1995	Reading strategy instruction	0.54		1.52	
Guthrie, Anderson & Alao, 1999	CORI		-0.20		-0.07
Guthrie, Van Meter & Hancock, 1998	CORI		0.39		1.11
Hahn, 1983	Comprehension monitoring training		1.04		
Hedin, 2008	Think Aloud Instruction	-0.52	-0.43	0.01	
Houtveen & Van de Grift, 2007	Metacognitive strategy instruction	0.43		0.26	
Jeffers, 1990	Metacognitive strategy instruction	0.16			

Johnson-Glenberg, 2000	RT	0.30	0.31		0.33
Johnson, 2002	Strategy instruction	0.19		0.23	
Kelly, 1984	Metacognitive awareness and generative verbal strategies (Below average)		-0.20		
	(Average/above average)		0.06		
Kelly, Moore & Tuck, 1994	RT	1.26	2.06		
Kim, Vaughn & Klingner, 2006	Computer-Assisted Collaborative Strategic Reading (CA-CSR)	0.45	0.87		1.10
Kinnunen & Vauras, 1995	Strategy training		0.87		
Klingner, Vaughn & Arguelles, 2004	Collaborative Strategic Reading				0.43
	(Learning disabled)	0.31			
	(Low achieving)	0.31			
	(Average/high achieving)	0.19			
Klingner, Vaughn, & Schumm, 1998	Collaborative Strategic Reading	0.30			
Labercane & Battle, 1987	RT	0.29			
Lau & Chan, 2007	Cognitive strategy instruction		0.41		0.51
Lederer, 2000	RT		0.41		0.63
Leu, Castek & Hartman, 2005	High intensity internet integration + RT	-0.20			
Levin, 1989	ISL + RT	0.10	-0.11		
Li, 1996	Strategy instruction		0.40	0.10	
Loranger, 1997	TSI	0.67			
Lovett, Borden & Warren-Chaplin, 1996	Strategy Reciprocal Teaching Program		-1.28		2.80[w]
Lubliner, 2001	Questioning + clarifying	-0.12			
Lysynchuk, Pressley & Vye, 1990	RT				
	(4^{th} grade)	0.50			
	(7^{th} grade)	0.35			
Mason, 2004	TWA		-0.05		
McElvain, 2005	Transactional Literature Circles	0.30			
McKeown, Beck, & Blake 2009	Strategies Instruction (year 1)		0.39		
McKeown, Beck, & Blake 2009	Strategies Instruction (year 2)		-0.02		

McNamara, O'Reilly & Best, 2006	iSTART		0.40		
Miranda, Villaescusa & Vidal-Abarca, 1997	Self-instruction training		2.78[w]	0.69	
Nelson, 2003	Metacognitive strategy instruction	0.40	0.47		
Nolan, 1991	Self-questioning + prediction	1.32			
O' Hara, 2007	CORI-STAR	0.73	1.89	0.89	2.30
Ocasio, 2006	RT		-1.97		
Padron, 1985	RT	-0.06			1.01
Palincsar & Brown, 1984	RT		1.86		1.30
Palinscar, 1987	RT		1.07		0.34
Paris & Oka, 1986	Informed Strategies for Learning			0.54	
Paris, Cross & Lipson, 1984	Informed Strategies for Learning	0.06	0.35	0.63	
Payne & Manning, 1992	Comprehension monitoring training	1.10		1.70	
Perez, 2002	COMPRENDER	0.24	0.86		
Philbrick, 2002	Strategy instruction		0.39	1.20	
Pickens & McNaughton, 1988	Comprehension strategy training	0.67			
Piercy, 1997	Multi-strategy instruction	-0.09			
Radcliffe, Caverly & Hand, 2008	P.L.A.N.		1.01		0.55
Rauenbusch & Bereiter, 1991	Comp strategies instruction	0.75			
Rogevich & Perin, 2008	TWA + written summarization		2.78[w]		
Russell, 1997	RT		0.19		
Saenz, Fuchs & Fuchs, 2005	PALS (Learning disabled) (Low achieving) (Average/high achieving)		 0.39 0.24 0.46		
Salomon & Globerson, 1989	Metacognitive Instruction with Reading Partner	1.42		2.87[w]	
Shortland-Jones, 1986	Meta-schema instruction	0.15			0.23
Smith, 2006	Reading comprehension strategy instruction		0.86		
Souvignier & Mokhlesgerami, 2006	Strategy instruction + cognitive self-regulation	0.12		0.53	0.39

Sung, Chang & Huang, 2008	Computer-Assisted Strategy Teaching and Learning Environment (CASTLE)				
	(High achievers)		0.21		1.10
	(Low achievers)		0.50		0.90
Takala, 2006	RT		0.16		-0.43
Taylor & Frye, 1992	Strategy Instruction				
	(5^{th} grade)		0.32		0.39
	(6^{th} grade)		-0.45		0.03
Tregaskes & Daines, 1989	Metacognitive strategy instruction		0.47		
Uttero, 1992	Instruction Modeling Cooperative Engagement	0.09	2.78[w]		
	(4^{th} grade)				
	(6^{th} grade)	0.37	2.78[w]		
Van den Bos, Brand-Gruwel & Aarnoutse, 1998	Comprehension strategies through reading and listening	0.37		0.34	0.65
Villani, 1987	Explicit strategy instruction		0.19		
Walraven & Reitsma, 1992	Strat instruction with prior knowledge	0.79		1.73	
Wigfield, Guthrie & Perencevich , 2008	CORI	0.46	0.78		0.40
Wise, Ring & Olson, 2000	Accurate reading in context	-0.14			
Woodward, 1991	Reading and Thinking Strategies	1.77[w]	2.78[w]		

[*] Construct 1 = standardized measures of reading comprehension; Construct 2 = non-standardized measures of reading comprehension; Construct 3 = strategy knowledge; Construct 4 = strategy use

[w] These are Winsorized values for ESs identified as outliers.

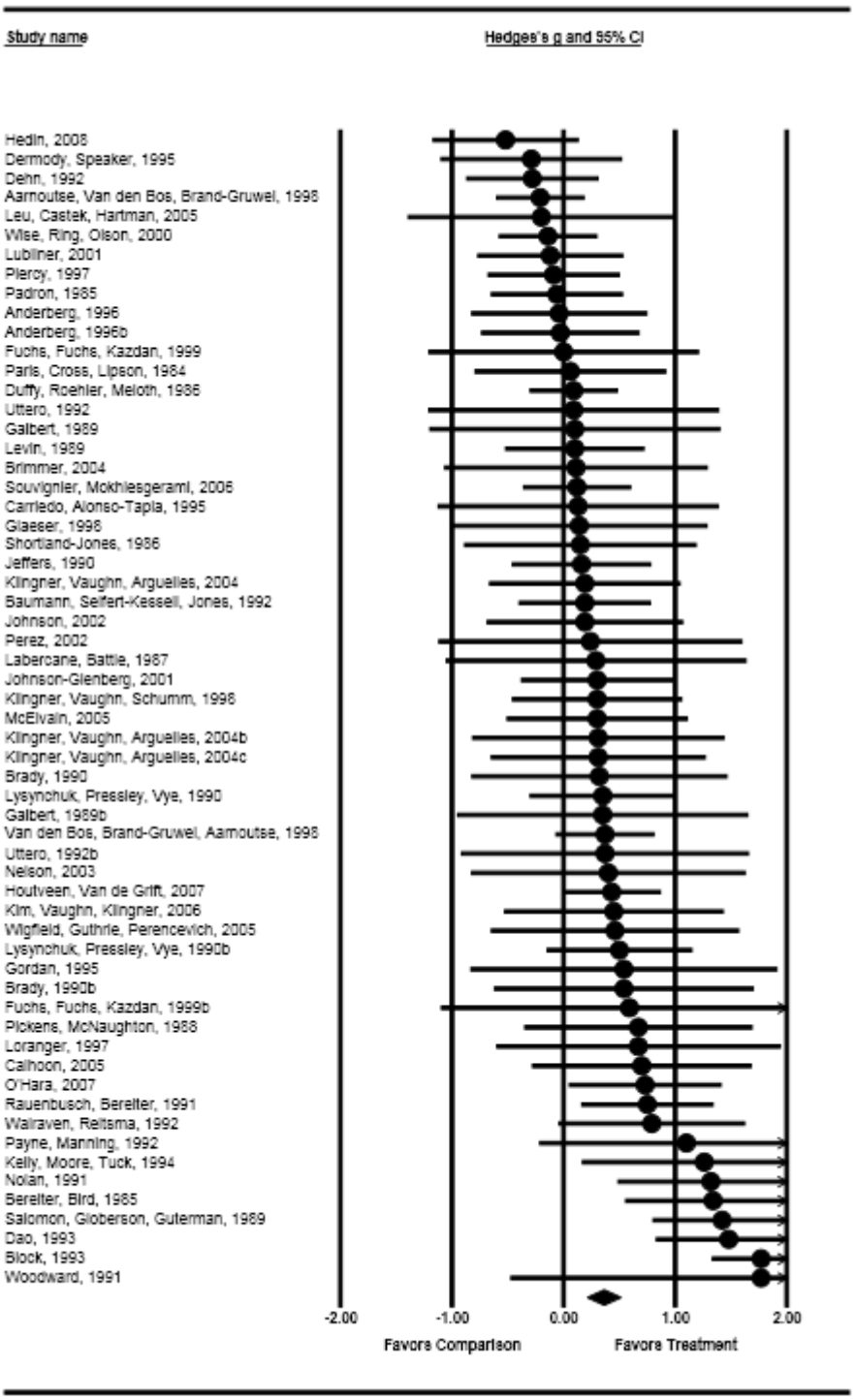

Figure 4: Forest Plot of Effect Sizes for Standardized Measures of Comprehension

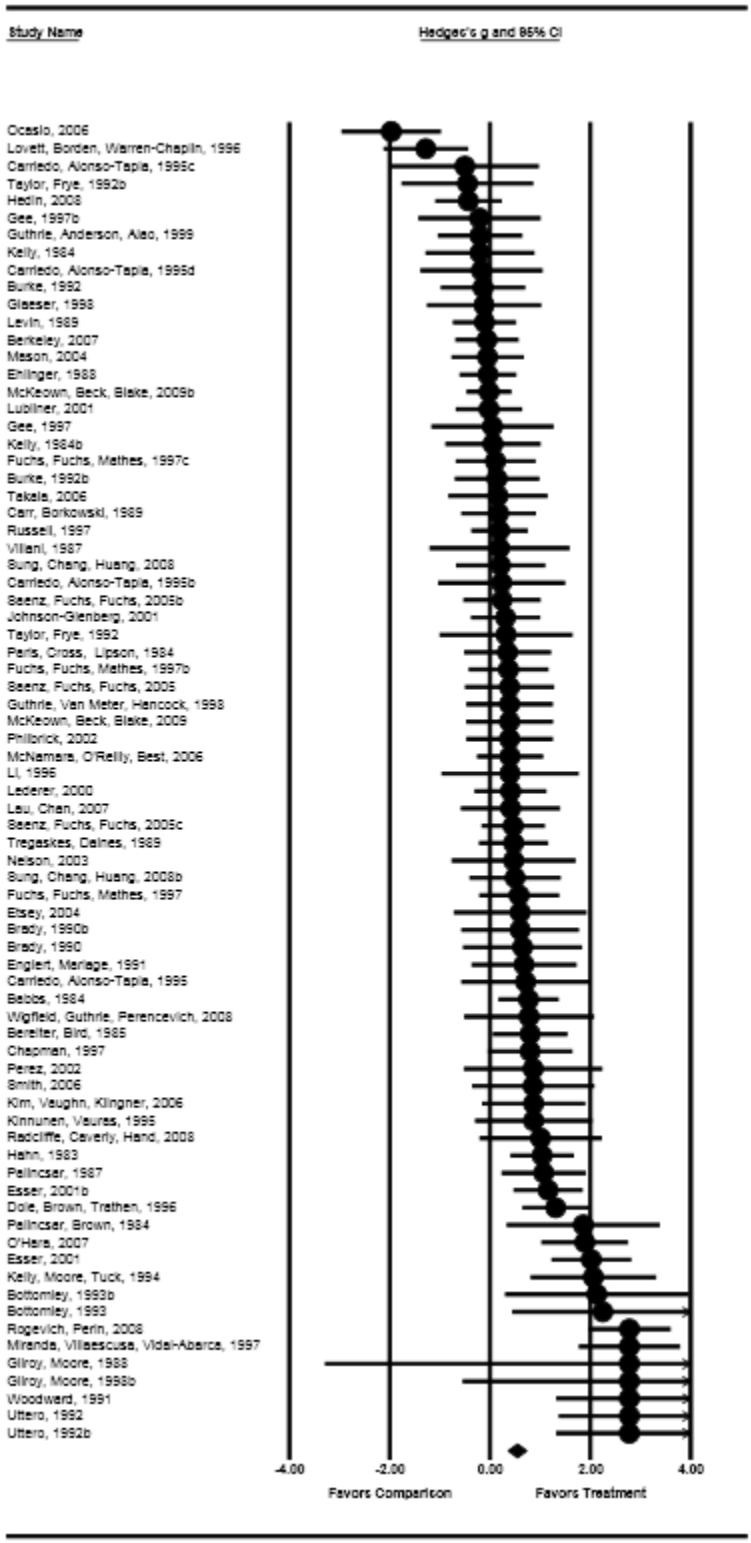

Figure 5: Forest Plot of Effect Sizes for Non-Standardized Measures of Comprehension

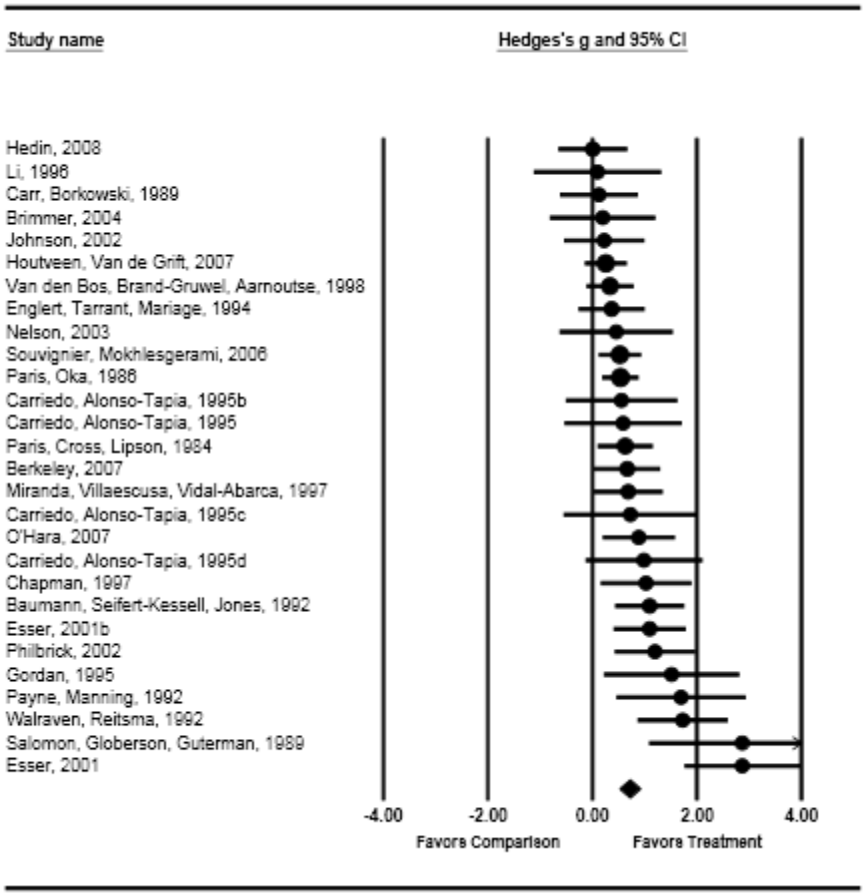

Figure 6: Forest Plot of Effect Sizes for Measures of Strategy Knowledge

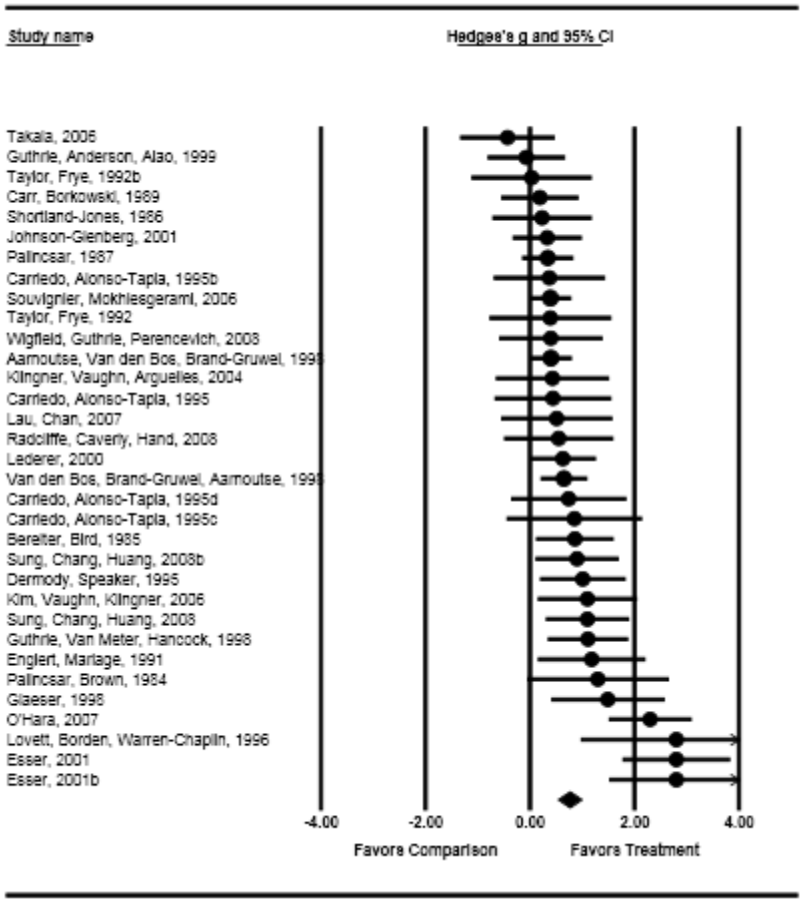

Figure 7: Forest Plot of Effect Sizes for Measures of Strategy Use

Table 6: Main Effects and Heterogeneity Statistics

						Heterogeneity		
Outcome	**N**	**Mean ES**	**SE**	**95% Confidence Interval**	**P-value**	**Q**	**P-value**	**I^2**
Standardized Measures of Reading Comprehension	60	0.36	0.08	0.21 – 0.51	0.00	127.08	0.00	53.57
Non-Standardized Measures of Reading Comprehension	76	0.55	0.09	0.36 – 0.73	0.00	233.58	0.00	67.89
Strategy Knowledge	28	0.73	0.10	0.53 – 0.92	0.00	55.57	0.00	51.41
Strategy Use	33	0.77	0.11	0.54 – 0.99	0.00	74.49	0.00	59.74

Table 7: Summary of Analyses to Assess the Risk of Publication Bias

	Begg and Mazumbar Rank Correlation		**Egger's Regression Intercept**			**Duval and Tweedie's Trim and Fill**	
Outcome	**Tau**	**p-value (2-tailed)**	**Intercept (SE)**	**95% Confidence Interval**	**p-value (2-tailed)**	**N of trimmed studies**	**Adjusted point estimate**
Standardized Comprehension	0.18	0.05	0.36 (0.51)	-0.66 – 1.39	0.48	0	--
Non-standardized Comprehension	0.22	0.01	1.48 (0.61)	0.26 – 2.70	0.02	2	0.54
Strategy Knowledge	0.33	0.01	1.76 (0.61)	0.50 – 3.02	0.01	0	--
Strategy Use	0.25	0.04	1.61 (0.70)	0.19 – 3.03	0.03	0	--

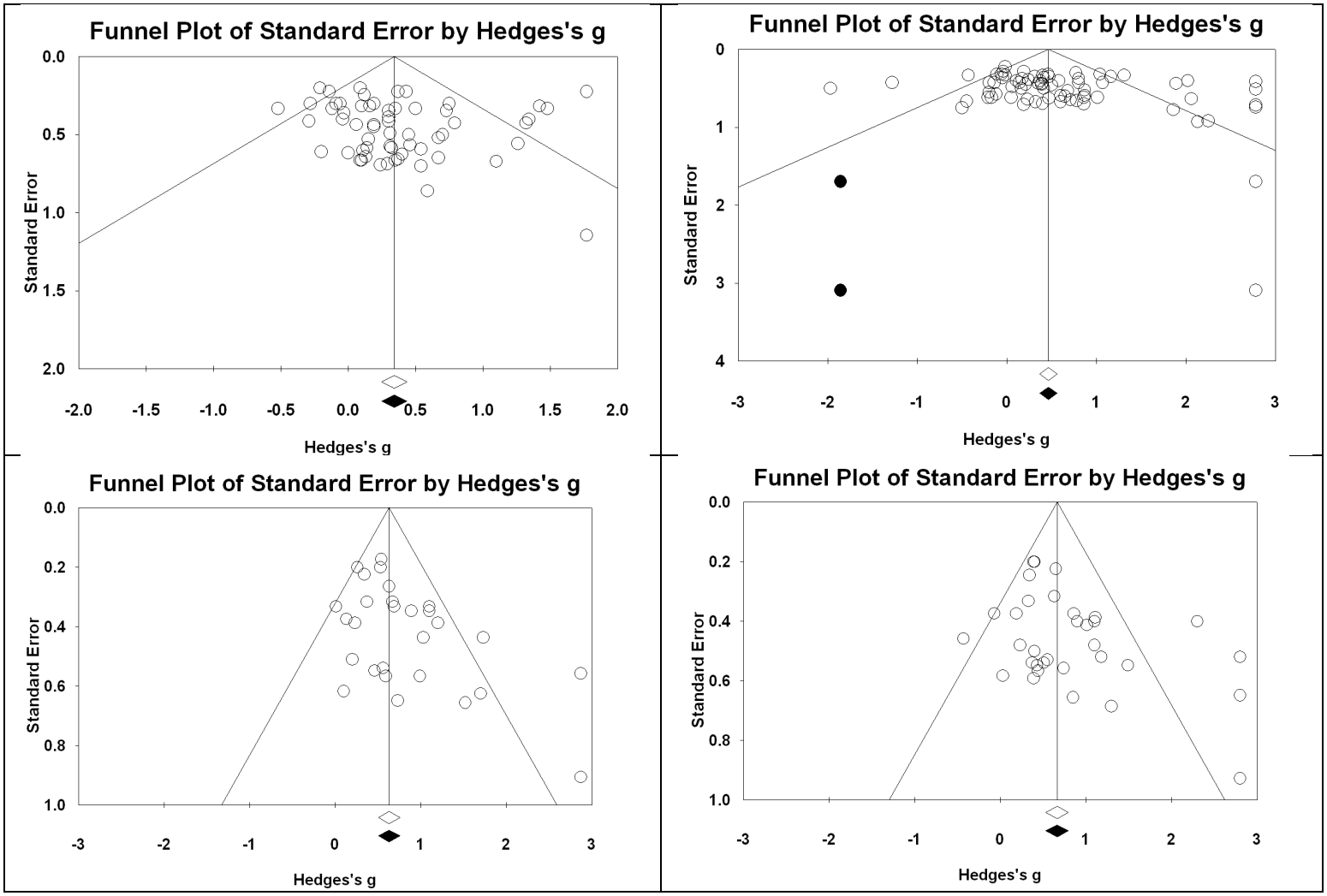

Figure 8: Funnel Plots to Examine Publication Bias
(Top left quadrant = standardized reading comprehension; top right = non-standardized comprehension; bottom left = strategy knowledge; bottom right = strategy use)

Table 8: Descriptive Statistics for Studies Included in the Meta-Analysis (Categorical Variables)

	Frequency
Methodological Characteristics	
Publication Type	
Journal article	59
Dissertation	40
Other	2
Type of Assignment to Condition	
Simple random	27
Random, with matching	15
Simple non-random	40
Non-random, with some attempt to equate groups	18
Level of Assignment	
Student	32
Group smaller than class	10
Class	41
School	17
Baseline Equivalence (n=127 samples)	
Treatment and comparison deemed equivalent on—	
Reading comprehension	85
Strategy knowledge	20
Strategy use	22
Other measure	56
Demographic characteristics	30
Treatment and comparison deemed equivalent on—	
0 of the above measures	25
1 of the above measures	36
2 of the above measures	25
3 of the above measures	33
4 of the above measures	7
5 of the above measures	0
Pretest adjustments (n=197 effect sizes)	
Effect sizes could be adjusted for pretest differences with data given	153
Effect sizes could not be adjusted for pretest differences	44
Amount of Attrition After Assignment (n=197 effect sizes)	
None	56
1 – 5%	32
6 – 10%	14
11 – 15%	11
More than 15%	34
Unknown	50
Fidelity is Systematically Measured	
Yes	34
No / not reported	66
Sample is drawn from—	
Multiple schools	48
Single school	50
Type of Comparison Condition	
No treatment control	11

Business as usual instruction	58
Alternative intervention	32
Content Similarity of Comparison and Treatment	
Differ on other factors besides strategy focus	30
Strategy focus is the only major difference	71
Sample Characteristics	
Most of sample (75%+) is in— (n=127 samples)	
4th grade	26
5th grade	17
6th grade	18
7th grade	8
8th grade	7
Mixed grade	51
Most of sample (75%+) is identified as—	
Having a learning or reading disability (LD/RD)	18
Second/English language learners (SLL)	10
Struggling/below average readers	55
General Characteristics of the Instructional Framework	
Treatment Framework	
Reciprocal Teaching	23
Informed Strategies for Learning	3
Peer-Assisted Learning Strategies	4
Collaborative Strategic Reading	2
Think-Aloud Instruction	7
Transactional Strategies Instruction	2
Concept-Oriented Reading Instruction	3
Acronym Instruction	4
Digital Versions	6
Non-Branded	45
Treatment Specificity	
General approach, non-curriculum	22
General curriculum	33
Highly specified curriculum	46
Duration	
Short treatment (100-600 minutes)	36
Moderate treatment (601-1440 minutes)	34
Long treatment (1441-9600 minutes)	30
Primary Instructor	
Researcher/research team	48
Half researcher/half teacher	4
Teacher	42
Computerized instruction	5
Other	2
Instructional Setting	
Content area class (science or social studies)	16
Other setting (reading class and other non-content settings)	85
Treatment Occurs During—	
Regular classroom activities	60
Supplemental pull-out program	37
Other (out-of-school enrichment, etc.)	4

Instructional Delivery	
Instructional Grouping During Instructor Directed/Assisted Practice	
Individual students	4
Small groups (2-6 students)	36
Large groups (7-12 students)	15
Whole class	44
Ability Grouping During Instructor Directed/Assisted Practice	
Instruction provided primarily in same ability groups	46
Instruction provided primarily in mixed ability groups	49
Nature of Student Practice	
Student Practice is included apart from instructor directed/assisted practice?	
No (instructor directed and/or assisted practice only)	26
Yes (students practice strategies without teacher or computerized prompts at some point)	74
Amount of Collaboration During Student Practice	
Students mostly work individually	26
Mix of collaborative and individual strategy practice	21
Students mostly work collaboratively with peers	25
Goal Setting During Student Practice	
Students set their own reading goals/purposes	29
Goals/purposes are mostly provided by the teacher	67
Comprehension Monitoring During Student Practice	
Students regularly evaluate comprehension	64
Students complete strategies without comprehension monitoring	33
Flexible Strategy Selection During Student Practice	
Students select strategies flexibly (e.g., to meet specific goals)	38
Students primarily use strategies in a pre-determined sequence	59
Other Features Related to Student Strategy Use:	
Student taught to monitor strategy behavior	17
Students taught to monitor strategy effectiveness	9
Text Environment	
Primary Genre of Instructional Texts	
Narrative	19
Mix of narrative and expository	24
Expository/Informational	46
Instruction mostly includes—	
"Real world" texts (trade books, magazines, web pages)	11
Basal or basal-like passages	70
Text Selection	
Teachers select most/all texts that students read	86
Students have limited opportunities to self-select texts	3
Students have extensive opportunities to self-select texts	8
Grade Level of Instructional Texts	
Students primarily read on assigned grade level	47
Students primarily read below assigned grade level	17
Instructional Content	
Strategies that are explicitly taught include—	
Analyzing/evaluating/reflecting	8
Attending to graphic sources/pictures	2
Clarifying (words and/or concepts)	52

Using graphic organizers (story maps and concept maps)	17
Identifying main ideas	34
Generating questions	72
Identifying author's purpose	2
Identifying genre	2
Inferring unstated information	15
Monitoring understanding of words/phrases	7
Monitoring comprehension	32
Monitoring text coherence	2
Paraphrasing/retelling	17
Predicting	67
Previewing/skimming	14
General fix-up strategies (re-reading and reading ahead)	23
Regulating reading speed	5
Searching for information	7
Establishing a purpose/goal for reading	22
Summarizing	78
Taking notes	1
Using prior knowledge	38
Using external resources (dictionaries, etc.)	6
Using social resources/asking for help	6
Using text structure/organizational clues	12
Visualizing/mental imagery	14
Big Four	
Content includes all four RT strategies	38
Content includes three or fewer of the RT strategies	61
Role of Emergent Strategies	
All strategies are pre-planned	92
Some strategies emerge during instructional interactions	5

Unless otherwise noted, these frequency counts are for 101 research reports. When totals do not equal 101, data are missing for some reports. The discrepancy between the number of reports included in the full analyses (n=98) and the number included in this table is due to the last-minute elimination of three reports with unclear effect size information.

Table 9: Descriptive Statistics for Studies Included in the Meta-Analysis (Continuous Variables)

Variable	**N**	**Min.**	**Max.**	**Mean**	**SD**
Total number of strategies explicitly taught	99	2	12	5.63	2.24
Treatment span (days)	94	2	160	55.16	45.66
Number of lessons	91	2	160	27.24	25.85
Lesson length (minutes)	101	20	120	42.39	14.65
Total duration of instruction (minutes)	100	104	9600	1549.02	1937.32
Emphasis on self-regulated comprehension (composite score)	100	0	4	1.40	1.17

Table 10: Summary of Treatment Effects by Framework and Sample Characteristics

	Standardized Measures of Reading Comprehension				*Non-Standardized Measures of Reading Comprehension*				*Strategy Knowledge*				*Strategy Use*			
MCSI Frameworks	**Mean ES**	**n**	**SE**	**p-value**	**Mean ES**	**n**	**SE**	**p-value**	**Mean ES**	**n**	**SE**	**p-value**	**Mean ES**	**n**	**SE**	**p-value**
Reciprocal Teaching	0.31	16	0.16	0.06	0.46	15	0.23	0.04	1.03	1	0.59	0.08	0.58	8	0.25	0.02
Informed Strategies for Learning	0.81	2	0.48	0.09	1.30	2	0.63	0.04	0.58	2	0.32	0.07	--	--	--	--
Peer-Assisted Learning Strategies	0.43	3	0.46	0.34	0.36	6	0.32	0.26	--	--	--	--	--	--	--	--
Collaborative Strategic Reading	0.28	4	0.33	0.41	--	--	--	--	--	--	--	--	0.43	1	0.76	0.57
Think Aloud Instruction	0.37	5	0.27	0.17	0.09	3	0.43	0.83	0.47	3	0.31	0.14	0.86	1	0.65	0.18
Transactional Strategies Instruction	0.67	1	0.80	0.40	2.13	1	1.14	0.06	--	--	--	--	--	--	--	--
Concept-Oriented Reading Instruction	0.62	2	0.46	0.17	0.71	4	0.42	0.09	0.89	1	0.52	0.09	0.95	4	0.34	0.00
Acronym Frameworks	--	--	--	--	0.85	5	0.37	0.02	0.37	1	0.51	0.47	0.87	2	0.52	0.10
Digital Frameworks	0.41	4	0.32	0.20	0.48	4	0.40	0.23	2.87	1	0.99	0.00	1.03	3	0.39	0.01
Non-Branded Frameworks	0.37	23	0.13	0.01	0.55	36	0.14	0.00	0.77	19	0.13	0.00	0.78	14	0.19	0.00

Sample Characteristics																
Mostly White	0.03	7	0.23	0.91	1.00	10	0.25	0.00	0.62	6	0.21	0.00	0.50	3	0.35	0.16
Mostly African American and Hispanic	0.14	8	0.21	0.50	0.62	9	0.26	0.02	1.70	2	0.37	0.00	1.97	3	0.40	0.00
Mostly LD/RD	0.43	8	0.22	0.05	1.03	11	0.24	0.00	1.14	5	0.22	0.00	1.12	6	0.28	0.00
Mostly SLL	0.08	4	0.30	0.79	0.46	7	0.30	0.13	--	--	--	--	0.90	1	0.64	0.16
Mostly average or above average in reading ability	0.51	6	0.25	0.04	0.20	15	0.21	0.34	0.18	2	0.43	0.67	0.81	5	0.31	0.01
Mostly below average in reading ability (struggling readers)	0.40	27	0.12	0.00	0.64	30	0.14	0.00	0.78	11	0.16	0.00	0.95	15	0.17	0.00
Mostly 4th graders	0.25	18	0.15	0.10	0.92	9	0.31	0.00	0.66	4	0.26	0.01	0.64	3	0.40	0.10
Mostly 5th graders	0.05	6	0.23	0.82	0.06	12	0.22	0.78	0.71	5	0.25	0.00	0.45	4	0.31	0.14
Mostly 6th graders	0.08	5	0.30	0.79	0.51	14	0.23	0.02	0.28	4	0.29	0.34	0.46	7	0.26	0.08
Mostly 7th graders	0.69	4	0.28	0.01	0.55	3	0.54	0.30	1.51	2	0.59	0.01	0.85	3	0.46	0.06
Mostly 8th graders	0.16	1	0.54	0.76	-0.05	6	0.32	0.88	0.99	1	0.67	0.14	0.74	1	0.74	0.32

Table 11: Correlations of Method, Measurement, and Sample Characteristics with Adjusted SMD Effect Sizes

	Standardized Measures of Reading Comprehension			*Non-Standardized Measures of Reading Comprehension*			*Strategy Knowledge*			*Strategy Use*		
	k	**R***	**P-value**	**k**	**R***	**P-value**	**k**	**R***	**P-value**	**k**	**R***	**P-value**
Quality Indicators												
Assignment type	60			76			28			33		
Random vs. nonrandom (all)		0.11	0.45		-0.15	0.18		0.26	0.12		0.20	0.23
Random vs. nonrandom with no effort to equate groups		0.07	0.72		-0.14	0.28		0.30	0.09		0.10	0.62
Random vs. nonrandom with some effort to equate groups		0.17	0.40		-0.15	0.33		-0.16	0.75		0.31	0.13
Equivalence of groups at baseline	60	-0.18	0.21	76	0.08	0.44	28	-0.01	0.95	33	0.16	0.34
Implementation fidelity systematically tracked	60	-0.21	0.15	76	-0.16	0.14	28	-0.18	0.29	33	-0.21	0.18
Content similarity of treatment and comparison conditions	60	0.01	0.94	76	-0.14	0.21	28	-0.15	0.40	33	0.00	0.98
Effect size is adjusted for pretest differences	60	-0.32	0.02	76	0.13	0.25	28	-0.20	0.24	33	0.16	0.33
Attrition proportion	42	0.16	0.32	58	-0.19	0.13	20	0.13	0.54	26	0.14	0.45
Other Method Variables												
Assignment level	60			76			28			33		
Student vs. other level		0.21	0.17		0.20	0.07		0.32	0.06		0.23	0.16
School vs. other level		0.04	0.81		0.08	0.47		-0.24	0.17		-0.20	0.22
Sample drawn from multiple schools	60	0.08	0.57	75	-0.08	0.46	27	-0.36	0.03	33	-0.52	0.00
Alternative intervention provided to comparison group (vs. business as usual or no treatment)	60	-0.11	0.46	76	-0.07	0.54	28	0.23	0.19	33	0.54	0.00

Amount of estimation required to compute effect size	60	0.21	0.15	--	--	--	--	--	--	--	--	--
Publication status (published vs. unpublished)	60	0.25	0.08	76	-0.02	0.87	28	-0.10	0.55	33	-0.47	0.00
Measurement Variables												
Test genre	--	--	--	76			--	--	--	33		
Narrative vs. expository test					0.01	0.94					-0.05	0.76
Expository vs. unknown/mixed					0.05	0.66					0.23	0.20
Test format	--	--	--	76			28			33		
Constructed response format					0.18	0.11		0.11	0.55		-0.00	0.99
Multiple choice format								-0.05	0.79		0.08	0.63
Test measures single strategy (vs. composite of several strategies)	--	--	--	--	--	--	--	--	--	33	-0.15	0.36
Sample Characteristics												
Majority of sample drawn from lowest performing grades**	52	-0.31	0.05	70	-0.27	0.02	25	-0.31	0.09	26	-0.45	0.01
Majority of sample drawn from highest performing grades***	--	--	--	--	--	--	25	0.25	0.17	--	--	--
Most of sample is identified as LD/RD	60	0.04	0.80	70	0.25	0.02	28	0.35	0.03	33	0.23	0.17
Most of sample is identified as second language learners (SLL)	24	-0.14	0.52	30	-0.10	0.57	--	--	--	33	0.03	0.84
Most of sample is identified as struggling readers	38	-0.06	0.73	54	0.12	0.37	--	--	--	22	0.06	0.78
Most of sample is identified as low SES	17	0.31	0.28	23	0.23	0.24	--	--	--	7	0.13	0.75

*standardized beta obtained from weighted regression with single predictor variable

**4th/5th for standardized measures of comprehension; 5th/8th for non-standardized measures of comprehension; 6th for strategy knowledge; and 5th/6th for strategy use

***7th grade for strategy knowledge

Table 12: Weighted Regression Statistics for Control Variables (Standardized Measures of Reading Comprehension)

	Estimate	**SE**	**P-value**	**Beta**
Quality Indicators				
Intercept	1.03	0.21	0.00	0.00
Baseline equivalence	-0.03	0.06	0.67	-0.06
Fidelity measured	-0.29	0.15	0.05	-0.26
Pretest adjustments made	-0.58	0.23	0.01	-0.35
All Selected Control Variables				
Intercept	1.08	0.36	0.00	0.00
Student-level assignment	-0.00	0.20	0.98	-0.00
Baseline equivalence	-0.00	0.08	0.96	-0.01
Fidelity measured	-0.35	0.21	0.09	-0.30
Publication status	0.30	0.17	0.08	0.27
Pretest adjustments made	-0.62	0.30	0.04	-0.39
Amount of estimation needed	-0.25	0.37	0.50	-0.13
Majority of sample in grades 4/5	-0.31	0.18	0.07	-0.28
Retained Control Variables Only				
Intercept	0.97	0.26	0.00	0.00
Fidelity measured	-0.32	0.17	0.06	-0.27
Publication status	0.27	0.16	0.09	0.24
Pretest adjustments made	-0.54	0.22	0.01	-0.34
Majority of sample in grades 4/5	-0.29	0.15	0.06	-0.25

Table 13: Weighted Regression Statistics for Control Variables (Non-Standardized Measures of Reading Comprehension)

	Estimate	**SE**	**P-value**	**Beta**
Quality Indicators				
Intercept	0.83	0.18	0.00	0.00
Random assignment	-0.25	0.20	0.21	-0.14
Fidelity measured	-0.27	0.19	0.16	-0.15
All Selected Control Variables				
Intercept	0.79	0.22	0.00	0.00
Random assignment	-0.23	0.21	0.28	-0.13
Student-level assignment	0.17	0.20	0.40	0.10
Fidelity measured	-0.39	0.18	0.03	-0.23
Majority of sample in grades 5/8	-0.52	0.19	0.01	-0.30
Most of sample with LD/RD	0.63	0.25	0.01	0.27
Constructed response format	0.29	0.20	0.14	0.17
Retained Control Variables Only				
Intercept	0.68	0.15	0.00	0.00
Fidelity measured	-0.44	0.17	0.01	-0.26
Majority of sample in grades 5/8	-0.53	0.18	0.00	-0.30
Most of sample with LD/RD	0.58	0.24	0.02	0.25
Constructed response format	0.40	0.18	0.02	0.23

Table 14: Weighted Regression Statistics for Control Variables (Measures of Strategy Knowledge)

	Estimate	**SE**	**P-value**	**Beta**
Quality Indicators				
Intercept	0.93	0.37	0.01	0.00
Random assignment	0.29	0.22	0.19	0.24
Fidelity measured	-0.30	0.21	0.15	-0.25
Pretest adjustments made	-0.30	0.32	0.34	-0.18
All Selected Control Variables				
Intercept	1.41	0.38	0.00	0.00
Random assignment	-0.04	0.25	0.87	-0.04
Sample is mostly LD/RD	0.38	0.37	0.31	0.24
Student-level assignment	-0.07	0.32	0.82	-0.07
Fidelity measured	-0.19	0.18	0.30	-0.19
Sample drawn from multiple schools	-0.21	0.28	0.46	-0.20
Comparison type (alternative intervention vs. others)	0.53	0.35	0.13	0.49
Pretest adjustments made	-0.67	0.44	0.13	-0.43
Sample is mostly 6^{th} graders	-0.85	0.30	0.00	-0.53
Sample is mostly 7^{th} graders	0.07	0.62	0.91	0.02
Retained Control Variables Only				
Intercept	1.38	0.25	0.00	0.00
Sample is mostly 6^{th} graders	-0.87	0.27	0.00	-0.55
Sample is mostly LD/RD	0.28	0.26	0.27	0.18
Fidelity measured	-0.19	0.16	0.22	-0.20
Sample drawn from multiple schools	-0.22	0.26	0.39	-0.23
Comparison type	0.44	0.25	0.07	0.43
Pretest adjustments made	-0.63	0.34	0.06	-0.40

Table 15: Weighted Regression Statistics for Control Variables (Measures of Strategy Use)

	Estimate	**SE**	**P-value**	**Beta**
Quality Indicators				
Intercept	0.80	0.23	0.00	0.00
Random Assignment	0.24	0.24	0.31	0.17
Fidelity measured	-0.27	0.24	0.25	-0.19
All Selected Control Variables				
Intercept	2.04	0.64	0.00	0.00
Random assignment	-0.36	0.26	0.17	-0.25
Student level assignment	0.34	0.40	0.40	0.14
Sample is drawn from multiple schools	-0.40	0.20	0.05	-0.29
Fidelity measured	-0.15	0.24	0.53	-0.10
Publication status	-0.75	0.60	0.21	-0.35
Comparison type (alternative intervention)	0.77	0.69	0.27	0.32
Test mostly includes expository texts	0.01	0.23	0.96	0.01
Most of sample identified as LD/RD	-0.02	0.31	0.96	-0.01
Majority of students from $5^{th}/6^{th}$ grade	-0.41	0.26	0.12	-0.28
Retained Control Variables Only				
Intercept	1.95	0.61	0.00	0.00
Random assignment	-0.28	0.25	0.25	-0.20
Sample is drawn from multiple schools	-0.39	0.19	0.04	-0.28
Publication status	-0.72	0.57	0.21	-0.33
Comparison type (alternative intervention)	0.87	0.66	0.19	0.36
Majority of students from 5th/6th grade	-0.47	0.25	0.06	-0.32

Table 16: Correlations of Hypothesized Moderator Variables with Adjusted SMD Effect Sizes, Accounting for Control Variables

	Standardized Measures of Reading Comprehension			*Non-Standardized Measures of Reading Comprehension*			*Strategy Knowledge*			*Strategy Use*		
Pedagogical Moderators	**k**	**R***	**P-value**	**k**	**R***	**P-value**	**k**	**R***	**P-value**	**k**	**R***	**P-value**
Treatment specificity	52	-0.02	0.88	70	0.10	0.34	24	0.03	0.88	26	0.11	0.40
Student practice included	52	-0.15	0.26	69	0.30	0.00	24	-0.01	0.96	26	-0.22	0.14
Instruction provided by teacher (vs. researcher or computer)	52	0.18	0.30	70	-0.22	0.05	24	0.06	0.78	26	-0.10	0.52
Instruction provided by computer (vs. researcher or teacher)	52	0.01	0.97	70	-0.15	0.15	24	0.20	0.21	26	0.15	0.29
Instruction provided in content area setting (vs. reading class)	52	-0.06	0.68	70	-0.06	0.58	24	-0.10	0.61	24	-0.18	0.15
Instruction provided as "pull-out" program (vs. during class or other)	52	0.07	0.62	70	0.27	0.01	24	-0.15	0.61	26	0.24	0.26
Instruction provided in whole class format (vs. small group or individual)	52	-0.03	0.82	67	0.00	0.95	24	0.12	0.62	26	-0.22	0.14
Instructional grouping is heterogeneous	50	-0.20	0.17	64	-0.04	0.75	23	0.12	0.69	24	0.00	0.98
Amount of collaboration during student practice	43	0.00	0.99	52	-0.09	0.45	20	-0.33	0.08	26	-0.25	0.04
Instruction includes narrative and expository texts (vs. single-genre instruction)	43	0.25	0.03	69	0.14	0.17	22	0.15	0.54	25	0.11	0.43
Instruction includes mostly expository texts (vs. narrative or mixed-genre)	43	-0.20	0.09	69	-0.14	0.20	23	-0.47	0.23	25	-0.11	0.43

Instruction includes mostly narrative texts (vs. expository or mixed-genre)	43	-0.13	0.29	69	0.01	0.91	23	0.01	0.95	--	--	--
Instruction mostly includes texts on assigned grade level	38	0.34	0.06	50	-0.03	0.82	16	-0.09	0.84	16	-0.14	0.46
Emphasis on self-regulated comprehension	51	0.00	0.99	70	0.16	0.12	23	0.15	0.38	26	0.07	0.71
Students taught to monitor strategy behavior	50	-0.27	0.04	70	0.10	0.31	24	-0.16	0.45	26	0.18	0.15
Duration of instruction	52	-0.05	0.72	70	0.21	0.05	23	-0.04	0.88	26	-0.01	0.94
Content Moderators												
Clarifying is taught as a strategy	50	-0.09	0.53	70	-0.11	0.28						
Graphic organizers	50	0.08	0.58	70	0.20	0.05						
Main idea	50	0.29	0.04	70	0.10	0.36						
Questioning	50	-0.14	0.30	70	0.00	0.95						
Inferring unstated information	50	-0.04	0.76	70	-0.06	0.58						
Comprehension monitoring	50	0.09	0.55	70	-0.04	0.70						
Paraphrasing	50	-0.08	0.57	70	-0.09	0.43						
Predicting	50	-0.04	0.80	70	-0.12	0.26						
Previewing	50	-0.06	0.66	70	0.23	0.04						
General fix-up strategies	50	0.13	0.40	70	0.02	0.88						
Setting purpose/goal for reading	50	0.06	0.67	70	0.07	0.51						
Summarizing	50	-0.04	0.80	70	-0.08	0.43						
Using prior knowledge	50	0.16	0.30	70	0.14	0.20						
Visualizing	50	0.14	0.33	70	0.00	0.99						
Strategies for analysis and reflection	50	0.25	0.08	70	0.27	0.01						
Big four strategies included	50	-0.44	0.00	70	-0.22	0.07						
Number of discrete strategies taught	50	0.13	0.35	70	0.10	0.34	23	0.14	0.55	26	0.05	0.72

*standardized β obtained from weighted regression with single predictor variable added to control variables.

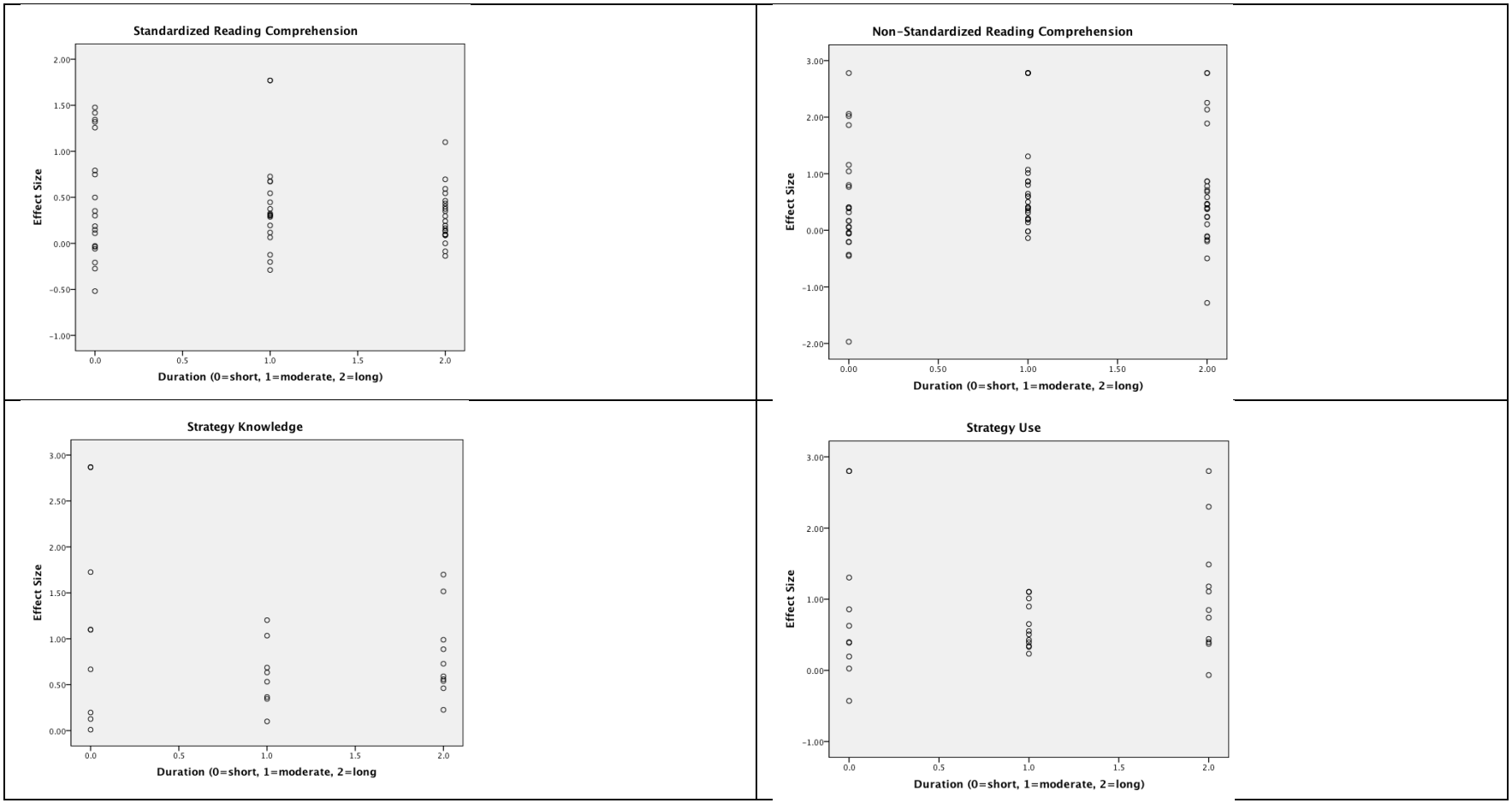

Figure 9: Scatterplots of Effect Size by Duration of Instruction

Table 17: Weighted Regression Statistics for Moderator and Control Variables Entered Simultaneously

	Estimate	SE	P-value	Beta
Standardized Reading Comprehension				
Intercept	0.76	0.37	0.04	0.00
Fidelity measured	-0.27	0.19	0.15	-0.24
Pretest adjustments made	-0.54	0.28	0.05	-0.38
Publication status	0.25	0.18	0.16	0.23
Majority of sample in grades 4/5	-0.32	0.18	0.08	-0.28
Student practice	0.22	0.20	0.28	0.15
Instruction provided by teacher	0.06	0.19	0.73	0.06
Heterogeneous grouping	-0.20	0.20	0.30	-0.17
Narrative and expository texts	0.05	0.18	0.79	0.04
Monitoring strategy behavior	-0.06	0.25	0.80	-0.03
Big four strategies*	-0.12	0.22	0.58	-0.11
Main idea	-0.02	0.25	0.94	-0.01
Using prior knowledge	0.14	0.21	0.49	0.12
Analysis and reflection	0.53	0.32	0.09	0.33
Non-Standardized Reading Comprehension				
Intercept	-0.03	0.25	0.91	0.00
Fidelity measured	-0.36	0.18	0.04	-0.22
Constructed response format	0.08	0.18	0.65	0.05
Majority of sample in grades 5/8	-0.51	0.16	0.00	-0.30
Majority of sample with LD/RD	0.25	0.24	0.29	0.11
Student practice included	0.39	0.21	0.07	0.18
Instruction provided by teacher	-0.08	0.21	0.70	-0.05
Instruction provided in pull-out format	0.51	0.21	0.02	0.28
Emphasis on self-regulated comprehension	0.01	0.08	0.86	0.02
Duration of instruction	0.17	0.13	0.19	0.17
Narrative and expository texts	-0.12	0.26	0.65	-0.06
Big four strategies*	-0.19	0.25	0.44	-0.10
Graphic organizers	0.46	0.22	0.03	0.24
Previewing	0.54	0.24	0.02	0.25
Analysis and reflection	0.97	0.37	0.01	0.29
Strategy Knowledge				
Intercept	1.44	0.30	0.00	0.00
Fidelity measured	-0.11	0.19	0.55	-0.11
Pretest adjustments made	-0.30	0.38	0.42	-0.20
Sample drawn from multiple schools	-0.52	0.30	0.08	-0.54
Comparison type	0.20	0.28	0.49	0.19
Sample is mostly 6th graders	-0.32	0.44	0.46	-0.21
Sample is mostly LD/RD	0.80	0.47	0.09	0.55
Instruction delivered by computer	1.23	1.00	0.22	0.21
Amount of collaboration	-0.19	0.12	0.10	-0.31

Instruction primarily includes expository texts	-0.46	0.41	0.27	-0.43
Strategy Use				
Intercept	2.28	0.71	0.00	0.00
Random assignment	-0.29	0.26	0.26	-0.20
Sample is drawn from multiple schools	-0.30	0.24	0.21	-0.22
Comparison type (alternative intervention)	0.54	0.80	0.50	0.22
Publication status	-0.87	0.59	0.14	-0.40
Majority of students from 5th/6th grade	-0.37	0.27	0.17	-0.26
Student practice included	-0.36	0.63	0.57	-0.19
Instruction provided in content area setting	0.02	0.27	0.95	0.01
Instruction provided in whole class format (versus small groups)	0.10	0.47	0.83	0.07
Instruction provided in pullout format	0.47	0.71	0.51	0.24
Amount of collaboration during student practice	-0.17	0.16	0.28	-0.19
Students taught to monitor strategy behavior	0.00	0.41	0.99	0.00

*Note: The regression output for this variable was obtained from a model that included the control variables plus the instructional variables, but not the individual strategy variables

CHAPTER V

DISCUSSION

This meta-analytic review includes intervention studies published between 1980 and 2009 in which students in grades 4-8 are taught to use two or more comprehension strategies. The collected studies were coded using a systematic data extraction scheme developed to address the central questions of the review. Information related to the characteristics of the student sample and instructional and methodological characteristics of each study were compiled in a database. Numerical effect sizes for each study for each major outcome measure were computed.

The mean effect of comprehension strategy instruction on each of the targeted outcome constructs was calculated to provide an overall summary of instructional effectiveness. Then, a series of moderator analyses were conducted to explain the sources of between-study variability for each outcome. Table 18 provides a visual depiction of how each moderator is related to treatment effectiveness for the four outcomes.

Overview of the Major Findings

In the sections that follow, the findings of this study are presented for each research question separately.

What is the average impact of MCSI on middle-grades student achievement?

The results of this meta-analysis confirm the conclusions reached in previous narrative and systematic reviews. Instruction in the use of multiple comprehension strategies has a positive impact on student achievement in grades 4 – 8. There is also some evidence that the impact of MCSI lasts beyond the intervention period.

The greatest impacts were observed for measures of strategy knowledge and strategy use. This is not surprising given that these skills are directly taught in MCSI. It is also not surprising that non-standardized measures of reading comprehension improved more than standardized measures, since the latter are generally more difficult to move than measures that are closer to the specific intervention content (NICHD, 2000; Rosenshine & Meister, 1994).

In addition to confirming that MCSI is associated with positive gains in achievement, the results of this study help clarify the size of these gains. At one time, researchers relied on a generic set of rules for interpreting the size of standardized treatment effects. An effect size of 0.80 or higher was considered large and anything in the 0.50 range was considered moderate (Cohen, 1988). However, these generic guidelines are often inadequate because they disregard the context in which the changes in a variable are observed. Hill, Bloom, Black, and Lipsey (2007) recommend an alternative approach in which an effect size is compared to the typical expected growth in a particular area. They estimate the expected growth on standardized measures of global reading ability to be 0.40 standard deviations as students move from fourth to fifth grade, 0.32 SD as they move from fifth to sixth grade, 0.23 SD from sixth to seventh, and 0.26 SD from seventh to eighth. Because tests of reading ability for students at this age tend to

focus heavily on comprehension, these estimates of academic growth can be used as benchmarks for interpreting the size of the treatment estimates obtained in the present study.

Using these benchmarks, the effect size estimate of 0.36 obtained in the current study for standardized measures of comprehension is fairly large, since an effect size of this magnitude constitutes approximately one year of growth in reading ability, and many of the reviewed studies lasted significantly less than a full year. Interpreting the magnitude of the summary effects for the remaining outcomes is more difficult because there are no empirical benchmarks to suggest how much growth is expected on these types of measures.

A word of caution is in order, however, regarding the magnitude of the summary effects obtained in this study. On average, the effect sizes obtained in more rigorously designed studies (i.e., studies that are better able to ensure equivalence of groups and/or control for group differences at baseline) are substantially smaller than those obtained in weaker studies. This is especially true for reading comprehension outcomes. The average impact of MCSI on both standardized and non-standardized measures is reduced by a factor of two when study quality is taken into account, while effects on strategy knowledge and strategy use appear to be more resistant to differences in study quality.

Does the impact of MCSI vary across studies?

Effect sizes for individual studies ranged from slightly negative to highly positive. The variability in effectiveness across studies is not surprising because there is a great deal of variation in study quality and instructional design. This is the reason given by the

members of the National Reading Panel to justify their choice not to calculate summary effects for text comprehension instruction. The findings of the present analysis support the claim that the expected impact of strategy instruction depends on how the intervention is implemented and how it is studied.

Is there evidence that some instructional frameworks are more effective than others?

A clear answer to this question cannot be reached because only a few of the most common MCSI frameworks have been repeatedly studied with middle grades students using experimental and quasi-experimental designs. A few tentative comments are warranted, however. First, this study shows that Reciprocal Teaching has a positive impact on reading comprehension. The average effect of RT on standardized measures obtained in this analysis (0.31) is in line with the median ES of 0.32 reported by Rosenshine and Meister (1994) over a decade ago.

A few of the other name-brand frameworks appear to be equally or more effective than RT at improving middle grades comprehension achievement. These include Peer-Assisted Learning Strategies, Think-aloud instruction, Transactional Strategies Instruction, and Concept-Oriented Reading Instruction. However, these effects are based on limited available evidence. Computer-assisted strategy instruction also appears to be effective. Finally, many of the non-branded frameworks are also associated with positive impact on reading comprehension, but the amount of variation in this category makes this summary effect difficult to interpret.

Is there evidence that some students benefit more from MCSI than others?

In general, positive impacts were observed across a range of student populations. There was some evidence that second language learners benefited less than others on standardized measures of comprehension, although they appear to benefit on other outcomes. Both struggling and non-struggling readers benefited from MCSI, with some evidence that below average readers benefited more on non-standardized measures than average or above average readers. There was also evidence that students with learning or reading disabilities benefited from MCSI.

Differences across grades for each construct were detected, but there were no clear patterns in the results to suggest that MCSI was especially beneficial or detrimental for any particular grade level.

Is there a relationship between instructional duration and effectiveness?

No strong systematic relationships were detected between instructional duration and effectiveness of treatment. There was some evidence that longer interventions were more effective for moving student achievement on non-standardized measures of comprehension, but the unique influence of duration was very small. This finding suggests that duration is not a major determinant of MCSI effectiveness, and that short-term and long-term interventions can be successful if implemented well.

How effective is MCSI when taught by classroom teachers, and how does this compare to the impact of MCSI taught by researchers and delivered by computer?

Interventions provided by teachers appear to be equally effective as those provided by researchers. This is an important finding because it suggests that classroom-based implementation of MCSI can be effective, given the right conditions and support.

Computerized instruction also appeared to be as effective as researcher or teacher-delivered instruction. In fact, there was a tendency for strategy knowledge outcomes to improve slightly more from computerized MCSI compared to instruction provided by live instructors. This finding lends support to recent efforts to build digital tutoring systems to enhance students' literacy development.

Does MCSI effectiveness differ by location?

There was no evidence that MCSI was more effective when provided during reading classes than in content area classes. Comprehension instruction can be effectively delivered alongside science and social studies content. There was some evidence that strategy instruction produced larger gains on non-standardized measures of reading comprehension when provided outside of the classroom, but in general, the findings suggest that MCSI can be effectively implemented as part of daily classroom practice.

What is the added benefit of student practice?

Although most MCSI interventions use a gradual release of responsibility instructional model, it is common for teachers to continue guiding students' strategy practice through the end of the intervention rather than providing opportunities for students to practice strategies on their own. The results of this study suggest that the

addition of student practice (without the teacher) is associated with positive gains in reading comprehension on both standardized and non-standardized measures but not for strategy knowledge and use. Although the added benefit of student practice is small, it warrants attention in future research as an important moderator of treatment effectiveness.

How does instructional grouping affect the impact of MCSI on student achievement?

There was no evidence that MCSI was more effective when provided in small groups versus whole classes. This again indicates that strategy instruction can be implemented effectively in classroom settings. There was some evidence that instruction was less effective in improving comprehension when students were grouped heterogeneously, although there is no explanation available for this trend. There may be some benefit to working with struggling and non-struggling readers separately; for example, it may facilitate individualization of instruction. This design element warrants additional study.

How does student collaboration affect the impact of MCSI on student achievement?

The amount of peer collaboration was not related to comprehension achievement for the studies included in this analysis. It was, however, negatively related to student performance on measures of strategy knowledge and use. This does not necessarily mean that collaboration is a negative instructional characteristic. It may be that increased amount of collaboration limits students' time to practice strategies individually, which

could hinder their performance on assessments that require students to demonstrate strategic expertise independently.

How does emphasis on self-regulated comprehension affect the impact of MCSI on student achievement?

There was no evidence that increased emphasis on self-regulated comprehension resulted in increased treatment effectiveness. This is a disappointing finding because a major hypothesis of this review was that providing opportunities for students to monitor their comprehension and make flexible decisions for strategy use would lead to greater achievement gains.

How does the genre of instructional materials affect the impact of MCSI?

In most cases, positive treatment effects were robust to differences in genre of instructional materials. One exception was the strategy knowledge outcome, which appeared to be hindered by interventions that focused exclusively on expository text. This suggests that some exposure to narrative material may help students develop declarative knowledge of strategic reading processes.

Is there a relationship between the number of strategies and impact?

Despite recent arguments that students might benefit more from deep proficiency with a few core strategies rather than exposure to many strategies (Block & Duffy, 2008), no systematic relationship was found between number of strategies and comprehension improvement.

Is there evidence that some strategies have more impact than others?

Some specific strategies were identified that appear to have a positive effect on comprehension. These include analysis/reflection, graphic organizers, and previewing. Notice these are not the most commonly taught strategies; in fact, analysis/reflection is one of the least frequently taught strategies in the MCSI literature. Also notable is the tendency for studies that included all four RT strategies (summarizing, clarifying, questioning, and predicting) to have lower effect sizes compared to studies that did not use all four of these strategies. This relationship does not hold after accounting for other instructional characteristics, but still warrants additional study.

Limitations of the Findings

Before discussing the implications of these findings for practice and research, a few limitations are worth mentioning. First, all meta-analyses are susceptible to bias resulting from study sampling. Careful and systematic procedures were used to locate and screen eligible studies, but there is no way of ensuring that every eligible study has been included.

A second limitation of this study is that inadequate reporting practices in the primary research literature made it difficult to reliably code some of the characteristics that were originally hypothesized as important predictors of treatment effectiveness. Some of these characteristics were eliminated from the coding protocol early on when their incidence of poor reporting became obvious. Other characteristics were retained, but their consistency of coding may have suffered from unclear reporting. To avoid this problem in the future, there is a real need for better reporting practices in the intervention

literature. This literature would benefit from more detailed descriptions of student samples in terms of grade-level distribution, ethnicity, linguistic heritage, and reading ability, and when possible, more frequent reporting of outcome data for separate groups of students. Furthermore, this literature would benefit from more detailed descriptions of the texts that were included during instruction and how they were selected.

Finally, it should be noted that the only feature that was experimentally manipulated in the collected studies was presence or absence of MCSI; therefore, the moderator analyses are purely exploratory. All characteristics that were found to enhance or detract from MCSI effectiveness will need to be systematically manipulated in future work to verify their causal relationship with treatment impact.

Implications for Practice

Findings from teacher development research indicate that strategy instruction is difficult and precise work (El-Dinary & Schuder, 1993). Teachers who effectively implement strategy instruction have deep understandings of the pedagogy—in particular, they can successfully manage the various decision-making points that arise while planning and implementing strategy lessons. This includes making decisions about which strategies to teach to which students, how to model and explain the strategies, and how much and in what contexts students need to practice the strategies (Anderson & Roit, 1993; Duffy, 1993). As explained at the beginning of this study, instantiations of MCSI vary considerably from study to study, producing an obfuscated knowledge base from which it has been difficult to isolate the essential components that teachers should know about and implement in their classrooms. The present analysis does not provide as much

guidance related to these components as was originally hoped. Many of the hypothesized moderating factors had no discernible relationship with MCSI effectiveness, suggesting that other unspecified factors warrant exploration.

The findings of this analysis do suggest a few recommendations that may help teachers as they plan and implement MCSI in their classrooms. The major implication of this meta-analysis for instructional practice is that multiple comprehension strategy instruction should be included as a key feature of middle grades literacy instruction. Even in the most rigorous studies, MCSI was found to positively impact student achievement on a variety of dimensions, including standardized measures of comprehension, which is no small feat. Although there is tremendous pedagogical variability in the MCSI literature, the fact that many of the characteristics that varied across studies did not have systematic relationships with treatment effectiveness suggests that MCSI can be implemented in various ways and still benefit students in grades 4-8.

Second, the findings of this study support the claim that teachers should consider strategies beyond the big four. Studies that included all four of the RT strategies were associated with positive gains in student achievement, but these gains were not noticeably larger than those obtained in studies that did not include these strategies. In fact, there is some evidence that MCSI is more effective when fewer of the RT strategies are used, but this finding does not hold after controlling for other instructional factors.

There is also evidence that teachers should teach students to preview texts before they begin reading, create graphic organizers as they read, and take time to analyze and reflect on their reading. While the findings do not provide clear answers for why these particular strategies are most helpful, one might speculate that graphic organizers and

previewing help mitigate the demands of content-area texts older students are expected to read. They help students anticipate text content, organize their thoughts, and provide a basis for recognizing and discussing conceptual relationships presented in a text (DiCecco & Gleason, 2002; Robinson, 1998). Strategies for analysis and reflection have the potential to help students acquire new knowledge from the texts they read, an important consideration for adolescent literacy instruction (Carnegie Council on Advancing Adolescent Literacy, 2010).

The tendency for these three strategies to produce higher effect sizes should not, however, be interpreted as license to focus exclusively on these strategies and ignore others. The presence or absence of most of the strategies included in the MCSI studies in this review had no discernible relationship with treatment effectiveness—either positive or negative—after accounting for other factors. This means that strategy instruction is generally effective regardless of the specific strategies that are included. The inclusion of previewing, graphic organizers, and higher-level analytical strategies was associated with additional effectiveness beyond the already positive impact of MCSI. Also note that these strategies always came bundled with other strategies, so their added value is on top of the effect of other strategies included in the instructional repertoires. This meta-analysis does not show that the use of previewing, graphic organizers, and analysis/reflection alone (without other strategies) is an effective instructional choice, since no studies were analyzed that taught these three strategies in isolation.

This review also provides evidence that teachers should create opportunities for students to practice strategies on their own without direct teacher intervention. To implement this recommendation, teachers will need to carefully scaffold their students'

strategic expertise through direct explanation, modeling, and guided strategy practice. Once students are ready, teachers should also be willing to temporarily fade into the crowd as students attempt to use strategies on their own—first in groups and then by themselves. This seems like an obvious recommendation, but it is warranted given that many versions of MCSI reviewed in this study fail to scaffold students all the way to independent practice, focusing instead on assisted practice with the teacher or collaborative practice with other students.

Implications for Research

The recommendations for future research that come out of this review are discussed below.

Purposeful Examinations of Under-Emphasized Factors

Several characteristics that were originally planned as moderators for this analysis could not be considered because they have been included in so few empirical studies. The first of these features that should be investigated is the possibility that MCSI is enhanced by selecting strategies to teach based on the needs of specific students. As explained in Chapter 2, only a few studies identified for this review made room for students' pre-existing strategies to become part of the curriculum. This trend precluded the possibility of examining the added benefit of emergent strategies. Examinations are needed that explore how and if an emphasis on emergent strategies might impact student achievement.

Certain features of the textual environment also warrant additional study due to lack of variability across studies. In particular, there is little discussion in this literature of students selecting books on their own, which limits the ability of this analysis to explore the role of text selection in moderating treatment effectiveness. There is also little discussion of how different choices of reading materials might impact instruction. Most studies included basal or basal-like materials selected by the researchers or teachers from readily accessible grade-level materials. Fewer studies included authentic materials such as trade books or websites that students might find intrinsically appealing. These features of the text environment deserve additional study.

Also, this review shows that independent reading is rarely prioritized as a crucial instructional component in the studies of strategy instruction for upper elementary and middle school students—despite the fact that "a great deal of time spent actually reading" (Duke & Pearson, 2002, p. 207) is often considered an important component of balanced literacy instruction (cf. Pressley, 2006). Since research on independent reading has yielded inconsistent results (NICHD, 2000), more work is needed that investigates the effectiveness of strategy instruction coupled with extensive independent reading practice.

Even though no relationship was observed between emphasis on self-regulated comprehension and treatment effectiveness, this is an area that deserves more attention. The majority of the studies primarily provide strategy practice in the form of instructional procedures that are heavily regulated by the teacher. This is evidenced by the low frequency of studies in which students were expected to set their own reading goals, select strategies flexibly, and monitor strategy effectiveness. Studies are needed that pay

explicit attention to the impact of these features of constructively responsive reading (Pressley et al., 1992).

Finally, it was not possible to examine the delayed impact of MCSI on student outcomes because this literature focuses primarily on immediate outcomes. Longitudinal work is warranted that examines how and if the effects of strategy instruction are maintained over time.

Systematic Manipulations of Moderating Factors

The moderating relationships identified in this study provide hypotheses that will need to be tested in future experimental and quasi-experimental work.

An important issue that will need to be addressed in future work is the relationship between student practice and treatment effectiveness. To verify that student practice is an essential component of MCSI, one might imagine a study that compares the impact of instructor directed and assisted practice with a condition that includes these two phases of instruction plus a carefully implemented student practice portion. If the results of this meta-analysis hold up, one would expect to see increased effectiveness when student practice is included. A related need is work that examines the relative benefits of peer collaboration and individual strategy practice during the student practice phase.

Another factor that will need to be addressed is the relationship of instructional duration to effectiveness. Duration is a particularly tricky characteristic to study because it is not simply a matter of dosage. Increased duration requires increased buy in from classroom teachers, closer collaborative relationships with school districts, and a host of other substantive factors that can themselves impact treatment effectiveness (R.

Anderson, personal communication). Furthermore, the relationship between duration and effectiveness may not be linear; it is entirely possible that after a certain amount of time, additional strategy lessons cease to be effective. This may especially be true when strategy-based instruction is implemented vertically across several grades—students who receive high quality MCSI in fourth-grade will likely need less of it in later grades, for example. In spite of these difficulties, duration warrants attention in future work because it is a practical concern. If teachers are expected to implement MCSI in their classrooms, they will need guidance regarding the amount of instruction needed to impact student achievement.

Additional examinations are also needed of the impact of strategy instruction provided by computers compared to instruction provided by teachers. A word of caution is in order here; the goal of these studies should not be to show that digital agents are more effective than live instructors, but to understand more clearly how teachers can use computerized agents as instructional tools to enhance their own practice.

Conclusion

This review provides a comprehensive summary of how multiple comprehension strategies instruction has been implemented and studied over the past three decades. The findings of the review provide strong empirical support for the use of MCSI to promote literacy achievement for students in grades 4 – 8. Also, practical recommendations can be taken from these findings to help educators plan and implement strategy instruction. Finally, the findings of this study provide directions for future research on reading comprehension pedagogy.

Table 18: Summary of Moderator Relationships

	Impact of MCSI on—			
	Reading Comprehension		Strategy Expertise	
	Standardized	Non-Standardized	Strategy Knowledge	Strategy Use
Specificity of intervention				
Duration of instruction		■		
Instructor = teacher				
Instructor = computer			■	
Location = pullout		■		
Location = content area class				
Student practice included	■	■		
Heterogeneous grouping	▨			
Group = whole class				
Amount of peer collaboration			▨	▨
Self-regulated comprehension				
Monitoring strategy behavior				
Text = expository			▨	
Text = assigned grade level				
Number of strategies				
Analysis and reflection	■	■		
Clarifying				
Graphic organizers		■		
Main idea				
Questioning				
Inferring unstated information				
Comprehension monitoring				
Paraphrasing				
Previewing		■		
General fix-up strategies				
Setting reading purpose				
Summarizing				
Using prior knowledge				
Visualizing				

Shaded bar = There is potentially a small positive relationship, after accounting for control variables and other moderators; dashed bar = There is potentially a small negative relationship; Blank = No relationship detected.

Appendix A

Report ID _____

Part 1/Sections A-E

A. Bibliographic Information

1	1st author	
2	2nd author	
3	3rd author	
4,5	Year	1-Journal _____ 2-Dissertation 3-Conference paper/report 4-Book chapter
6	Contrast # _____ / _____ from this report	

B. Instructional Framework Information – MCSI Treatment Condition (if two or more qualifying conditions are used, use separate sheets/ID numbers)

Characterizing the instructional framework		
7	Condition Name	
8	Name of Tx instruction	Name brand: 1-RT; 2-ISL; 3-PALS; 4-CSR; 5-ThAloud; 6-TSI; 7-CORI; 8-Other _____ Acronym: 9-POSSE; 10-PLAN; 11-TWA; 12-Other _____ 13-Digital name _____ 14-Generic name _____
9	Tx instructional program includes PD for instructors?	0-no 1-yes 9-nr
10	Tx instructional program provides—	1-general framework for thinking about strategic reading (e.g., an activity or way for students to use strategies, but not a clear plan for how to teach this activity) 2a-general curriculum for what gets taught in what order, without pre-packaged materials (e.g., there is a scope and sequence, but not a script, binder, or packaged set of materials) 2b-highly specified curriculum, with pre-written lessons and/or teacher script (e.g., all lesson plans are included in a binder) 4-either a general or specified curriculum, but can't be distinguished 9-nr

C. Instructional Framework Information – Comparison Condition

Characterizing the comparison condition		
12	Condition Name	
13	Nature of comparison instruction	0-na; not a group comparison study 1-no tx control (while tx group receives MCSI, control students receive nothing, recess, etc.) 2-business as usual (typical practice, with nothing extra added) 3-alternative PD (for teachers) 4-alternative intervention (for students) 9-nr
14	Content focus of comparison instruction	0-na; not a group comparison study 1-shares no/few characteristics with MCSI tx instruction (not focused on comprehension; other factors besides strategy instruction distinguish the two conditions) 2-shares some/many characteristics with MCSI tx instruction (comprehension focus; strategy focus is the predominant distinguishing factor) 3-is a single strategy framework 4-is an alternative MCSI framework (do not compute ES contrasts using this control) 9-nr

D. Content Characteristics (Treatment only)

Specific strategies taught			
18	Which strategies are taught? (circle all that apply) Note: only mark strategies mentioned as strategies by study authors	1-Analyzing and evaluating 2-Attending to graphics/pictures 3-Clarifying- word meaning (context) 4-Clarifying- text understanding (context) 5-Clarifying- general (context) (only when 3 or 4 is not clear) 6-Creating semantic/concept maps 7-Creating story maps 8-Drawing 9-Finding main ideas 10-Generating questions 11-Identifying author's purpose 12-Identifying genre 13-Inferring unstated info 14-Monitoring- words/phrases 15-Monitoring- text comprehension 17-Monitoring- text coherence	18-Paraphrasing/retelling (own words) 19-Predicting 20-Previewing 21-Reading ahead 22-Reflecting and relating 23-Regulating reading speed 24-Rereading 25-Searching for information 26-Setting goal/purpose for reading 27-Summarizing 28-Taking notes 29-Using prior knowledge 30-Using external resources (e.g., dictionary) 31-Using social resources (asking for help) 32-Using text structure/headings/organization 33-Visualizing/mental imaging Other:
19	Possibility for emergent strategies	0-none (all strategies preplanned) 1-some (most strategies preplanned, some bubbled up) 2-substantial (few preplanned, with intentional focus on emergent strategies) 9-nr	
Additional content related to self-regulated comprehension			
Does instruction explicitly emphasize the following?			
21	how to monitor one's own strategy use/behavior? (e.g., checking off strategies as they are used)	0-no 1-yes 9-nr	
22	how to choose strategies flexibly (e.g., choosing strategies for specific reasons)?	0-no 1-yes 9-nr	
23	how to monitor strategy effectiveness? (e.g., evaluating whether or not selected strategies are working)	0-no 1-yes 9-nr	

E. Pedagogical Characteristics

Duration			
24	tx spans __# days	or ____weeks _____ ____months nr ____school year	
25	#lessons total	_____ nr	
26	Min/lesson, average	_____ nr	
Instructional scaffolding - relative amount of instructor directed instruction (i.d), assisted practice (a.p.), and student practice (s.p.)			
Does instructional sequence include—			
27a	instructor-directed instruction (= the instructor does most of the work while students watch; e.g., introducing and/or modeling strategies)	0-no 1-yes 9-nr	
27b	assisted practice (= the instructor and students use strategies together; e.g., whole-class or small-group strategy discussions)	0-no 1-yes 9-nr	
27c	student practice (= students work indiv or collab without or mostly without the teacher)	0-no 1-yes 9-nr	
Instructional personnel			
30	Who is the primary instructor?	1-mostly researcher (mark 0 below for both items) 2-half researcher/half teacher (mark 0 below for both items) 3-mostly teacher 4-volunteer/tutor (answer #31; mark 0 below for #32) 5-mostly computer (mark 0 below for both items) 6-computer with one of the above _____ 9-nr	

31	Teacher agency	0-na 1-low agency (mostly planned/scripted by researcher) 2-some/high agency 9-nr	
32	Teacher specialty	0-na 1-content area teacher-gen ed. 2-reading/language teacher-gen ed. 3-special education teacher 4-reading/language specialist 9-nr	
Location			
33	Where does instruction primarily take place?	1-regular general ed classroom activities – reading class 2-regular special ed classroom activities – reading class (resource) 3-regular general ed classroom activities – content class 4-regular special ed classroom activities – content class (resource) 5-regular general ed classroom activities – hybrid reading/content class 6-regular special ed classroom activities – hybrid reading/content class (resource) 7-pull-out program (in school) 8-afterschool/enrichment (in school setting) 10-out of school setting/lab 99-nr	
Instructional grouping			
34,35	Grouping – instructor directed	0-na (not much i.d.) 1-individual (student interacts with teacher 1-to-1) 2-small group (2-6) 3-large group (7-12) 4-whole class (13+) 9-nr	If 2, 3, or 4— 0-na 1-same ability groups 2-mixed ability groups 3-tutor/tutee pairs 9-nr
36,37	Grouping – assisted practice	0-na (not much a.p.) 1-individual (student interacts with teacher 1-to-1) 2-small group (2-6) 3-large group (7-12) 4-whole class (13+) 9-nr	If 2, 3, or 4— 0-na 1-same ability groups 2-mixed ability groups 3-tutor/tutee pairs 9-nr
38,39	During student practice—	0-na (not much student practice) 1-students mostly work individually 2-equal mix of collaborative and individual work (e.g., collab first, then indiv.) 3-students mostly work collaboratively with peers 9-nr	If 2 or 3— 0-na 1-same ability a-pairs b-groups 2-mixed ability a-pairs b-groups 3-tutor/tutee pairs 9-nr
Nature of student practice			
During student practice phase, are students given opportunities to practice—			
52	setting goals/ microgoals for their reading?	0-no (teacher sets most goals) 1-yes 9-nr	
53	monitoring their compre while reading?	0-no 1-yes 9-nr	
54	selecting strategies flexibly?	0-no (teacher assigns/selects strategies) 1-yes 9-nr	
55a	monitoring their own strategy use/behavior?	0-no 1-yes 9-nr (Note: e.g., keeping a strategy checklist to mark off which strategies they have used)	
55b	evaluating the effectiveness of strategies?	0-no 1-yes 9-nr (Note: e.g., purposeful thinking about whether or not strategies are working)	
56a	verbally explaining their strategy use?	0-no 1-yes 9-nr (Note: in general, mark yes if students are discussing strategies together out loud)	
56b	explaining their strategy use in writing?	0-no 1-yes 9-nr	
Use of procedural facilitators			
57	Is a procedural facilitator used?	0-no (mark 0 below for both items) 1-yes 9-nr	

58	If yes, what type?	0-na 1-prompt sheet 2-mnemonic 3-both		
59	If yes, which statement is most true?	0-na 1-nothing fades away (students continue using the prompt/cue sheet for duration of instruction) 2-the tangible facilitator fades, but students continue using the inscribed strategy sequence 3-the procedural facilitator fades away completely (i.e., strategy use becomes more flexible over time) 9-nr		
Text environment				
60	Text genre	1-mostly narrative 2-narrative/expository mix 3-mostly expository 9-nr		
61,62, 63	Text mode	i.d. 0-na (not much t.d.) 1-mostly print 2-mostly print, some digital 3-mostly digital, some print 4-mostly digital 9-nr	a.p. 0-na (not much a.p.) 1-mostly print 2-mostly print, some digital 3-mostly digital, some print 4-mostly digital 9-nr	s.p. 0-na (not much s.p.) 1-mostly print 2-mostly print, some digital 3-mostly digital, some print 4-mostly digital 9-nr
64,65, 66	Text source	i.d. 0-na (not much t.d.) 1-mostly inauthentic 2-mostly basal/textbook 3-mostly authentic 9-nr	a.p. 0-na (not much a.p.) 1-mostly inauthentic 2-mostly basal/textbook 3-mostly authentic 9-nr	s.p. 0-na (not much s.p.) 1-mostly inauthentic 2-mostly basal/textbook 3-mostly authentic 9-nr
67	Students use self-selected texts?	0-none – instructor assigns texts 1-limited, but some 2-substantial 9-nr		
68	Students all read the same materials?	0-no – some evidence that texts are matched to readers 1-yes/mostly yes		
69,70	Text level	1-students mostly read texts below assigned grade level (text level < student grade level, for most students) 2-students mostly read texts at assigned grade level (text level = student grade level, for most students) 3-students mostly read texts above assigned grade level (text level > student grade level, for most students) 9-nr	1-students mostly read below ability level (text level < student reading level, for most) 2-students mostly read at ability level (text level = student reading level, for most) 3-students mostly read above ability level (text level > student reading level, for most) 9-nr	

Part II/Sections F-H

F. Student Characteristics

(if outcome data is presented separately by grade, ability level, etc., use separate sheet for each one, and make note here)

Notes:

		Treatment	*Comparison*	*Overall*
72	rawN (at assignment)			
73	If another N is used for these %, write it here—			
74	# in each grade	___4th ___5th ___6th ___7th ___8th ___other	___4th ___5th ___6th ___7th ___8th ___other	___4th ___5th ___6th ___7th ___8th ___other
75	% in each grade	___4th ___5th ___6th ___7th ___8th ___other	___4th ___5th ___6th ___7th ___8th ___other	___4th ___5th ___6th ___7th ___8th ___other
76	Pred grade level	4 6 8 5 7 0-other nr	4 6 8 5 7 0-other nr	4 6 8 5 7 0-other nr
77	# each race/ethn	___W ___AA ___H ___As ___other	___W ___AA ___H ___As ___other	___W ___AA ___H ___As ___other

78	% each race/ethn	___W ___AA ___H ___As ___other	___W ___AA ___H ___As ___other	___W ___AA ___H ___As ___other
79	Pred race/ethnicity	1-W 3-H 5-other 2-AA 4-As nr	1-W 3-H 5-other 2-AA 4-As nr	1-W 3-H 5-other 2-AA 4-As nr
80	# LD/RD/SPED			
81	% LD/RD/SPED			
82	# SLL			
83	% SLL			
84	# low SES/ “at risk”			
85	% low SES/ “at risk”			
86	# av/above av readers			
87	% av/above av readers			
88	# struggling readers*			
89	% struggling readers*			
90	# struggling comprehenders			
91	% struggling comprehenders			
92	# struggling decoders			
93	% struggling decoders			
94	# male			
95	% male			

*code this section if the authors do not break out by compr/decoding ability; **If specific numbers are not reported, estimate using these categories: *none*, *few* (<1/3), *half* (1/3-2/3), *most* (>2/3); if estimate is not possible, write *nr*

G. Research Design Characteristics

96	Study type	1-eligible group comparison (continue below) 2-ineligible group comparison (stop here; mark 0 for rest of section) 3-pre/post comparison (stop here; mark 0 for rest of section) 4-qualitative only (stop here; mark 0 for rest of section)		
97	Assignment type	0-na 1-nonrandom a-no effort made to equate groups b-some attempt to equate groups (e.g., matching) 2-random a-no matching b-with matching, group-wise c-with matching, individual 9-nr		
98	Statistical controls	0-no analytical strategies (statistical controls) 1-analytical strategies (statistical controls) used to account for potential group differences		
99	Assignment level	0-na 1-student 2-group smaller than class 3-class 4-school 9-nr		
100	Assg occurs within:	0-na 1-no blocking 2-groups smaller than class 3-class 4-schools 9-nr		
101	Groups equiv at baseline on compr?	0-no statistical comparison made/unclear results 1-equivalent 2-tx group favored 3-comparison group favored		
102	Groups equiv at baseline on strat knowledge?	0-no statistical comparison made/unclear results 1-equivalent 2-tx group favored 3-comparison group favored		
103	Groups equiv at baseline on strat use?	0-no statistical comparison made/unclear results 1-equivalent 2-tx group favored 3-comparison group favored		
104	Groups equiv at baseline on other reading measure? _____	0-no statistical comparison made/unclear results 1-equivalent 2-tx group favored 3-comparison group favored		
105	Groups equiv at baseline on other measure? _____	0-no statistical comparison made/unclear results 1-equivalent 2-tx group favored 3-comparison group favored		
106	Groups equiv at baseline on demographics?	0-no comparison made/unclear results 1-equivalent 2-not equivalent		
107	Implementation fidelity tracked?	0-no 1-yes a-author claims fidelity achieved b-author acknowledges problems in impl		
		Treatment	*Control*	*Overall*
108	Clustering scheme	(a)___groups of (b)_____ students (c)___classes of (d)_____ students (e)___schools of (f)_____ students	___groups of _____ students ___classes of _____ students ___schools of _____ students	___groups of _____ students ___classes of _____ students ___schools of _____ students

H. Impact Information for ES Calculations; use as many worksheets as needed

Summary of Extracted Data (complete worksheets in this order)

110	Reading comprehension measure?	IMM 0-no 1-yes; how many? _____ DEL 0-no 1-yes; how many? _____
111	Strategy knowledge measure?	IMM 0-no 1-yes; how many? _____ DEL 0-no 1-yes; how many? _____
112	Strategy application measure?	IMM 0-no 1-yes; how many? _____ DEL 0-no 1-yes; how many? _____
113	Comprehension monitoring measure?	0-no 1-yes
114	Reading attitude measure?	0-no 1-yes

Worksheet # ____ for Report ID _____ Authors _____ **Contrast # _____ Tx = _____ C=_____**							
A	Construct	1-reading comprehension 2-strategy knowledge 3-strategy application a-indiv strat, prompted b-mult strat, prompted c-student-init strategy use					
B	Text genre	1-mostly narrative 2-mix of narr and expos 3-mostly expository 9-nr 0na					
C	Test type	1-standardized/nationally or state normed 2-researcher-designed (designed for this study) 3-previously used measure, nonstand 9-nr					
D	Test format	1-multiple choice 2-constructed response/writing 3-observation/interview/think-aloud 9-nr					
E	Test name						
F	Pretest adj?	0-no (complete all columns below) 1-yes (omit pre column below)					
		Pre (G)Pretest reported? 0N 1Y (H) Same scale as post? 0N 1Y		*Post* (I) G1____days since tx ended		*Delayed Post* (J) G2____days since tx ended	
		TX	Comp	TX	Comp	TX	Comp
K	obsN						
L	Mean						
M	sd						
N	t-value						
O	p-value						
P	F-value						
Q	Other						
R	Attrition Index: TX (rawN-obsN)/rawN = C (rawN-obsN)/rawN =						
S	Notes						

Appendix B

Summary of Inter-Rater Reliability Statistics

Item	Exact Agreement	Cohen's Kappa	Correlation
Screening Decisions			
Dissertation abstract screening (omit or collect for full review)	11/12=92%	0.83	
Full study screening (eligible or ineligible for analysis)	11/12=92%	0.83	
MCSI condition selection	10/10=100%		
Comparison condition selection	9/10=90%		
Outcome measure selection	38/41=93%	0.76	
Descriptive Coding			
8-Framework type	18/20=90%	0.86	
9-Presence of professional development	18/20=90%	0.80	
10-Treatment specificity	17/20=85%	0.78	
13-Type of comparison condition	15/20=75%	0.43	
14-Content similarity of comparison condition	16/20=80%	0.62	
18-1-Analyzing and evaluating	19/20=95%	0.64	
18-2-attending to graphics or pictures	18/20=90%	0.00	
18-3-clarifying word meaning	17/20=85%	0.69	
18-4-clarifying text understanding	17/20=85%	0.35	
18-5-clarifying general	18/20=90%	0.61	
18-6-creating semantic maps or concept maps	19/20=95%	0.64	
18-7-creating story maps	19/20=95%	0.00	
18-8-drawing	19/20=95%	0.00	
18-9-finding main ideas	17/20=85%	0.63	
18-10-generating questions	18/20=90%	0.76	
18-11-identifying author's purpose	19/20=95%	0.64	
18-12-identifying genre	20/20=100%		
18-13-inferring unstated information	20/20=100%	1.00	
18-14-monitoring words or phrases	18/20=90%	0.00	
18-15-monitoring text comprehension	15/20=75%	0.48	
18-17-monitoring text coherence	20/20=100%		
18-18-paraphrasing or retelling	17/20=85%	0.50	
18-19-predicting	19/20=95%	0.90	
18-20-previewing	19/20=95%	0.77	
18-21-reading ahead	20/20=100%	1.00	
18-22-reflecting and relating	20/20=100%	1.00	
18-23-regulating reading speed	19/20=95%	0.64	
18-24-rereading	20/20=100%	1.00	
18-25-searching for information	20/20=100%	1.00	
18-26-setting goal or purpose for reading	19/20=95%	0.86	
18-27-summarizing	19/20=95%	0.85	
18-28-taking notes	20/20=100%		
18-29-using prior knowledge	14/20=70%	0.40	
18-30-using external resources	20/20=100%	1.0	

18-31-using social resources	20/20=100%		
18-32-using text structure, headings, etc.	19/20=95%	0.64	
18-33-visualizing	20/20=100%	1.00	
19-Opportunity for emergent strategies	20/20=100%		
21-Student taught to monitor strategy behavior	19/20=95%	0.77	
22-Students taught to choose strategies flexibly	12/20=60%	0.24	
23-Students taught to monitor strategy effectiveness	17/20=85%	0.50	
24-Treatment span in days	16/20 matches = 80%		0.99
25-Number of lessons	14/20 matches = 70%		0.99
26-Average number of minutes per lesson	14/20 matches = 70%		0.68
27c-Presence of student practice	18/20=90%	0.69	
30-Primary instructor	17/20=85%	0.79	
32-Teacher area/specialty	16/20=80%	0.61	
33-Location of instruction	15/20=75%	0.62	
34-Group size during instructor directed instruction	16/20=80%	0.67	
35-Group type during instructor directed instruction	15/20=75%	0.58	
36-Group size during assisted practice	13/20=65%	0.46	
37-Group type during assisted practice	14/20=70%	0.50	
39-Group type during student practice	14/20=70%	0.45	
52-Students set goals/purposes during student practice	19/20=95%	0.88	
53-Students monitor comprehension during student practice	15/20=75%	0.41	
54-Students select strategies flexibly (for specific purposes or difficulties) during student practice	14/20=70%	0.44	
55a-Students monitor strategy behavior during student practice	18/20=90%	0.62	
55b-Students evaluate strategy effectiveness during student practice	17/20=85%	0.50	
60-Genre of instructional materials	15/20=75%	0.61	
64-Source of materials during instructor directed practice	16/20=80%	0.65	
65-Source of materials during assisted practice	15/20=75%	0.51	
66-Source of materials during student practice	16/20=80%	0.58	
67-Use of self-selected materials	19/20=95%	0.83	
69-Level of instructional materials (relative to students' assigned grade	15/20=75%	0.63	

level)			
70-Level of instructional materials (relative to students' ability levels)	11/20=55%	0.31	
Sample Characteristics (n=18 studies)			
Tx raw N at assignment			1.0
Control raw N at assignment			1.0
Overall raw N at assignment			1.0
Tx # 4th graders			0.999
Tx perc of 4th graders			1.0
Tx # 5th graders			1.0
Tx perc of 5th graders			0.999
Tx # 6th graders			0.998
Tx perc of 6th graders			0.995
Tx # 7th graders			0.975
Tx perc of 7th graders			0.960
Tx # 8th graders			1.0
Tx perc of 8th graders			1.0
Tx # other graders			0.865
Tx perc of other graders			0.686
Tx pred grade			0.954
Control # 4th graders			1.0
Control perc of 4th graders			0.997
Control # 5th graders			1.0
Control perc of 5th graders			1.0
Control # 6th graders			0.999
Control perc of 6th graders			0.996
Control # 7th graders			0.963
Control perc of 7th graders			0.960
Control # 8th graders			1.0
Control perc of 8th graders			1.0
Control # other graders			0.830
Control perc of other graders			0.686
Control pred grade			0.983
Overall # 4th graders			1.0
Overall perc of 4th graders			1.0
Overall # 5th graders			1.0
Overall perc of 5th graders			1.0
Overall # 6th graders			1.0
Overall perc of 6th graders			1.0
Overall # 7th graders			0.972
Overall perc of 7th graders			0.960
Overall # 8th graders			1.0
Overall perc of 8th graders			1.0
Overall # other graders			0.839
Overall perc of other graders			0.686
Overall pred grade			0.984
Tx # White			0.962
Tx # African American			0.737
Tx # Hispanic			0.405
Tx # Asian			0.995
Tx # Other			0.878

Tx perc White	0.791
Tx perc African American	0.870
Tx perc Hispanic	0.534
Tx perc Asian	0.987
Tx perc Other	0.902
Control # White	0.968
Control # African American	0.909
Control # Hispanic	0.636
Control # Asian	0.992
Control # Other	0.632
Control perc White	0.879
Control perc African American	0.901
Control perc Hispanic	0.451
Control perc Asian	0.980
Control perc Other	0.926
Overall # White	0.957
Overall # African American	0.833
Overall # Hispanic	0.534
Overall # Asian	0.994
Overall # Other	0.983
Overall perc White	0.951
Overall perc African American	0.898
Overall perc Hispanic	0.689
Overall perc Asian	0.478
Overall perc Other	1.0
Tx # LD	1.0
Tx perc LD	1.0
Control # LD	0.965
Control perc LD	0.984
Overall # LD	0.992
Overall perc LD	0.996
Tx # SLL	1.0
Tx perc SLL	1.0
Control # SLL	1.0
Control perc SLL	1.0
Overall # SLL	1.0
Overall perc SLL	1.0
Tx # LOW SES	0.987
Tx perc LOW SES	0.971
Control # LOW SES	0.990
Control perc LOW SES	0.972
Overall # LOW SES	0.985
Overall perc LOW SES	0.969
Tx # AV/ABOVE AV READERS	0.994
Tx perc AV/ABOVE AV READERS	0.981
Control # AV/ABOVE AV READERS	0.992
Control perc AV/ABOVE AV READERS	0.981
Overall # AV/ABOVE AV READERS	0.991
Overall perc AV/ABOVE AV	0.981

READERS			
Tx # STRUGGLING READERS			0.983
Tx perc STRUGGLING READERS			0.973
Control # STRUGGLING READERS			0.984
Control perc STRUGGLING READERS			0.973
Overall # STRUGGLING READERS			0.986
Overall perc STRUGGLING READERS			0.977
Tx # MALE			0.998
Tx perc MALE			0.970
Control # MALE			0.994
Control perc MALE			0.968
Overall # MALE			1.0
Overall perc MALE			0.991
Design Information			
Assignment type	14/18=78%	0.65	
Use of statistical controls	17/18=94%	0.77	
Assignment level	16/18=89%	0.91	
Baseline equivalence-comprehension	16/18=89%	0.78	
Baseline equivalence- strategy knowledge	16/18=89%	0.72	
Baseline equivalence- strategy use	16/18=89%	0.73	
Baseline equivalence- other reading measure	12/18=67%	0.32	
Baseline equivalence- other measure	16/18=89%	0.73	
Baseline equivalence- demographics	15/18=83%	0.60	
Fidelity is measured	14/18=78%	0.61	
Tx # classes			0.961
Tx # in each class			0.991
Tx # schools			1.0
Tx # in each school			0.950
Control # classes			0.991
Control # in each class			0.985
Control # schools			0.999
Control # in each school			0.956
Overall # classes			0.986
Overall # in each class			0.995
Overall # schools			1.0
Overall # in each school			0.944
Outcome Data			
Construct	31/32=97%	0.95	
Genre	15/18=83%	0.70	
Test type	16/18=89%	0.81	
Format	18/18=100%	1.00	
Post test mean is already pretest adjusted	14/18=78%	0.55	
Pre-test reported	18/18=100%	1.0	
Pretest same scale	17/18=94%	0.82	

Tx pre N	1.0
Tx pre mean	1.0
Tx pre SD	1.0
Control pre N	1.0
Control pre mean	1.0
Control pre SD	0.998
Post # days	0.664
Tx post N	1.0
Tx post mean	0.999
Tx post SD	1.0
Control post N	1.0
Control post mean	1.0
Control post SD	1.0
Delayed # days	1.0
Tx delayed N	1.0
Tx delayed mean	1.0
Tx delayed SD	1.0
Control delayed N	1.0
Control delayed mean	1.0
Control delayed SD	1.0
Tx attrition at post	0.460
Control attrition at post	0.892

Appendix C

Meta-Analysis Coding Manual
Multiple Comprehension Strategy Instruction (MCSI) in Grades 4-8
Part I/Sections A-E

Section A – Bibliographic Information

General info: If a single study is reported across several reports, use the most recent and/or most comprehensive report as the major publication; if the study is reported in nonpublished (e.g., dissertation) and published form, use the published form as the major publication. Study information can be coded from both major and minor reports, but the major report should be used for coding section A.

#	Descriptor	Instructions/Examples
1	1st author	Enter the name of the first author of the major publication (last name, initials). Example: *Davis, D.S.*
2	2nd author	Enter the name of the second author of the major publication (last name, initials).
3	3rd author	Enter the name of the third author (last name, initials); if there are four or more authors, also enter […].
4	Year	Enter publication year of the major publication.
5	Publication type	Mark the publication type for the major publication. 1=journal; write name of journal on the line provided 2=dissertation (or thesis, etc.) 3=conference paper or technical report 4=book chapter
6	Contrast #	Enter the contrast number for this coding sheet and the total number of contrasts taken from this report. Reports have multiple contrasts if more than one qualifying instructional framework is tested against a single comparison framework. Example: enter *1 / 2* if there are two contrasts from this publication and this is the first one.

Section B: Instructional Framework Information – MCSI Treatment Condition

General info: Remember to use separate coding sheets for each contrast if multiple contrasts are reported for a single study.

#	Descriptor	Instructions/Examples
7	Tx Condition name	Enter the name of the MCSI instructional program, exactly as given in the report

8	Nature of tx instruction	Mark the most appropriate category for the nature/type of MCSI framework being used. (Note: these are all described in APPENDIX A) <u>Name-brand frameworks (#1-8)</u> (Note: a name-brand framework is a program with a recognizable name or a commercialized name) 1=Reciprocal Teaching 2=Informed Strategies for Learning 3=Peer Assisted Learning Strategies 4=Collaborative Strategic Reading 5=Think Aloud Instruction 6=Transactional Strategy Instruction 7=Concept Oriented Reading Instruction 8=The instructional framework has a name brand (not a strategy acronym), but it is not one of the ones listed here; also write the name on the line provided <u>Acronymized procedures (#9-12)</u> (Note: This is a specific type of name-brand program; these have names that are strategy acronyms) 9=P.O.S.S.E. 10=P.L.A.N. 11=T.W.A. 12=The instructional framework is named with a strategy acronym, but it is not one of the ones listed here; also write the name on the line provided 13=The instructional framework is a digital version of one of the above or a unique framework delivered mostly by computer; also write the name on the line provided; if it is a digital adaptation of another framework, write the # of that framework on the line as well. 14=The instructional framework has a generic name (example: *strategy instruction*); also write the name on the line provided

9	Presence of PD	0=the instructional program does not provide any sort of professional development for classroom teachers (example: the instruction is delivered by researchers, so teachers do not have to learn to do the instruction) 1=yes, the instructional program provides professional development to teachers specific to the MCSI framework being studied 9=not enough information is given in the study report to reasonably infer a 0 or 1 code
10	General or specific	1=the MCSI instruction provides a general framework for thinking about strategic reading, but participating instructors (teachers or researchers) are expected to develop their own lessons as they go; use this code when the developers of the curriculum provide an activity or series of activities for the students to complete (e.g., a reciprocal teaching activity) but do not provide a lesson sequence or packaged set of lessons for how to teach students to do this activity 2a=the MCSI instruction provides a general curriculum (e.g., guidelines for what gets taught, which might specify instructional order) before instruction begins; use this code when the developers begin with a pre-planned curriculum or lesson sequence, but the individual lessons are not scripted or planned out in advance (i.e., there is no binder of lesson plans and materials developed before instruction begins) 2b=the MCSI instruction provides a set of fairly rigid, pre-written lessons before instruction begins; use this code when the developers give the instructors a heavily specified set of lessons (which may include scripts) before instruction begins (i.e., there is a binder of lessons and materials) 4=it is definitely 2a or 2b, but there is not enough information to code one of these separately 9= not enough information is given in the study report to reasonably code this item

Section C: Instructional Framework Information – Comparison Condition

#	Descriptor	Instructions/Examples
12	Comparison Condition Name	Enter name of comparison condition exactly as given in the study report.
13	Nature of comparison condition	0=not applicable; this is not a group comparison study 1=the comparison serves as a "no treatment control"; mark this when the comparison group receives no attention or non-academic attention (e.g., recess) during the time when the MCSI group receives comprehension instruction 2=the comparison is a "business as usual" condition; mark this when comparison students receive typical practice with nothing extra added or changed 3=the comparison condition is an alternative professional development for teachers; mark this when the comparison teachers receive some type of non-MCSI training while treatment teachers receive MCSI training; this is likely the case when the intervention is framed as a PD experience for teachers 4=the comparison condition is an alternative intervention for students; mark this when the comparison students receive some non-MCSI academic treatment while the treatment students receive MCSI; this is likely the case when the intervention is delivered directly to students (i.e., the researcher delivers it instead of working with teachers) 9=not enough information is given in the study report to reasonably code this item
14	Content similarity	0=not applicable; this is not a group comparison study 1=the comparison condition shares no or few content characteristics with the MCSI tx condition; mark this when the comparison condition is not focused on reading comprehension and/or when other factors besides the strategy focus distinguish the two conditions 2=the comparison condition shares some or many characteristics with the MCSI condition; mark this when the comparison condition is focused on comprehension (or the same content information) and the only major difference between the two is the strategy emphasis 3=the comparison condition is a single strategy framework; mark this when the comparison condition includes strategy instruction, but only one strategy is taught 4=the comparison condition is an alternative MCSI framework 9=not enough information is given in the study report to reasonably code this item

Section D: Content Characteristics (MCSI Treatment Condition Only)

General notes: If multiple contrasts are included, this section will need to be coded separately for each MCSI condition; A single strategy can serve multiple purposes (example: summarization could help a student monitor and/or enhance comprehension, depending on how it is described by the authors)

#	Descriptor	Instructions/Examples
15	Monitoring	0=mark this when the authors do not describe comprehension monitoring as one of the purposes served by the strategies that are taught 1=the authors describe the strategies that are taught as serving a monitoring purpose; mark this when the instruction includes strategies for recognizing comprehension difficulties, identifying parts of the text that don't make sense, etc. (Note: this is a partially redundant code; as a general rule, you will only mark yes for this code when *monitoring* or *clarifying* are marked as strategies for item #18) REMOVED
16	Repairing	0=mark this when the authors do not describe comprehension repair (fix-up) as one of the purposes served by the strategies that are taught 1=the authors describe the strategies that are taught as serving a repair/fix up purpose; mark this when the instruction includes strategies for correcting or clarifying comprehension difficulties REMOVED
17	Enhancing	0=mark this when the authors do not describe comprehension enhancement as one of the purposes served by the strategies that are taught 1=the authors describe the strategies that are taught as enhancing comprehension, even when specific comprehension difficulties have not been identified; mark this when the instruction includes strategies for bolstering students' interactions, connections, etc. This code signifies that strategies are used even when students fully understand the text. REMOVED
18	Strategies taught	Circle the number of each strategy included in the tx instructional program; multiple strategies will be circled; when authors use an idiosyncratic strategy label, circle the label that best approximates the strategy they are describing; as a general rule, only mark strategies that the authors identify as strategies (e.g., if they list summarizing and questioning as the strategies they are teaching, but later in the article they provide an example of a graphic organizer, do not code graphic organizer as a strategy) (Note: the definition of strategy used to guide this coding is *a strategy is a mental tool a reader uses on purpose to aid comprehension*)

19	Emergent strategies	0=there is no possibility for new or revised strategies to emerge during instructional interactions; mark this when all the instructed strategies are preplanned before instruction begins 1=there is some possibility that new or revised strategies will emerge during instructional interactions; mark this when the strategies are mostly preplanned, but the authors acknowledge that new strategies might "bubble up" as the lessons progress 2=there is a substantial emphasis placed on emergent strategies; mark this when a few strategies are preplanned, but the authors explicitly look for students' pre-existing strategies or solicit new and revised strategies from students during the lessons 9=not enough information is given in the study report to reasonably code this item
21	Content_behavior monitoring	0=the instructional program does not explicitly emphasize how to monitor one's own strategy behavior 1=the instructional program does explicitly emphasize how to monitor one's own strategy behavior; use this code when students are taught/expected to check off the strategies they have used to make sure they have used them all; this is a form of behavioral monitoring 9=not enough information is given in the study report to reasonably code this item
22	Content_flexible coordination	0=the instructor does not explicitly explain/model/teach how to choose strategies flexibly; use this code when the instructor mainly teaches how to follow a strategy sequence that does not allow for flexible strategy use 1=the instructor does explicitly explain/model/teach how to choose strategies flexibly (i.e., that strategies have to be chosen for specific reasons) 9=not enough information is given in the study report to reasonably code this item
23	Content_strategy monitoring	0=the instructor does not explicitly explain/model/teach how to monitor whether or not selected strategies are effective (i.e., that you have to decide if a strategy is working) 1=the instructor does explicitly explain/model/teach how to monitor whether or not selected strategies are effective (i.e., that you have to decide if a strategy is working); only use this code when students are taught to think about or discuss whether or not their selected strategies are effective/helpful 9=not enough information is given in the study report to reasonably code this item

Section E: Pedagogical Characteristics

General notes: If multiple contrasts are included, this section will need to be coded separately for each MCSI condition

#	Descriptor	Instructions/Examples
24	span	Enter the number of school days the treatment spans; if span is reported in weeks, months, or years, enter the reported number on the appropriate line and then use the following to calculate days: 1 week=5 days; 1 month=20 days; 1 year=160 days; this is NOT necessarily the # of lessons or # of instructional days (example: if 20 lessons are spaced over 40 days, the span is 40 days); if a range is given, use the midpoint nr=not enough information is given in the study report to reasonably code this item
25	# lessons total	Enter the total number of lessons delivered to students; if lessons are reported in chunks, sum across the chunks (example: 3 intro lessons + 10 guided practice lessons + 10 independent practice lessons = 23 lessons total); if a range is given, use the midpoint; mark nr if only days/span is reported or if a reasonable estimate cannot be determined
26	Minutes per lesson	Enter the average length of each lesson in minutes; if lessons are reported in chunks with different lengths, calculate the average length to the nearest hundredth (example: two 7 minute lessons + one 21 minute lesson = 11.67 min/lesson on average); if a range is given, use the midpoint; mark nr if a reasonable estimate cannot be determined
27a	Instructor directed (i.d.)	*(Note: instructor-directed instruction, i.d. = the instructor does most of the work while students watch; e.g., introducing or modeling strategies)* 0=there is no point in the instructional sequence when the teacher/instructor does most of the strategy work 1=at some point in the instructional sequence, usually at the beginning, the teacher/instructor does most of the strategy work (e.g., provides examples or demonstrations while students watch/listen) 9=nr
27b	Assisted practice (a.p.)	*(Note: assisted practice, a.p. = the instructor and students work together to use strategies; e.g., whole-class or small-group strategy discussions)* 0=there is no point in the instructional sequence when students and teacher/instructor use strategies together 1=at some point in the instructional sequence, students and teacher/instructor use strategies together (e.g., the teacher guides students' strategy use with a shared text) 9=nr

27c	Student practice (s.p.)	*(Note: student practice, s.p. = students read and practice strategies individually or collaboratively, without teacher intervention; the teacher/instructor backs away completely or almost completely)* 0=there is no point in the instructional sequence when students use strategies without teacher assistance 1=at some point in the instructional sequence, students use strategies on their own, without teacher assistance (e.g., independent practice or peer-group practice without teacher sitting with the group) 9=nr
28a	i.d. vs a.p.	(Note: for items #28a-28c, mark one answer; no ties; only mark nr when the relative amount of the delivery methods cannot be inferred from the report; only mark 0 when neither method is used) 1=instructor-directed instruction is more prominent than assisted practice 2=assisted practice is more prominent than instructor-directed instruction 9=nr 0=neither is used
28b	i.d. vs s.p.	1= instructor-directed instruction is more prominent than student practice 2=student practice is more prominent than instructor-directed instruction 9=nr 0=neither is used
28c	a.p. vs s.p.	1=assisted practice is more prominent than student practice 2=student practice is more prominent than assisted practice 9=nr 0=neither is used

30	Instructor	1=instruction is primarily delivered by the researcher or research team (e.g., graduate assistants); also mark this when the researcher is the teacher (e.g., action research projects) 2=instruction is delivered jointly (about half and half) by the research team and a teacher or other school employee 3=instruction is primarily delivered by a teacher who is not the researcher (e.g., when the researcher trains the teachers to carry out the intervention) 4=instruction is primarily delivered by a volunteer or adult tutor who is not a teacher 5=instruction is primarily delivered by computer software or a digital pedagogical agent 6=instruction is delivered jointly by computer and a live instructor; also write the code that would be used for the live instructor on the line 9=not enough information is given in the study report to reasonably code this item
31	Teacher agency	(Note: this code only applies when codes 3, 4, or 6 are marked for #30 above; otherwise, mark 0-na) 0=not applicable; mark this when the instruction is delivered by the researcher or computer 1=teacher has low agency; mark this when the teacher has little control over the instruction (example: the lessons are mostly preplanned or scripted) 2=teacher has some agency or high agency; mark this when the teacher has some control over lesson design and implementation 9=not enough information is given in the study report to reasonably code this item
32	Teacher specialty	(Note: this code only applies when codes 3, 4, or 6 are marked for #30 above; otherwise, mark 0-na) 0=not applicable; mark this when the instruction is delivered by the researcher or computer 1=the instruction is delivered by a general education content area teacher (example: science or social studies teacher) 2=the instruction is delivered by a general education reading/language teacher 3=the instruction is delivered by a special education teacher 4=the instruction is delivered by a reading/language specialist (e.g., a speech pathologist or a pull-out reading specialist) 9=not enough information is given in the study report to reasonably code this item

33	Location	1=instruction primarily takes place during regular class activities in a general education reading class 2=instruction primarily takes place during regular class activities in a special education or resource reading class 3=instruction primarily takes place during regular class activities in a general education content area class (example: science or social studies class) 4=instruction primarily takes place during regular class activities in a special education or resource content area class (example: science or social studies class) 5=instruction primarily takes place during regular class activities in a general education hybrid reading/content class (example: science and reading taught together as part of an interdisciplinary unit) 6=instruction primarily takes place during regular class activities in a special education or resource hybrid reading/content class (example: science and reading taught together) 7=instruction primarily takes place in a pull-out program during school hours 8=instruction primarily takes place in an afterschool or enrichment program in the school setting (e.g., afterschool tutoring, summer school program) 10=instruction primarily takes place in an out-of-school or lab setting 99=not enough information is given in the study report to reasonably code this item
34	Grouping_instructor directed	0=not applicable; mark this when there is no instructor directed instruction (when you mark 0 for #27a above) 1=during instructor-directed instruction, an individual student interacts with the instructor in a one-to-one setting 2=during instructor-directed instruction, a small group of students (2-6) interact with the instructor 3=during instructor-directed instruction, a large group of students (7-12) interact with the instructor 4=during instructor-directed instruction, the whole class (13+) interacts with the instructor 9=not enough information is given in the study report to reasonably code this item; also use this if you coded 9 for #27a above

35	Grouping_id_ability	(Note: this code only applies if you marked 2, 3, or 4 for #34; otherwise, mark 0-na) 0=not applicable; instructor-directed instruction does not use grouping 1=when grouped for instructor-directed instruction, students are in same ability groups (homogenous with respect to reading comprehension or global reading ability) 2=when grouped for instructor-directed instruction, students are in mixed ability groups (heterogeneous with respect to reading comprehension or global reading ability) 3=when grouped for instructor-directed instruction, students are in tutor/tutee pairs (a higher achieving student is assigned to assist a lower achieving student; this is a special type of mixed ability grouping) 9=not enough information is given in the study report to reasonably code this item
36	Grouping_assisted practice	0=not applicable; mark this when there is no assisted practice (when you mark 0 for #27b above) 1=during assisted practice, an individual student interacts with the instructor in a one-to-one setting 2=during assisted practice, a small group of students (2-6) interact with the instructor 3=during assisted practice, a large group of students (7-12) interact with the instructor 4=during assisted practice, the whole class (13+) interacts with the instructor 9=not enough information is given in the study report to reasonably code this item; also use this if you coded 9 for #27b above
37	Grouping_ap_ability	(Note: this code only applies if you marked 2, 3, or 4 for #36; otherwise, mark 0-na) 0=not applicable; assisted practice does not use grouping 1=when grouped for assisted practice, students are in same ability groups (homogenous with respect to reading comprehension or global reading ability) 2=when grouped for assisted practice, students are in mixed ability groups (heterogeneous with respect to reading comprehension or global reading ability) 3=when grouped for assisted practice, students are in tutor/tutee pairs (a higher achieving student is assigned to assist a lower achieving student; this is a special type of mixed ability grouping) 9=not enough information is given in the study report to reasonably code this item

38	Grouping_student practice	0=not applicable; mark this when there is no student practice (when you mark 0 for #27c above) 1=during student practice, students mostly work individually 2=during student practice, there is a relatively equal mix of collaborative and individual work (example: students work in groups first then on their own) 3=during student practice, students mostly work collaboratively with peers (in partners or groups) 9=not enough information is given in the study report to reasonably code this item
39	Grouping_sp_ability	(Note: this code only applies if you marked 2 or 3 for #38; otherwise, mark 0-na) 0=not applicable; student practice does not use grouping 1a=when grouped for student practice, students are in same ability pairs (2 students; homogenous with respect to reading comprehension or global reading ability) 1b= when grouped for student practice, students are in same ability groups (3+ students; homogenous with respect to reading comprehension or global reading ability) 2a=when grouped for student practice, students are in mixed ability pairs (2 students; heterogeneous with respect to reading comprehension or global reading ability) 2b= when grouped for student practice, students are in mixed ability groups (3+ students; heterogeneous with respect to reading comprehension or global reading ability) 3=when grouped for student practice, students are in tutor/tutee pairs (a higher achieving student is assigned to assist a lower achieving student; this is a special type of mixed ability grouping) 9=not enough information is given in the study report to reasonably code this item
40	Direct explanation	0=the instructor does not directly explain strategies 1=the instructor does directly explain strategies (defines a strategy, how it works, how to use it) 9=not enough information is given in the study report to reasonably code this item REMOVED
41	Model strategies	0=the instructor does not demonstrate or model strategies 1=the instructor does demonstrate or model strategies (give an example of the strategy in use, think aloud, etc.) 9=not enough information is given in the study report to reasonably code this item REMOVED

42	Facilitate strategy discussion	0=the instructor does not facilitate or lead discussions using strategies 1=the instructor does facilitate or lead discussions of text using strategies (e.g., the teacher asks strategy questions/gives strategy prompts while reading a text together with the class) 9=not enough information is given in the study report to reasonably code this item REMOVED
43	Facilitate tangible products	0=the instructor does not facilitate the completion of tangible products; mark this when the instructor does not lead students through the completion of worksheets, graphic organizers, etc. 1=the instructor does facilitate the completion of tangible products 9=not enough information is given in the study report to reasonably code this item REMOVED
44	Expl vs. model	(Note: for items #44-49, mark one answer; no ties; only mark nr when neither descriptor can be inferred from the report; only mark 0 when neither method is used) 0=neither is used 1=the instructor does more direct explanation than modeling/demonstrating 2=the instructor does more modeling/demonstrating than direct explanation 9=not enough information is given in the study report to reasonably code this item REMOVED
45	Expl vs. facilt discussion	0=neither is used 1=the instructor does more direct explanation than facilitating strategy discussions 2=the instructor does more facilitating of strategy discussions than direct explanation 9=not enough information is given in the study report to reasonably code this item REMOVED
46	Expl vs. tang products	0=neither is used 1=the instructor does more direct explanation than facilitating completion of tangible products 2=the instructor does more facilitating of tangible products than direct explanation of strategies 9=not enough information is given in the study report to reasonably code this item REMOVED
47	Model vs. facilt discussion	0=neither is used 1=the instructor does more modeling/demonstrating than facilitating strategy discussions 2=the instructor does more facilitating of strategy discussion than modeling/demonstrating 9=not enough information is given in the study report to reasonably code this item REMOVED

48	Model vs. tang products	0=neither is used 1=the instructor does more modeling/demonstrating than facilitating completion of tangible products 2=the instructor does more facilitating of tangible products than modeling/demonstrating 9=not enough information is given in the study report to reasonably code this item REMOVED
49	Facilt discussion vs. tang products	0=neither is used 1=the instructor does more facilitating of strategy discussions than facilitating completion of tangible products 2=the instructor does more facilitating the completion of tangible products than facilitating strategy discussions 9=not enough information is given in the study report to reasonably code this item REMOVED
52	Opportunities for goal-setting	*Note: for items 52-56b, think primarily about the student practice portion of the instructional sequence (i.e., the portion of instruction where students are given time to practice strategies in their own reading; if there is no student practice, or very little, code this section thinking about the assisted practice portion)* 0=students are not given opportunities to set goals or microgoals for their reading; mark this when the teacher sets most of the goals or no goals are ever explicitly discussed 1=students are given opportunities to set goals or microgoals for their reading 9=not enough information is given in the study report to reasonably code this item
53	Opportunities to monitor comprehension	0=students are not given opportunities to monitor comprehension while reading (to think about whether or not what is being read makes sense) 1=students are given explicit opportunities to monitor comprehension while reading (to think about whether or not what is being read makes sense) 9=not enough information is given in the study report to reasonably code this item
54	Opportunities to select strategies flexibly	0=students are not given opportunities to select strategies based on specific needs/goals/reasons; mark this when the teacher assigns or prompts most strategies 1=students are given opportunities to select strategies based on specific needs/goals/reasons 9=not enough information is given in the study report to reasonably code this item

55a	Opportunities for mon strategy behavior	0=students are not given opportunities to monitor or evaluate their strategy behavior 1=students are given opportunities to monitor or evaluate their strategy behavior (e.g., to check off which strategies they have used or to count how many strategies they have used to make sure they have followed directions) 9=not enough information is given in the study report to reasonably code this item
55b	Opportunities for monitoring strategy effectiveness	0=students are not given opportunities to monitor or evaluate the usefulness of strategies they have selected 1=students are given opportunities to monitor or evaluate the usefulness of strategies they have selected 9=not enough information is given in the study report to reasonably code this item
56a	Opportunities to verbally explain strategy use	0=students are not given opportunities to verbally explain their use of strategies or their strategic reasoning 1=students are given opportunities to verbally explain their use of strategies or their strategic reasoning (e.g., students engage in think-alouds as part of their student practice; or they share their strategies with other students); use this code whenever students use strategies in collaboration with teacher/peers during guided or student practice (e.g., they say things like, “My prediction is…”) 9=not enough information is given in the study report to reasonably code this item
56b	Opportunities to explain strategy use in writing	0=students are not given opportunities to explain their use of strategies or their strategic reasoning in writing 1=students are given opportunities to explain their use of strategies or their strategic reasoning in writing during instructional time (not counting summative assessments) (e.g., students write out their think-alouds as part of their student practice) 9=not enough information is given in the study report to reasonably code this item
57	Procedural facilitator	(Note: a procedural facilitator is something that makes the process of using strategies easier, such as a cue sheet that reminds students of the strategies they should use or a mnemonic acronym that helps students remember strategies) 0=procedural facilitators are not used 1=procedural facilitators are used 9=not enough information is given in the study report to reasonably code this item

58	Proc facilitator_type	0=not applicable; mark this when a procedural facilitator is not used (when you mark 0 for #57 above) 1=a prompt sheet is used (example: a strategy poster on the wall or a strategy bookmark students keep at their desks) 2=a mnemonic is used (example: an acronym students memorize to help them remember the four strategies they should use) 3=both a prompt sheet and a mnemonic are used 9=not enough information is given in the study report to reasonably code this item
59	Proc facilitator_fading	0=not applicable; mark this when a procedural facilitator is not used (when you mark 0 for #57 above) 1=nothing fades away; mark this when the student is expected to continue using the facilitator for the remainder of instruction 2=the tangible aspects of the facilitator fade away before the end of instruction; mark this when the students stop using the written reminders, but they are expected to continue following the same strategy sequence 3=the procedural facilitator fades away completely before the end of instruction; mark this when the strategy sequence becomes more flexible over time 9=not enough information is given in the study report to reasonably code this item
60	Genre	1=the instructional program mostly uses narrative text 2=the instructional program uses a mix of narrative and expository text 3=the instructional program mostly uses expository text 9=not enough information is given in the study report to reasonably code this item
61	Mode_instructor directed	0=not applicable; mark this when there is no instructor-directed instruction 1=during instructor-directed instruction, students mostly interact with print texts 2=during instructor-directed instruction, students mostly interact print text, with some digital texts interspersed 3=during instructor-directed instruction, students interact mostly with digital text, with some print text mixed in 4=during instructor-directed instruction, students mostly interact with digital text 9=not enough information is given in the study report to reasonably code this item

62	Mode_assisted practice	0=not applicable; mark this when there is no assisted practice 1=during assisted practice, students mostly interact with print texts 2=during assisted practice, students mostly interact print text, with some digital texts interspersed 3=during assisted practice, students interact mostly with digital text, with some print text mixed in 4=during assisted practice, students mostly interact with digital text 9=not enough information is given in the study report to reasonably code this item
63	Mode_student practice	0=not applicable; mark this when there is no student practice 1=during student practice, students mostly interact with print texts 2=during student practice, students mostly interact print text, with some digital texts interspersed 3=during student practice, students interact mostly with digital text, with some print text mixed in 4=during student practice, students mostly interact with digital text 9=not enough information is given in the study report to reasonably code this item
64	Text source_instructor directed	0=not applicable; mark this when there is no instructor-directed instruction 1=during instructor-directed instruction, students primarily interact with inauthentic texts, like worksheets or specially designed passages 2=during instructor-directed instruction, students primarily interact with basal or textbook materials, or textbook-like materials (example: excerpts that are meant to approximate authentic reading materials) 3=during instructor-directed instruction, students primarily interact with authentic text, like trade books, magazines, or active websites 9=not enough information is given in the study report to reasonably code this item
65	Text source_assisted practice	0=not applicable; mark this when there is no assisted practice 1=during assisted practice, students primarily interact with inauthentic texts, like worksheets or specially designed passages 2=during assisted practice, students primarily interact with basal or textbook materials, or textbook-like materials (example: excerpts that are meant to approximate authentic reading materials) 3=during assisted practice, students primarily interact with authentic text, like trade books, magazines, or active websites 9=not enough information is given in the study report to reasonably code this item

66	Text source_student practice	0=not applicable; mark this when there is no student practice 1=during student practice, students primarily interact with inauthentic texts, like worksheets or specially designed passages 2=during student practice, students primarily interact with basal or textbook materials, or textbook-like materials (example: excerpts that are meant to approximate authentic reading materials) 3=during student practice, students primarily interact with authentic text, like trade books, magazines, or active websites 9=not enough information is given in the study report to reasonably code this item
67	Self-selected text	0=students do not use self-selected text; mark this when instructor chooses most reading materials 1=students have some limited access to self-selected text; mark this when most texts are chosen for students but there are some opportunities for individual choice 2=students' use of self-selected text is substantial; mark this when self-selection is heavily emphasized as part of the instructional program 9=not enough information is given in the study report to reasonably code this item
68	Reading same materials	0=students do not all read the same materials; mark this when there is at least some evidence that texts are matched to individual readers such that multiple texts are getting used in the classroom at the same time 1=students mostly read the same materials as their peers; no evidence that texts are matched to individual readers 9=not enough information is given in the study report to reasonably code this item
69	Text grade level	1=students mostly read texts below their assigned grade level (example: 5th grade students reading 3rd grade texts) 2=students mostly read texts at their assigned grade level; text level = assigned grade level for most students (or the average student) 3=students mostly read texts above their assigned grade level; text level is above assigned grade level for most students (or the average student) 9=not enough information is given in the study report to reasonably code this item

70	Text ability level	1=students mostly read texts below their ability levels; text level is below student reading ability level (e.g., 5th grade students who read at 5th grade level on average are primarily reading 3rd grade texts) 2=students mostly read texts at their ability levels; text readability level = student reading ability level for most students (or the average student); also use this when students read "just right" books 3=students mostly read texts above their ability levels (example: 5th grade struggling readers are assigned to read 5th grade materials) 9=not enough information is given in the study report to reasonably code this item

Section F – Student Characteristics

General info: tx=treatment; c=comparison condition; ov=overall (tx+c); when needed, enter NA for *not applicable* and NR for *not reported*; when possible, enter a reasonable estimate instead of entering NR; if specific numbers cannot be estimated, you can use these categories: none, few (<1/3), half (1/3-2/3), most (>2/3); when using these categories, use words instead of numerals to indicate that you are heavily estimating.

When outcome information is presented for subgroups separately, use separate sheets for each one; for example, if results are presented separately for fourth- and fifth-grade students, use a separate set of coding documents for each grade level.

If demographic information is reported at the school or district level only, enter this information in the *overall* column and mark NR for the *tx* and *c* columns.

#	Descriptor	Instructions/Examples
72tx	Tx raw N	Enter the raw sample size at time of assignment for the treatment group, or best estimate; this should be the student-level N; if N is reported at class or school level only, estimate number of students; unless the authors specify otherwise, assume 1 class = 25 students
72c	Comparison group raw N	Enter the raw sample size at time of assignment for the comparison group, or best estimate; this should be the student-level N; if N is reported at class or school level only, estimate number of students
72ov	Overall raw N	Enter the raw sample size at time of assignment for both groups combined, or best estimate; this should be the student-level N; if N is reported at class or school level only, estimate number of students
73tx	Tx other N	If another N is used for calculating percentages in later columns in this section, enter that number here for the treatment group (example: when the demographics are presented for the final N but not the raw N); if this is not applicable, enter NA

73c	Comparison group other N	If another N is used for calculating percentages in later columns in this section, enter that number here for the comparison group (example: when the demographics are presented for the final N but not the raw N); if this is not applicable, enter NA
73ov	Overall other N	If another N is used for calculating percentages in later columns in this section, enter that number here for the comparison group (example: when the demographics are presented for the final N but not the raw N); if this is not applicable, enter NA
74tx4	#4th graders in tx	Enter the number of xth grade students in the y group, or best estimate
74tx5	#5th graders in tx	"
74tx6	#6th graders in tx	"
74tx7	#7th graders in tx	"
74tx8	#8th graders in tx	"
74tx other	#other grades in tx	"
74c4	#4th graders in c	"
74c5	#5th graders in c	"
74c6	#6th graders in c	"
74c7	#7th graders in c	"
74c8	#8th graders in c	"
74c other	#other grades in c	"
74ov4	#4th graders overall	"
74ov5	#5th graders overall	"
74ov6	#6th graders overall	"
74ov7	#7th graders overall	"
74ov8	#8th graders overall	"
74ov other	#other grades overall	"
75tx4	%4th graders in tx	Enter the percentage of y group students who are xth graders; calculate this value by dividing the number of xth graders by the y group N and multiplying by 100
75tx5	%5th graders in tx	"
75tx6	%6th graders in tx	"
75tx7	%7th graders in tx	"
75tx8	%8th graders in tx	"

75tx other	%other grades in tx	“
75c4	%4th graders in c	“
75c5	%5th graders in c	“
75c6	%6th graders in c	“
75c7	%7th graders in c	“
75c8	%8th graders in c	“
75c other	%other grades in c	“
75ov4	%4th graders overall	“
75ov5	%5th graders overall	“
75ov6	%6th graders overall	“
75ov7	%7th graders overall	“
75ov8	%8th graders overall	“
75ov other	%other grades overall	“
76tx	Pred grade level in tx group	Circle the grade level that is most represented in this group (i.e., the grade level with the highest percentage); if there is an exact tie, circle all tying grades; write NR if this code cannot be estimated
76c	Pred grade level in c group	“
76ov	Pred grade level overall	“
77txW	# of White students in tx	Enter the number of students identified with this race/ethnicity label in the y group
77txAA	# of African American students in tx	“
77txH	# of Hispanic students in tx	“
77txAs	# of Asian students in tx	“
77txOther	# of students from other ethnicities in tx group	“
77cW	# of White students in c	“
77cAA	# of African American students in c	“

77cH	# of Hispanic students in c	“
77cAs	# of Asian students in c	“
77cOther	# of students from other ethnicities in c group	“
77ovW	# of White students overall	“
77ovAA	# of African American students overall	“
77ovH	# of Hispanic students overall	“
77ovAs	# of Asian students overall	“
77ovOther	# of students from other ethnicities overall	“
78txW	% of White students in tx	Enter the percentage of students in the y group who are identified with this race/ethnicity label; calculate this by dividing the # by the total N and multiplying by 100
78txAA	% of African American students in tx	“
78txH	% of Hispanic students in tx	“
78txAs	% of Asian students in tx	“
78txOther	% of students from other ethnicities in tx	“
78cW	% of White students in c	“
78cAA	% of African American students in c	“
78cH	% of Hispanic students in c	“
78cAs	% of Asian students in c	“
78cOther	% of students from other ethnicities in c	“

78ovW	% of White students overall	“
78ovAA	% of African American students overall	“
78ovH	% of Hispanic students overall	“
78ovAs	% of Asian students overall	“
78ovOther	% of students from other ethnicities overall	“
79tx	Predominant race/ ethnicity in tx	Circle the race/ethnicity that is most represented in this group (i.e., the one with the highest percentage); if there is an exact tie, circle all tying labels; select NR if this code cannot be estimated
79c	Predominant race/ ethnicity in c	“
79ov	Predominant race/ ethnicity overall	“
80tx	# of students in tx with LD/RD/SpEd designation	Enter the number of students in the y group who have been officially diagnosed/identified as having a learning disability, reading disability, or special education status; do not count students with ADHD or other behavioral/emotional designations in this category unless they also have an identified learning/reading disability
80c	# of students in c with LD/RD/SpEd designation	“
80ov	# of students overall with LD/RD/SpEd designation	“
81tx	% of students in tx with LD/RD/SpEd designation	Enter the percentage of students in the y group who have been officially diagnosed/ identified as having a learning disability, reading disability, or special education status; calculate this value by dividing the # by the total and multiplying by 100
81c	% of students in c with LD/RD/SpEd designation	“
81ov	% of students overall with LD/RD/SpEd designation	“

82tx	# of SLL students in tx	Enter the # of students in the y group who are learning to read/use strategies in a second language; this includes English language learners (ELLs) if the study is conducted in an English-speaking setting, but it could also include Mandarin speaking students learning to read in Spanish in a school in Spain, etc.
82c	# of SLL students in c	"
82ov	# of SLL students overall	"
83tx	% of SLL students in tx	Enter the percentage of students in the y group who are designated as second language learners; calculate this value by dividing the # by the total and multiplying by 100
83c	% of SLL students in c	"
83ov	% of SLL students overall	"
84tx	# low SES students in tx	Enter the number of students in the y group described as low SES, economically disadvantaged, at risk, or other generic labels indicating perceived potential for diminished school success
84c	# low SES students in c	"
84ov	# low SES students overall	"
85tx	% low SES students in tx	Enter the percentage of students in the y group described as low SES, economically disadvantaged, at risk, or other generic labels indicating perceived potential for diminished school success
85c	% low SES students in c	"
85ov	% low SES students overall	"
86tx	# average/above average readers in tx	Enter the number of students in the y group who are described as average or above average readers (i.e., nonstruggling readers), or best estimate; this refers to students who are average or above average in both decoding and comprehension
86c	# average/above average readers in c	"
86ov	# average/above average readers overall	"
87tx	% average/above average readers in tx group	Enter the percentage of students in the y group who are described as average or above average readers (i.e., nonstruggling readers)

87c	% average/above average readers in c group	“
87ov	% average/above average readers overall	“
88tx	# struggling readers in tx	Enter the number of students in the y group who are described as struggling readers (i.e., below grade level in reading); this category is a generic version of items 90 and 92; use this section when the authors do not provide descriptions of students’ ability levels that are broken out by decoding and comprehension
88c	# struggling readers in c	“
88ov	# struggling readers overall	“
89tx	% struggling readers in tx	Enter the percentage of students in the y group who are described as struggling readers (i.e., below grade level in reading)
89c	% struggling readers in c	“
89ov	% struggling readers overall	“
90tx	# struggling comprehenders in tx	Enter the number of students in the y group who are described as struggling comprehenders (i.e., below grade level in reading comprehension)
90c	# struggling comprehenders in c	“
90ov	# struggling comprehenders overall	“
91tx	% struggling comprehenders in tx	Enter the percentage of students in the y group who are described as struggling comprehenders (i.e., below grade level in reading comprehension)
91c	% struggling comprehenders in c	“
91ov	% struggling comprehenders overall	“
92tx	# struggling decoders in tx	Enter the number of students in the y group who are described as struggling decoders (i.e., below grade level in decoding/word recognition/fluency)
92c	# struggling decoders in c	“

92ov	# struggling decoders overall	“
93tx	% struggling decoders in tx	Enter the percentage of students in the y group who are described as struggling decoders (i.e., below grade level in decoding/word recognition/fluency)
93c	% struggling decoders in c	“
93ov	% struggling decoders overall	“
94tx	# males in tx	Enter the number of males in the y group
94c	# males in c	“
94ov	# males overall	“
95tx	% males in tx	Enter the percentage of males in the y group
96c	% males in c	“
96ov	% males overall	“

Section G: Research Design Characteristics

#	Descriptor	Instructions/Examples
96	Study type	1=eligible group comparison (at least one MCSI group is compared to at least one non-MCSI group); complete remaining codes in this section 2=ineligible group comparison (it is a quasi/experimental study, but there is no non-MCSI group); remainder of section is not applicable 3=pre/post or repeated measures comparison, without a comparison group; remainder of section is not applicable 4=qualitative design only; remainder of section is not applicable
97	Assignment type	0=na; use this when the study is not a group comparison study 1a=nonrandom assignment, with no effort made to equate groups a priori 1b=nonrandom assignment, with some attempt to equate groups (e.g., matching students or classrooms) 2a=simple random assignment, with no matching 2b=random assignment with matching, group-wise (e.g., classrooms are matched, then assigned) 2c=random assignment, with individual matching (e.g., students are matched, then assigned) 9=nr

98	Statistical controls	0=analytical strategies are not used to account for potential group differences 1=analytical strategies (e.g., statistical controls) are used to account for potential group differences
99	Assignment level	0=na 1=student level (individual students were assigned to groups) 2=groups smaller than class (e.g., reading groups, dyads) 3=class level (class is the unit of assignment) 4=school level (school is the unit of assignment) 9=nr
100	Assignment occurs within: (i.e., blocking)	*Note: this code often has to be inferred from the study design, since few researchers explicate their blocking scheme* 0=na 1=no blocking (individuals are assigned, with no clustering; e.g., in an out-of-school setting) 2=groups smaller than class (i.e., individuals are assigned, but blocked within pre-existing groups) 3=class (i.e., individuals are assigned, but within pre-existing classes) 4=school I.e., individuals or classes are assigned, but within schools)
101	Groups equiv at baseline on comprehension	0=na/no statistical comparisons made for reading comprehension, or unclear results 1=groups are equivalent at baseline on reading comprehension; use this when the researchers use an appropriate statistical test that shows nonsignificant differences before the intervention begins; also use this when the researchers report baseline data that could be used to confirm equivalence if the coder were to do the statistical test (e.g., when means, n, and standard deviations are reported for a pre-test, even if the researcher does not conduct the statistical comparison) 2=there is evidence that the tx group is higher on this construct 3=there is evidence that the comparison group is higher on this construct

102	Groups equiv at baseline on strategy knowledge	0=na/no statistical comparisons made for strategy knowledge, or unclear results 1=groups are equivalent at baseline on strategy knowledge; use this when the researchers use an appropriate statistical test that shows nonsignificant differences before the intervention begins; also use this when the researchers report baseline data that could be used to confirm equivalence if the coder were to do the statistical test (e.g., when means, n, and standard deviations are reported for a pre-test, even if the researcher does not conduct the statistical comparison) 2=there is evidence that the tx group is higher on this construct 3=there is evidence that the comparison group is higher on this construct
103	Groups equiv at baseline on strategy use	0=na/no statistical comparisons made for strategy use, or unclear results 1=groups are equivalent at baseline on strategy use; use this when the researchers use an appropriate statistical test that shows nonsignificant differences before the intervention begins; also use this when the researchers report baseline data that could be used to confirm equivalence if the coder were to do the statistical test (e.g., when means, n, and standard deviations are reported for a pre-test, even if the researcher does not conduct the statistical comparison) 2=there is evidence that the tx group is higher on this construct 3=there is evidence that the comparison group is higher on this construct
104	Groups equiv at baseline on other reading measure	*Note: if applicable, write construct name on the line provided* 0=na/no statistical comparisons made for another reading measure, or unclear results 1=groups are equivalent at baseline on another reading measure (e.g., fluency, decoding, global reading, reading attitude/motivation); use this when the researchers use an appropriate statistical test that shows nonsignificant differences before the intervention begins; also use this when the researchers report baseline data that could be used to confirm equivalence if the coder were to do the statistical test (e.g., when means, n, and standard deviations are reported for a pre-test, even if the researcher does not conduct the statistical comparison) 2=there is evidence that the tx group is higher on this construct 3=there is evidence that the comparison group is higher on this construct

105	Groups equiv at baseline on other measure	*Note: if applicable, write construct name on the line provided* 0=na/no statistical comparisons made for an additional measure, or unclear results 1=groups are equivalent at baseline on another measure (e.g., general motivation, self-efficacy, content knowledge); use this when the researchers use an appropriate statistical test that shows nonsignificant differences before the intervention begins; also use this when the researchers report baseline data that could be used to confirm equivalence if the coder were to do the statistical test (e.g., when means, n, and standard deviations are reported for a pre-test, even if the researcher does not conduct the statistical comparison) 2=there is evidence that the tx group is higher on this construct 3=there is evidence that the comparison group is higher on this construct
106	Groups equiv at baseline on demographics	0=na/no statistical comparisons made for demographic characteristics, or unclear results 1=groups are equivalent at baseline on demographic characteristics (e.g., equal representation of gender, SES status, ELL status, etc. across groups) 2=there is evidence that the groups are demographically nonequivalent
107	Implementation fidelity	0=implementation fidelity is not tracked/measured by the study authors 1a=implementation fidelity is tracked/measured by the study authors, who claim fidelity is mostly achieved 1b=implementation fidelity is tracked/measured by the study authors, who acknowledge problems in implementation or evidence of substantial slippage
108txA	#groups in tx	Enter the number of groups (smaller than class; e.g., reading groups) in the treatment group, if this type of grouping is used; if this type of grouping is not used, enter NA
108txB	#students per group in tx	Enter the number of students, on average, in each of the groups (smaller than class) in the treatment group; if this clustering is not used, enter NA; if exact numbers are presented for each group, enter each number as reported, with commas separating each one
108txC	#classes in tx	Enter the number of classes in the treatment group
108txD	#students per class in tx	Enter the number of students, on average, per class in the treatment group; if exact numbers are presented for each class, enter each number as reported, with commas separating each one
108txE	#schools in tx	Enter the number of schools in the treatment group
108txF	#students per school in tx	Enter the number of students, on average, in each school in the treatment group; if exact numbers are presented for each school, enter each number as reported, with commas separating each one

108cA	#groups in c	Enter the number of groups (smaller than class; e.g., reading groups) in the c group, if this type of grouping is used; if this type of grouping is not used, enter NA
108cB	#students per group in c	Enter the number of students, on average, in each of the groups (smaller than class) in the c group; if this clustering is not used, enter NA; if exact numbers are presented for each group, enter each number as reported, with commas separating each one
108cC	#classes in c	Enter the number of classes in the c group
108cD	#students per class in c	Enter the number of students, on average, per class in the c group; if exact numbers are presented for each class, enter each number as reported, with commas separating each one
108cE	#schools in c	Enter the number of schools in the c group
108cF	#students per school in c	Enter the number of students, on average, in each school in the c group; if exact numbers are presented for each school, enter each number as reported, with commas separating each one
108ovA	#groups overall	Enter the number of groups (smaller than class; e.g., reading groups) overall, if this type of grouping is used; if this type of grouping is not used, enter NA
108ovB	#students per group overall	Enter the number of students, on average, in each of the groups (smaller than class) overall; if this clustering is not used, enter NA; if exact numbers are presented for each group, enter each number as reported, with commas separating each one
108ovC	#classes overall	Enter the number of classes overall
108ovD	#students per class overall	Enter the number of students, on average, per class overall; if exact numbers are presented for each class, enter each number as reported, with commas separating each one
108ovE	#schools overall	Enter the number of schools overall
108ovF	#students per school overall	Enter the number of students, on average, in each school overall; if exact numbers are presented for each school, enter each number as reported, with commas separating each one

Section H: Impact Information for ES Calculations

General info: immediate posttest=the first measure after instruction ends; delayed posttest=the last measurement point reported in the study for a particular construct.

Reading comprehension measure = a test that measures a student's ability to construct a mental representation of a text; often, this is measured using a passage comprehension format (i.e., student reads a passage or series of passages and answers questions or responds to prompts); most studies will include at least one of these.

Strategy knowledge measure = a test that measures a student's ability to name, define, or explain strategies (declarative knowledge).

Strategy application measure = a test that measures a student's ability to use/apply strategies while reading; e.g., a performance assessment that asks students to read and annotate their strategies in the margins, or an interview protocol that asks students to read and describe their strategy use.

Comprehension monitoring measure = a test that measures a student's ability to recognize inconsistencies in a text (e.g., an error detection test).
Reading attitude measure = a test that measures a student's affect/motivation/attitude for reading (e.g., a questionnaire that asks students how much they enjoy reading or how often they read).

#	Descriptor	Instructions/Examples
110imm	Immediate compr measure	0=no immediate comprehension measure is reported 1=an immediate comprehension measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
110del	Delayed compr measure	0=no delayed comprehension measure is reported 1=a delayed comprehension measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
111imm	Immediate strategy knowledge measure	0=no immediate strategy knowledge measure is reported 1=an immediate strategy knowledge measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
111del	Delayed strategy knowledge measure	0=no delayed strategy knowledge measure is reported 1=a delayed strategy knowledge measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
112imm	Immediate strategy application measure	0=no immediate strategy application measure is reported 1=an immediate strategy application measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
112del	Delayed strategy application measure	0=no delayed strategy application measure is reported 1=a delayed strategy application measure is reported; enter number of eligible contrasts on the line; use a different worksheet for each one
113	Comprehension monitoring	0=no outcome measure reported 1=at least one outcome for this construct is reported; no ES worksheets are needed for this construct at this time
114	Reading attitude	0=no outcome measure reported 1=at least one outcome for this construct is reported; no ES worksheets are needed for this construct at this time

Effect Size Info Worksheet

General info: use as many worksheets as needed for each study; enter the worksheet number, report ID, study authors, and contrast # (a of b) for each one; also, enter the tx name and the comparison name for each worksheet

#	Descriptor	Instructions/Examples
A	Construct	1=this is a reading comprehension measure 2=this is a strategy knowledge measure 3a=this is a strategy application measure in which the student is prompted to use an individual strategy (e.g., a measure of questioning ability) 3b=this is a strategy application measure in which the student is prompted to use several strategies (e.g., a composite measure of questioning, summarizing, and predicting ability) 3c=this is a strategy application measure in which specific strategies are not directly prompted; instead, the student initiates strategies (e.g., a think-aloud protocol)
B	Text genre	1=the test mostly includes narrative text 2=the test includes a mix of narrative and expository texts (assume this category for most standardized measures, unless the authors specify otherwise) 3=the test mostly includes expository text 9=nr; use this when a reasonable estimate cannot be made 0=na; use this when the test does not include passages/texts (e.g., an interview in which the students are given a list of strategies to describe)
C	Test type	1=standardized/nationally or state normed test 2=researcher-designed test (designed for this study) 3=previously used measure (e.g., a nonstandardized measure used in a previous study; a measure that has been normed in previous work but only with smaller populations) 9=nr
D	Test format	1=the test is mostly multiple-choice format (or fill in the blank, etc); responses are provided with minimal explanation or elaboration 2=the test is mostly constructed response (the student has to produce extended answers in writing) 3=the test is mostly observational/interview/think-aloud 9=nr
E	Test name	Enter the name of the test, exactly as stated by the authors

F	Pretest adj	0=the authors do not report pre-test adjusted post-test data for this construct (enter pre- and post information separately below) 1=the authors report pre-test adjusted post-test data (i.e., adjusted means and SDs); use these values in the post-test columns below *Note: if the analyses are pretest adjusted, but the authors only report regression coefficients instead of adjusted post-test means and SDs, mark no for this item and code pre-test column below*
G	Pretest reported	0=the authors do not report sufficient pretest information for this construct 1=the authors do present sufficient pretest information for this construct
H	Pretest same scale	0=the pretest data is not reported on the same scale/measure as the post-test data 1=the pretest and posttest data are directly comparable (same measure/scale)
I	Timing of imm post-test	Enter the number of days between the end of instruction and the administration of the immediate measure; if there is no reason to assume there was a delay, enter 0.
J	Timing of delayed post-test	Enter the number of school days between the end of instruction and the administration of the delayed measure; if the delay is reported in weeks, calculate days by multiplying by 5
KpreTX	Observed N of the tx group at pretest	Enter the reported N (i.e., the N used for statistical analyses) for the appropriate group and testing administration
KpreC	Observed N of the c group at pretest	"
KpostTX	Observed N of the tx group at posttest	"
KpostC	Observed N of the c group at posttest	"
KdelTX	Observed N of the tx group at delayed posttest	"
KdelC	Observed N of the c group at delayed posttest	"
LpreTx	Tx group mean at pretest	Enter the reported group mean for the appropriate group and testing administration
LpreC	C group mean at pretest	"
LpostTX	Tx group mean at post-test	"
LpostC	C group mean at post-test	"

LdelTX	Tx group mean at delayed posttest	“
LdelC	C group mean at delayed posttest	“
MpreTx	Tx group SD at pretest	Enter the reported standard deviation for the appropriate group and testing administration; this should be a student-level SD; if SD is presented at aggregate level only, record this value and make a note indicating the level
MpreC	C group SD at pretest	“
MpostTX	Tx group SD at post-test	“
MpostC	C group SD at post-test	“
MdelTX	Tx group SD at delayed posttest	“
MdelC	C group SD at delayed posttest	“
N,O,P,Q	Other reported values	If means and standard deviations are not reported, enter the reported t-value, p-value, or F-value, Beta, etc. that best corresponds to the tx-comparison contrast for this construct; enter the value in the TX column for the appropriate test administrations
Rtx, Rc	Attrition index for tx	Using the formula on the coding document, calculate the proportion of students who left the each group between assignment and the immediate post-test results ([rawN-obsN]/rawN) ; if only the obsN is reported, write NR; if a separate index cannot be calculated for the tx and c groups, calculate an overall attrition index instead in the notes section: ([totalrawN=totalobsN]/totalrawN)

REFERENCES

[*]Included in the meta-analysis

[+]Included in the descriptive review

[*+]Aarnoutse, C., Brand-Gruwel, S., & Oduber, R. (1997). Improving reading comprehension strategies through listening. *Educational Studies*, 33(2), 209-227.

[*]Aarnoutse, C.A.J., Van den Bos, K.P., & Brand-Gruwel, S. (1998). Effects of listening comprehension training on listening and reading. *The Journal of Special Education,* 32(2), 115-126.

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69-90). New York: Routledge.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, 61(5), 364-373.

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review*, 10(2), 129-154.

[*]Anderberg, A. (1996). *The effects of reciprocal teaching techniques on reading comprehension for limited English proficient students* (Unpublished doctoral dissertation). University of Connecticut.

Anderson, V. (1991). A teacher development project in transactional strategy instruction for teachers of severely reading-disabled adolescents. Paper presented at the Annual Meeting of the National Reading Conference, Palm Springs, CA.

[+]Anderson, V. A., & Roit, M. (1993). Planning and implementing collaborative strategy instruction for delayed readers in Grades 6-10. *The Elementary School Journal*, 94(2), 121-137.

Atwell, N. (1987). *In the Middle: Writing, Reading, and Learning with Adolescents*. Upper Montclair, NJ: Boynton/Cook.

Atwell, N. (1998). *In the Middle: New Understandings About Writing, Reading, and Learning*. Portmouth, NH: Heinemann.

[*+]Babbs, P. J. (1984). Monitoring cards help improve comprehension. *The Reading Teacher,* 38(2), 200-204.

Baker, L., & Brown, A. L. (1984). Cognitive monitoring in reading. In J. Flood (Ed.), *Understanding reading comprehension: Cognition, language, and the structure of prose*. Newark, DE: International Reading Association.

Baker, L., Dreher, M. J., & Guthrie, J. T. (2000). Why teachers should promote reading engagement. In L. Baker, M. J. Dreher & J. T. Guthrie (Eds.), *Engaging young readers: Promoting achievement and motivation*. New York: Guilford Press.

Baumann, J. F. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly*, 20(1), 93- 115.

[+]Baumann, J. F., Hooten, H., & White, P. (1999). Teaching comprehension through literature: A teacher-research project to develop fifth graders' reading strategies and motivation. *The Reading Teacher*, 53(1), 38-51.

[*+]Baumann, J. F., Seifert-Kessell, N., & Jones, L. A. (1992). Effect of think-aloud instruction on elementary students' comprehension monitoring abilities. *Journal of Reading Behavior*, 24(2), 143-172.

[*+]Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction*, 2(2), 131-156.

[*]Berkeley, S. (2007). *Reading comprehension strategy instruction and attribution retraining for secondary students with disabilities* (Unpublished doctoral dissertation). George Mason University.

Biancarosa, G. (2005). After third grade. *Educational Leadership, 63*(2), 16-22.

Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

[*+]Block, C. C. (1993). Strategy instruction in a literature-based reading program. *The Elementary School Journal*, 94(2), 139-151.

Block, C. C., & Duffy, G. G. (2008). Research on teaching comprehension. In C. C. Block & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices* (Vol. 2). New York: Guilford Press.

Block, C.C. & Parris, S.R. (2008). *Comprehension instruction: Research-based best practices* (Vol. 2). New York: Guilford Press.

Borenstein, M. (2005). Software for publication bias (pp. 193-220). In H.R. Rothstein, A.J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex, England.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. West Sussex, UK: John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive Meta-analysis Version 2. Englewood, NJ: Biostat.

*Bottomley, D.M. (1993). *The effects of two procedures for teaching strategic reading on fourth-grade students' reading comprehension: Transactional strategies instruction and Reciprocal Teaching* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Bradford, L.E. (1991). *Metacognition and reading instruction: The effects of reciprocal teaching on reading comprehension of poor readers* (Unpublished doctoral dissertation). Indiana University.

*Brady, P.H. (1990). *Improving the reading comprehension of middle school students through reciprocal teaching and semantic mapping strategies* (Unpublished doctoral dissertation). University of Oregon.

Brand-Gruwel, S., Aarnoutse, C. A. J., & van den Bos, K. P. (1998). Improving text comprehension strategies in reading and listening settings. *Learning and Instruction,* 8(1), 63-81.

*Brimmer, K.M. (2004). *Using thinking aloud procedures to promote reading comprehension of expository texts with intermediate grade level students* (Unpublished doctoral dissertation). Oakland University.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior, 22*, 1-4.

Brown, A. L., & Palincsar, A. S. (1982). Inducing strategic learning from texts by means of informed, self-control training (Technical report no. 262). Champaign, IL: Center for the Study of Reading.

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(1), 18-37.

+Bruce, M. E., & Chan, L. K. S. (1991). Reciprocal teaching and transenvironmental programming: A program to facilitate the reading comprehension of students with reading difficulties. *Remedial and Special Education*, 12(5), 44-54.

*Burke, L. (1992). *The effects of an intervention program for improving reading comprehension and motivational attitudes with eighth grade social studies students* (Unpublished doctoral dissertation). University of Toledo.

*+Calhoon, M. B. (2005). Effects of a peer-mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities*, 38(5), 424-433.

Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York: Carnegie Corporation of New York.

*+Carr, M., & Borkowski, J. G. (1989). Attributional training and the generalzation of reading strategies with underachieving students. *Learning and Individual Differences*, 1(3), 327-341.

*+Carriedo, N., & Alonso-Tapia, J. (1995). Comprehension strategy training in content areas. *European Journal of Psychology of Education*, 10(4), 411-431.

Casteel, C. P., Isom, B. A., & Jordan, K. F. (2000). Creating confident and competent readers: Transactional strategies instruction. *Intervention in School and Clinic*, 36(2), 67-74.

Chall, J. S., Jacobs, V., & Baldwin, L. (1990). *The reading crisis*. Cambridge, MA: Harvard University Press.

*Chapman, M.L. (1997). *Instructing narrative text: Using children's concept of story with reciprocal teaching activities to foster story understanding and metacognition* (Unpublished doctoral dissertation). University of Michigan.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition) Hillsdale, NJ: Lawrence Erlbaum.

Collins, A., & Smith, E. E. (1982). Teaching the process of reading comprehension. In D. K. Detterman & R. J. Sternberg (Eds.), *How and how much can intelligence be increased* (pp. 173-185). Norwood, NJ: Ablex.

Collins, V. L., Dickson, S. V., Simmons, D. C., & Kame'enui, E. J. (1996). Metacognition and its relation to reading comprehension: A synthesis of the research: National Center to Improve the Tools of Educators.

Conley. (2008). Cognitive strategy instruction for adolescents: What we know about the promise, what we don't know about the potential. *Harvard Educational Review*, 78(1), 84-106.

Cook, T.D., Cooper, H., Cordray, D.S., Hartmann, H., Hedges, L.V., Light, R.J., Louis, T.A., & Mosteller, F. (1992). *Meta-analysis for explanation: A casebook.* New York: Russell Sage.

[*]Dao, M.N.T.H. (1993). *An investigation into the application of the reciprocal teaching procedure to enhance reading comprehension with educationally at-risk Vietnamese American pupils* (Unpublished doctoral dissertation). University of California-Berkeley and University of California-San Francisco.

Davey, B. (1983). Think aloud: Modeling the cognitive processes of reading comprehension. *Journal of Reading*, 27(1), 44-47.

[+]De Corte, E., Verschaffel, L., & Van De Den, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *British Journal of Educational Psychology*, 71, 531-559.

Dehn, M.J. (1992). *The effects of informed strategy training and computer mediated text on comprehension monitoring and reading comprehension* (Unpublished doctoral dissertation). University of South Dakota.

[*]Dermody, M.M. (1987). *Effects of metacognitive strategy training on fourth graders' reading comprehension* (Unpublished doctoral dissertation). University of New Orleans.

[*]Dermody, M.M. & Speaker, R.B. (1995). Effects of reciprocal strategy training in prediction, clarification, question generation, and summarization on fourth graders' reading comprehension. In K.A. Hinchman, D.J. Leu, & C.K. Kinzer (Eds.), Perspectives on Literacy Research and Practice: 44th Yearbook of the National Reading Conference (pp. 190-196). Chicago: The National Reading Conference.

DiCecco, V.M. & Gleason, M.M. (2002). Using graphic organizers to attain relational knowledge from expository text. *Journal of Learning Disabilities*, 35(4), 306-320.

[*+]Dole, J. A., Brown, K. J., & Trathen, W. (1996). The effects of strategy instruction on the comprehension performance of at-risk students. *Reading Research Quarterly*, 31(1), 62-88.

Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 347-372). New York: Routledge.

Duffy, G. (1992). Learning from the study of practice: Where we must go with strategy instruction. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Duffy, G. G. (1993). Rethinking strategy instruction; Four teachers' development and their low achievers' understandings. *The Elementary School Journal*, 93(3), 231-247.

Duffy, G. G., & Roehler, L. R. (1987). Teaching reading skills as strategies. *The Reading Teacher*, 40(4), 414-418.

*+Duffy, G. G., Roehler, L. R., Meloth, M. S., Vavrus, L. G., Book, C., Putnam, J., et al. (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), 237-251.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M. S., et al. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 22*(3), 347-368.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 205-242). Newark, DE: International Reading Association.

Durkin, D. (1978-1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, 14(4), 481-533.

Duval, S. (2005). The trim and fill method (pp. 127-144). In H.R. Rothstein, A.J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex, England.

*Ehlinger, J.M. (1988). *The relative merits of characteristics of teacher verbal modeling in influencing comprehension and comprehension monitoring of eighth grade readers* (Unpublished doctoral dissertation). Indiana University.

El-Dinary, P. B. (1994). The teaching practices of transactional-strategies-instruction teachers as revealed through collaborative interviewing: Reading research report no. 23: National Reading Research Center.

El-Dinary, P. B. (2002). Challenges of implementing transactional strategies instruction for reading comprehension. In C. C. Block & M. Pressley (Eds.), *Comprehension Instruction: Research-based Best Practices* (pp. 201-215). New York, NY: Guilford Press.

El-Dinary, P. B., & Schuder, T. (1993). Seven teachers' acceptance of transactional strategies instruction during their first year using it. *The Elementary School Journal*, 94(2), 207-219.

Ellis, E. S., & Graves, A. W. (1990). Teaching rural students with learning disabilities: A paraphrasing strategy to increase comprehension of main ideas. *Rural Special Education Quarterly, 10*, 2-10.

*+Englert, C. S., & Mariage, T. V. (1991). Making students partners in the comprehension process: Organizing the reading "POSSE". *Learning Disability Quarterly*, 14, 123-138.

*+Englert, C. S., Tarrant, K. L., Mariage, T. V., & Oxer, T. (1994). Lesson talk as the work of reading groups: The effectiveness of two interventions. *Journal of Learning Disabilities*, 27(3), 165-185.

*Esser, M.S. (2001). *The effects of metacognitive strategy training and attribution retraining on reading comprehension in African-American students with learning disabilities* (Unpublished doctoral dissertation). University of Wisconsin-Milwaukee.

*Etsey, K.F. (2004). *The effects of comprehension strategy instruction on Ghanaian English language learners' comprehension processes and text understanding* (Unpublished doctoral dissertation). University of Iowa.

Fisher, D., & Frey, N. (2008). What does it take to create skilled readers? Facilitating the transfer and application of literacy strategies. *Voices in the Middle*, 15(4), 16-22.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

+Franklin, M. R., Roach, P. B., Clary, J., E., & Ley, T. C. (1992). Overcoming the reading comprehension barriers of expository texts. *Educational Research Quarterly*, 16(1), 5-14.

Freebody, P. (2001). Theorising new literacies in and out of school. *Language and Education, 15*(2/3), 105-116.

*+Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34(1), 174-206.

*+Fuchs, L. S., Fuchs, D., Kazdan, S., & Allen, S. (1999). Effects of peer-assisted learning strategies in reading with and without training in elaborated help giving. *The Elementary School Journal*, 1999(3), 201-219.

+Fung, I. Y. Y., Wilkinson, A. G., & Moore, D. W. (2002). L1-assisted reciprocal teaching to improve ESL students' comprehension of English expository text. *Learning and Instruction*, 13, 1-31.

*Galbert, J.L.F. (1989). *An experimental study of reciprocal teaching of expository text with third, fourth, and fifth grade students enrolled in Chapter 1 reading* (Unpublished doctoral dissertation). Ball State University.

*Gee, E.J. (1997). *The interactive and combined effects of domain-specific knowledge and strategic knowledge on reading comprehension* (Unpublished doctoral dissertation). Utah State University.

Gee, J. P. (2001). Reading as situated language: A sociocognitive perspective. *Journal of Adolescent & Adult Literacy, 44*(8), 714-725.

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.

*+Gilroy, A., & Moore, D. W. (1988). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities with ten primary school girls. *Educational Psychology*, 8(1/2), 41-49.

*Glaeser, B.J.C. (1998). *The effects of an instructional model for improving reading comprehension achievement of students with learning disabilities, normally-achieving, at-risk, and gifted students in a multi-age, inclusive general education classroom* (Unpublished doctoral dissertation). University of Kansas.

Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8.

*Gordon, N.G. (1995). *Strategic reading instruction and reattribution training for students who are "at-risk" or learning disabled* (Unpublished thesis). Simon Fraser University.

Greenwood, C. R., & Delquadri, J. (1995). Classwide peer tutoring and the prevention of school failure. *Preventing School Failure*, 39(4), 21-25.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9). Retrieved May 1, 2009 from www.tcrecord.org

Guthrie, J. T. (2001). Contexts for engagement and motivation in reading [Electronic Version]. Reading Online, 4 from http://www.readingonline.org/articles/art_index.asp?HREF=/articles/handbook/guthrie/index.html

Guthrie, J. T. (2004). Teaching for literacy engagement. *Journal of Literacy Research*, 36(1), 1-30.

*+Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *The Elementary School Journal*, 99(4), 343-366.

Guthrie, J. T., & Cox, K. E. (2001). Classroom conditions for motivation and engagement in reading. *Educational Psychology Review*, 13(3), 283-302.

*+Guthrie, J. T., Van Meter, P., Hancock, G. R., Alao, S., Anderson, E., & McCann, A. (1998). Does Concept-Oriented Reading Instruction increase strategy use and conceptual learning from text? *Journal of Educational Psychology*, 90(2), 261-278.

*+Guthrie, J. T., Van Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., et al. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, 31(3), 306-332.

*+Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., et al. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology*, 96(3), 403-423.

*+Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, 92(2), 331-341.

Haggard, M. R. (1988). Developing critical thinking with the Directed Reading-Thinking Activity. *The Reading Teacher*, 41(6), 526-533.

*Hahn, A.L. (1983). *Training sixth graders to use general comprehension monitoring strategies for expository text processing* (Unpublished doctoral dissertation). University of Marlyand-College Park.

Haller, E.P., Child, D.A., & Walberg, H.J. (1988). Can comprehension be taught?: A quantitative synthesis of "metacognitive" studies. *Educational Researcher*, 17(9), 5-8.

Harris, K. R., & Graham, S. (1999). Programmatic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly, 22*(4), 251-262.

Harris, K. R., & Pressley, M. (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children*, 57(5), 392-404.

Harvey, S., & Goudvis, A. (2000). *Strategies that work: Teaching comprehension to enhance understanding*. Portland, ME: Stenhouse.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107-128.

Hedges, L.V. (2004a). Effect sizes in multi-site designs using assignment by cluster, unpublished manuscript.

Hedges, L.V. (2004b) Correcting significance tests for clustering, unpublished manuscript.

Hedges, L.V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341-370.

Hedges, L.V., & Hedberg, E.C. (2007). Instraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60-87.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

[*]Hedin, L.R. (2008). *The effects of thinking aloud on the comprehension and monitoring of sixth graders* (Unpublished doctoral dissertation). Univ of Illinois at Urbana-Champaign.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2007). *Empirical benchmarks for interpreting effect sizes in research*: MDRC Working Papers on Research Methodology.

Horner, S. L., & Shwery, C. S. (2002). Becoming an engaged, self-regulated reader. *Theory Into Practice*, 41(2), 103-109.

[*+]Houtveen, A. A. M., & van de Grift, W. J. C. M. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, 18(2), 173-190.

Idol, L., & Croll, V. J. (1987). Story-mapping training as a means of improving reading comprehension. *Learning Disability Quarterly*, 10, 214- 229.

Israel, S.E., & Duffy, G.G. (2009), *Handbook of research on reading comprehension*. New York: Routledge.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22(3, 4), 255-278.

[*]Jeffers, B.T. (1990). *A comparative study of metacognitive strategies in eighth grade reading improvement classes* (Unpublished doctoral dissertation). University of North Texas.

[+]Jiménez, R. T. (1997). The strategic reading abilities and potential of five low-literacy Latina/o readers in middle school. *Reading Research Quarterly*, 32(3), 224-243.

[+]Jiménez, R. T., & Gámez, A. (1996). Literature-based cognitive strategy instruction for middle school Latina/o students. *Journal of Adolescent and Adult Literacy*, 40(2), 84-91.

[*]Johnson, J.C. (2002). *Providing the toolbox: Reading comprehension strategy instruction to 4th grade students* (Unpublished doctoral dissertation). University of Minnesota.

[*+]Johnson-Glenberg, M. C. (2000). Training reading comprehension in adequate decoders/poor comprehenders: Verbal versus visual strategies. *Journal of Educational Psychology,* 92(4), 772-782.

[+]Johnson-Glenberg, M. C. (2005). Web-based training of metacognitive strategies for text comprehension: Focus on poor comprehenders. *Reading and Writing*, 18, 755-786.

Keene, E. O., & Zimmerman, S. (1997). *Mosaic of thought: Teaching comprehension in a reader's workshop*. Portsmouth, NH: Heinemann.

[*]Kelly, L.R. (1984). *Reading comprehension enhancement: A metacognitive generative strategies approach* (Unpublished doctoral dissertation). University of California-Berkeley.

[*+]Kelly, M., Moore, D. W., & Tuck, B. F. (1994). Reciprocal teaching in a regular primary school classroom. *Journal of Educational Research*, 88(1), 53-61.

[*+]Kim, A., Vaughn, S., Klingner, J. K., Woodruff, A. L., Reutebuch, C. K.,& Kouzekanani, K. (2006). Improving the reading comprehension of middle school students with disabilities through computer-assisted collaborative strategic reading. *Remedial and Special Education*, 27(4), 235-249.

Kim, J. S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis,* 28(4), 335-355.

[*+]Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction*, 5, 143-165.

[+]Kitano, M. K., & Lewis, R. B. (2007). Examining the relationship between reading achievement and tutoring duration and content for gifted culturally and linguistically diverse students from low-income backgrounds. *Journal for the Education of the Gifted*, 30(3), 295-325.

[+]Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *The Elementary School Journal*, 96(3), 275-293.

Klingner, J. K., & Vaughn, S. (2000). The helping behaviors of fifth graders while using collaborative strategic reading during ESL content classes. *TESOL Quarterly*, 34(1), 69-98.

[*+]Klingner, J. K., Vaughn, S., Arguelles, M. E., Hughes, M. T., & Leftwich, S. A. (2004). Collaborative strategic reading: "Real-world" lessons from classroom teachers. *Remedial and Special Education*, 25(5), 291-302.

[*+]Klingner, J. K., Vaughn, S., & Schrumm, J. S. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *The Elementary School Journal*, 99(1), 3-22.

[+]Kramarski, B., & Feldman, Y. (2000). Internet in the classroom: Effects on reading comprehension, motivation and metacognitive awareness. *Education Media International*, 37(3), 149-155.

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.

[*+]Labercane, G., & Battle, J. (1987). Cognitive processing strategies, self-esteem, and reading comprehension of learning disabled students. *B.C. Journal of Special Education*, 11(2), 167-185.

[*+]Lau, K. (2006). Implementing strategy instruction in Chinese language classes: A school-based Chines reading strategy instruction programme. *Educational Research,* 48(2), 195-209.

[*+]Lau, K., & Chan, D. W. (2007). The effects of cognitive strategy instruction on Chinese reading comprehension among Hong Kong low achieving students. *Reading and Writing*, 20, 833-857.

[*+]Lederer, J. M. (2000). Reciprocal teaching of social studies in inclusive elementary classrooms. *Journal of Learning Disabilities*, 33(1), 91-106.

Lee, J. Grigg, W., & Donahue, P. (2007). *The nation's report card: Reading 2007* (NCES 2007-496). Washington, DC: National Center for Education Statistics.

[+]Le Fevre, D. M., Moore, D. W., & Wilkinson, A. G. (2003). Tape-assisted reciprocal teaching: Cognitive bootstrapping for poor decoders. *British Journal of Educational Psychology*, 73, 37-58.

[*]Leu, D.J., Castek, J., Hartman, D.K., Coiro, J., Henry, L.A., Kulikowich, J.M., & Lyver, S. (2005). Evaluating the development of scientific knowledge and new forms of reading comprehension during online learning. Final Report submitted to the North Central Regional Educational Laboratory/Learning Point Associates.

Leu, D. J., Kinzer, C. K., Coiro, J. L., & Cammack, D. W. (2004). Toward a theory of new literacies emerging from the Internet and other information and communication technologies. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (Vol. 5). Newark, DE: International Reading Association.

[*]Levin, MC (1989). *An experimental investigation of Reciprocal Teaching and Informed Strategies for Learning taught to learning-disabled intermediate school students* (Unpublished doctoral dissertation). Columbia University Teachers College.

[*]Li, R. (1996). *Using the case method to improve reading comprehension of elementary students* (Unpublished doctoral dissertation). University of Virginia.

Lipsey, M.W. (2008). Group combo calculator. {Excel calculator].

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage.

[+]Little, Q., & Richards, R. T. (2000). Teaching learners--learners teaching: Using reciprocal teaching to improve comprehension strategies in challenged readers. *Reading Improvement*, 37(4), 190.

[*+]Loranger, A. L. (1997). Comprehension strategy instruction: Does it make a difference? *Reading Psychology*, 18(1), 31-68.

[*+]Lovett, M. W., Borden, S. L., Warren-Chaplin, P. M., Lacerenza, L., DeLuca, T., & Giovinazzo, R. (1996). Text comprehension training for disabled readers: An evaluation of reciprocal teaching and text analysis training programs. *Brain and Language*, 54, 447-480.

[*]Lubliner, S. (2001). *The effects of cognitive strategy instruction on students' reading comprehension* (Unpublished doctoral dissertation). University of San Francisco.

Lutz, S. L., Guthrie, J. T., & Davis, M. H. (2006). Scaffolding for engagement in elementary school reading instruction. *The Journal of Educational Research*, 100(1), 3-20.

[*+]Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1990). Reciprocal teaching improves standardized reading-comprehension performance in poor comprehenders. *The Elementary School Journal*, 90(5), 469-484.

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 663-679.

[+]Manset-Williamson, G., & Nelson, J. M. (2005). Balanced, strategic reading instruction for upper-elementary and middle school students with reading disabilities: A comparative study of two approaches. *Learning Disability Quarterly*, 28, 59-74.

[*+]Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology*, 96(2), 283-296.

[+]Mason, L. H., Snyder, K. H., Sukhram, D. P., & Kedem, Y. (2006). TWA + PLANS strategies for expository reading and writing: Effects for nine fourth-grade students. *Exceptional Children*, 73(1), 69-89.

Massey, D. D. (2009). Self-regulated comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 389-399). New York: Routledge.

Mastropieri, M.A., Scruggs, T.E., Bakken, J.P., & Whedon, C. (1996). Reading comprehension: A synthesis of research in learning disabilities. In T.E. Scruggs & M.A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities: Intervention research* (Volume 10, Part B, pp. 201- 227). Greenwich, CT: JAI.

[+]McCrudden, M. T., Perkins, P. G., & Putney, L. G. (2005). Self-efficacy and interest in the use of reading strategies. *Journal of Research in Childhood Education*, 20(2), 119-131.

[*]McElvain, C. (2005). *Transactional literature circles and the reading comprehension of at-risk English learners in the mainstream classroom* (Unpublished doctoral dissertation). University of San Francisco.

McHugh, C.M. (2004). Calculations for Correcting Test Statistics and Standard Errors for Clustering [Excel spreadsheet].

[*+]McKeown, M.G., Beck, I.L., & Blake, R.G.K. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, 44(3), 218-252.

[+]McKeown, R. G., & Gentilucci, J. L. (2007). Think-aloud strategy: Metacognitive development and monitoring comprehension in the middle school second-language classroom. *Journal of Adolescent and Adult Literacy*, 51(2), 136-147.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, 38(1), 1-30.

*+McNamara, D. S., O'Reilly, T. P., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, 34(2), 147-171.

Meichenbaum, D., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology*, 77(2), 115-126.

*+Miranda, A., Villaescusa, M. I., & Vidal-Abarca, E. (1997). Is attribution retraining necessary? Use of self-regulation procedures for enhancing the reading comprehension strategies of children with learning disabilities. *Journal of Learning Disabilities*, 30(5), 503-512.

Moore, A.R. (1993). *Effects of strategy training and classwide peer tutoring on the reading comprehension of students with learning disabilities* (Unpublished doctoral dissertation). Indiana University.

Moss, B. (2005). Making a case and a place for effective content area literacy instruction in the elementary grades. *The Reading Teacher*, 59(1), 46-55.

National Center for Education Statistics. (2007). *The nation's report card.* Retrieved August 6, 2009 from http://nationsreportcard.gov/reading%5F2007

National Council For Teachers of English. (1996). *NCTE/IRA standards for English language arts*. Retrieved February 9, 2009, from http://www.ncte.org/standards

National Institute of Child Health and Human Development (NICHD). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

*Nelson, E. (2003). *The effect of metacognitive strategy instruction on fifth grade comprehension of expository text* (Unpublished doctoral dissertation). University of Bridgeport.

+Nelson, J. M., & Manset-Williamson, G. (2006). The impact of explicit, self-regulatory reading comprehension strategy instruction on the reading-specific self-efficacy, attributions, and affect of students with reading disabilities. *Learning Disability Quarterly*, 29(3), 213-230.

*+Nolan, T. E. (1991). Self-questioning and prediction: Combining metacognitive strategies. *Journal of Reading*, 35(2), 132-138.

*Ocasio, T.L. (2006). *A comparison of two instructional programs to develop strategies to improve reading comprehension* (Unpublished doctoral dissertation).Widener University.

Ogle, D. M. (1986). K-W-L: A teaching model that develops active reading of expository text. *The Reading Teacher*, 39(6), 564-570.

[*]O'Hara, J.D. (2007). *The influence of supplemental instructional approaches upon the comprehension, metacognitive awareness and motivation of struggling third- and fourth-grade readers* (Unpublished doctoral dissertation). University of Maryland-College Park.

[+]Olson, C. B., & Land, R. (2007). Cognitive strategies approach to reading and writing instruction for English language learners in secondary school. *Research in the Teaching of English*, 41(3), 269-303.

[*]Padrón, Y.N. (1985). *Utilizing cognitive reading strategies to improve English reading comprehension of Spanish-speaking bilingual students* (Unpublished doctoral dissertation). University of Houston.

[*]Palincsar, A.S. (1987). *Collaborating for collaborative learning of text comprehension*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Palincsar, A. S. (2007). Reciprocal teaching from 1982 to 2006: The role of research, theory, and representation in the transformation of instructional research. *56th Yearbook of the National Reading Conference*, 41-52.

[*+]Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.

[+]Palincsar, A. S., Brown, A. L., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist*, 22(3/4), 231-253.

[*+]Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology*, 78(6), 1239-1252.

[+]Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development*, 55, 2083-2093.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.

[*+]Paris, S. G., & Oka, E. R. (1986a). Children's reading strategies, metacognition, and motivation. *Developmental Review*, 6, 25-56.

Paris, S. G., & Oka, E. R. (1986b). Self-regulated learning among exceptional children. *Exceptional Children*, 53(2), 103-108.

Paris, S. G., & Paris, A. H. (2001). Classroom application of research on self-regulated learning. *Educational Psychologist*, 36(2), 89-101.

*+Payne, B. D., & Manning, B. H. (1992). Basal reader instruction: Effects of comprehension monitoring training on reading comprehension, strategy use and attitude. *Reading Research and Instruction*, 32(1), 29-38.

Pearson, P.D. (2009). How we got to where we are: A historical perspective on reading comprehension research. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. New York: Routledge.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8(3), 317-344.

*Perez, D.R. (2002). *COMPRENDER: Comprehension: Orchestrating meaning and purpose for reading English among newly developing English language learners* (Unpublished doctoral dissertation). Kansas State University.

*Philbrick, A.R. (2002). *The effects of metacognitive reading instruction on students' comprehension of social studies text* (Unpublished doctoral dissertation). University of Arkansas.

*+Pickens, J., & McNaughton, S. (1988). Peer tutoring of comprehension strategies. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 8(1/2), 67-80.

*Piercy, T.D. (1997). *The effects of multi-strategy instruction upon reading comprehension* (Unpublished doctoral dissertation). University of Maryland.

Pressley, G. M. (1976). Mental imagery helps eight-year olds remember what they read. *Journal of Educational Psychology*, 68, 355-359.

Pressley, M. (2003). Metacognition and self-regulated comprehension. In A. Farstrup & J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 291-309). Newark, DE: International Reading Association.

Pressley, M. (2008). What the future of reading research could be. In C. C. Block & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices* (Vol. 2, pp. 391-413). New York: Guilford Press.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pressley, M., Brown, R., El-Dinary, P. B., & Afflerbach, P. (1995). The comprehension instruction that students need: Instruction fostering constructively responsive reading. *Learning Disabilities Research & Practice*, 10(4), 215-224.

Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., et al. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal*, 92(5), 513-555.

Pressley, M., Johnson, C. J., Symons, S., McGoldrick, J. A., & Kurita, J. A. (1989). Strategies that improve children's memory and comprehension of text. *The Elementary School Journal*, 90(1), 3-32.

Pressley, M., Wharton-McDonald, R., Mistretta-Hampton, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth- and fifth-grade classrooms in upstate New York. *Scientific Studies of Reading*, 2(2), 159-194.

[+]Proctor, C. P., Dalton, B., & Grisham, D. L. (2007). Scaffolding English language learners and struggling readers in a universal literacy environment with embedded strategy instruction and vocabulary support. *Journal of Literacy Research*, 39(1), 71-93.

[*+]Radcliffe, R., Caverly, D., Hand, J., & Franke, D. (2008). Improving reading in a middle school science classroom. *Journal of Adolescent and Adult Literacy*, 51(5), 398-408.

[+]Radcliffe, R., Caverly, D., Peterson, C., & Emmons, M. (2004). Improving textbook reading in a middle school science classroom. *Reading Improvement*, 41(3), 145-156.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: Office of Education Research and Improvement.

Raphael, T. E., George, M., Weber, C. M., & Nies, A. (2009). Approaches to teaching reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension*. New York: Routledge.

[*+]Rauenbusch, F., & Bereiter, C. (1991). Making reading more difficult: A degraded text microworld for teaching reading comprehension strategies. *Cognition and Instruction*, 8(2), 181-206.

Reis, S.M., Eckert, R.D., McCoach, D.B., Jacobs, J.K., & Coyne, M. (2008). Using enrichment reading practices to increase reading fluency, comprehension, and attitudes. *The Journal of Educational Research*, 101(5), 299-314.

[+]Rich, R. Z., & Blake, S. (1994). Collaborating for autonomy: Inducing strategic behaviors in students with learning disabilities. *Journal of Educational and Psychological Consultation*, 5(4), 359-372.

Robinson, D.H. (1998). Graphic organizers as aids to text learning. *Literacy Research and Instruction*, 37(2), 85-105.

[*+]Rogevich, M. E., & Perin, D. (2008). Effects on science summarization of a reading comprehension intervention for adolescents with behavior and attention disorders. *Exceptional Children*, 74(2), 135-154.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research*, 64, 479-530.

Rothstein, H.R., Sutton, A.J., & Borenstein, M. (2005). Publication bias in meta-analysis. In H.R. Rothstein, A.J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex, England.

[*]Russell, V.M. (1997). *Effects of reciprocal teaching on reading and oral language proficiency and reader self-perception of sixth-grade ESL students* (Unpublished doctoral dissertation). Fordham University.

[*+]Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, 71(3), 231-247.

[*+]Salomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a reading partner. *Journal of Educational Psychology*, 81(4), 620-627.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science*, 8, 173-190.

Schmitt, M.C. (1990). A questionnaire to measure children's awareness of strategic reading processes. *The Reading Teacher* 43(7), 454-458.

Schuder, T. (1993). The genesis of transactional strategies instruction in a reading program for at-risk students. *The Elementary School Journal*, 94(2), 183-200.

Sencibaugh, J.M. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement*, 44, 6-22.

[*]Shortland-Jones, B. (1986). *The development and testing of an instructional strategy for improving reading comprehension based on schema and metacognitive theories* (Unpublished doctoral dissertation). University of Oregon.

[+]Silvén, M., & Vauras, M. (1992). Improving reading through thinking aloud. *Learning and Instruction: The Journal of the European Association for Research on Learning and Instruction*, 2(2), 69-88.

Singer-Gabella, M. & Tiedemann, P. (2008). Tasks and trajectories in learning to teach. *Journal of Curriculum Studies*, 40(4), 457-480.

Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, 50 (2), 315-342.

[*]Smith, J.C. (2006). *The efficacy of explicit instruction of reading comprehension strategies to English language learners in elementary school* (Unpublished master's thesis). California State University-Long Beach.

[+]Soriano, M., Vidal-Abarca, E., & Miranda, A. (1996). Comparacion de dos procedimientos de instruccion en comprension y aprendizaje de textos: Instruccion directa y ensenanza reciproca. *Infancia y Aprendizaje*, 74(57-65).

[*+]Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16, 57-71.

Sterne, J.A.C. & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis (pp. 99-110). In H.R. Rothstein, A.J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 1-7). West Sussex, England.

Street, B. (2003). What's "new" in New Literacy Studies? Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education, 5*(2), 77-91.

[*+]Sung, Y., Chang, K., & Huang, J. (2008). Improving children's reading comprehension and use of strategies through computer-based strategy training. *Computers in Human Behavior*, 24, 1552-1571.

Swanson, (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities*, 32(6), 504-532.

[*+]Takala, M. (2006). The effects of reciprocal teaching on reading comprehension in mainstream and special (SLI) education. *Scandinavian Journal of Educational Research*, 50(5), 559-576.

Taylor, B.M., & Beach, R.W. (1984). The effects of text structure instruction on middle grades students' comprehension and production of expository text. *Reading Research Quarterly*, 19(2), 134-146.

*+Taylor, B., & Frye, B. J. (1992). Comprehension strategy instruction in the intermediate grades. *Reading Research and Instruction*, 32(1), 39-48.

Tennessee Department of Education. (n.d.). *Tennessee Curriculum Standards*. Retrieved December 9, 2008, from www.state.tn.us/education/ci/curriculum.shtml

Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., Stancavage, F., Durno, D., Javorsky, R., & Haan, C. (2007). *National assessment of Title I, final report: Volume II: Closing the reading gap: Findings from a randomized trial of four reading interventions for striving readers* (NCEE 2008-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

*+Tregaskes, M. R., & Daines, D. (1989). Effects of metacognitive strategies on reading comprehension. *Reading Research and Instruction*, 29(1), 52-60.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

*Uttero, D.A. (1992). *The effects of the instruction-modeling-cooperative engagement model on children's print comprehension in science* (Unpublished doctoral dissertation). University of Massachusetts-Lowell.

*+Van den Bos, K. P., Brand-Gruwel, S., & Aarnoutse, C. A. J. (1998). Text comprehension strategy instruction with poor readers. *Reading and Writing: An Interdisciplinary Journal,* 10, 471-498.

+Van Keer, H. (2004). Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology*, 74, 37-70.

+Van Keer, H., & Verhaeghe, J. P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *The Journal of Experimental Education*, 73(4), 291-329.

+Van Keer, H., & Verhaeghe, J. P. (2005). Comparing two teacher development programs for innovating reading comprehension instruction with regard to teachers' experiences and student outcomes. *Teaching and Teacher Education*, 21, 543-562.

[*]Villani, J.S. (1987). *The effect of instruction in reading strategies on self-perceptions of scholastic competence among middle-school students* (Unpublished doctoral dissertation). University of Maryland-College Park.

Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wagoner, S. A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, 18(3), 328-346.

[*+]Walraven, M., & Reitsma, P. (1993). The effect of teaching strategies for reading comprehension to poor readers and the possible surplus effect of activating prior knowledge. *National Reading Conference Yearbook*, 42, 243-250.

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*(3), 420-432.

[*+]Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McRae, A., et al. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. *Psychology in the Schools*, 45(5), 432-445.

Wilhelm, J. D. (2001). *Improving comprehension with think-aloud strategies*. New York: Scholastic.

Wilson, D. (1996). Effect Size Determination Program (Version 1.0) [Excel macro].

Wilson, D. (2005). Meta-Analysis Modified Weighted Multiple Regression (Version 2005.5.23) [SPSS macro].

Wilson, D. (2005). Mean Effect Size Calculator (Version 2005.5.23) [SPSS macro].

[*+]Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology*, 77, 197-235.

Wong, B. Y. L., & Jones, W. (1982). Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training. *Learning Disability Quarterly*, 5(3), 228-240.

[*]Woodward, D.C. (1991*). An investigation of the efficacy of Paris' Reading and Thinking Strategies on the reading comprehension and metacognition in reading in learning disabled students* (Unpublished thesis). Simon Fraser University.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3-17.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (Vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates.