DOCUMENT RESUME

ED 420 844 CS 013 236

AUTHOR	Chiu, Chris W. T.
TITLE	Synthesizing Metacognitive Interventions: What Training Characteristics Can Improve Reading Performance?
PUB DATE	1998-04-00
NOTE	45p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE	Information Analyses (070) -- Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Classroom Techniques; Elementary Secondary Education; Learning Strategies; Literature Reviews; Meta Analysis; *Metacognition; Reading Comprehension; *Reading Processes; *Reading Research; Research Methodology; Small Group Instruction
IDENTIFIERS	*Research Synthesis

ABSTRACT

In this paper, meta-analysis is used to identify components that are associated with effective metacognitive training programs in reading research. Forty-three studies, with an average of 81 students per study, were synthesized. It was found that metacognitive training could be more effectively implemented by using small-group instruction, as opposed to large-group instruction or one-to-one instruction. Less intensive programs were more effective than intensive programs. Program intensity was defined as the average number of days in a week that instruction was provided to students. Students in higher grades were more receptive to the intervention. Measurement artifacts, namely teaching to the test and use of nonstandardized tests and the quality of the studies synthesized played a significant role in the evaluation of the effectiveness of the metacognitive reading intervention. Appendixes contain ERIC keyword search; the coding instrument; coding instructions; interrater reliability; and formulas for the generalized least square regression coefficients and associated standard errors.) (Contains 1 figure, 4 tables of data, 55 references, and a list of 43 primary studies evaluated.)

Synthesizing Metacognitive Interventions:

What Training Characteristics Can Improve Reading Performance?

Chris W. T. Chiu

Michigan State University

1

Author Notes

A paper presented to the symposium, "Metacognition: Assessment & Training," at the Annual Meeting of the American Educational Research Association, in San Diego, California. April 13-17, 1998

Portions of this research were submitted to fulfill the requirement of an apprenticeship project for the degree of Doctor of Philosophy. Reprints of the symposium paper and the apprenticeship can be downloaded from the World Wide Web at http://pilot.msu.edu/~chiuwing/AERA98.doc and http://pilot.msu.edu/~chiuwing/WriteUp5.doc, respectively.

Correspondence concerning this manuscript should be addressed to the author at 1428 Haslett Rd, Apt 104, E. Lansing, MI 48823. Electronic mail may be sent via the Internet to chiuwing@pilot.msu.edu.

Abstract

In this paper, meta-analysis is used to identify components that are associated with effective metacognitive training programs in reading research. Forty-three studies, with an average of 81 students per study, were synthesized. It was found that metacognitive training could be more effectively implemented by using small-group instruction, as opposed to large-group instruction or one-to-one instruction. Less intensive programs were more effective than intensive programs. Program intensity was defined as the average number of days in a week that instruction was provided to students. Students in higher grades were more receptive to the intervention. Measurement artifacts, namely teaching to the test and use of nonstandardized tests and the quality of the studies synthesized played a significant role in the evaluation of the effectiveness of the metacognitive reading intervention.

THEORETICAL FRAMEWORK

Introduction

On Defining Metacognition

Regardless of the subtle differences in defining metacognition, there is a common ground on which reading researchers tend to agree that metacognition, in general, refers to "thinking about thinking." Reading researchers Forrest-Pressley and Waller (1984) wrote that "Metacognition is a construct that refers, first, to what a person knows about his or her cognitions and second, to the ability to control these cognitions. . . . Cognition refers to the actual processes and strategies that are used by the reader" (p. 6). Many researchers (Billingsley & Wildman, 1990; Haller, Child, & Walberg, 1988; Jacobs & Paris, 1987; Spires, 1990) have pointed out that the origin of metacognition can be traced back to research on young children conducted by Flavell and collaborators in the 1970s (Flavell, 1971; Flavell & Wellman, 1977).

To implement metacognitive intervention is to provide training on a strategy (Snowman, 1984) that purposely groups specific skills (e.g., summarization, and monitoring and resolving text comprehension obstacles) for the sake of enhancing reading performance. Metacognition is of particular interest to reading researchers because it is considered to be teachable (Haller et al., 1988; Jacobs & Paris, 1987; Paris, Cross, & Lipson, 1984; Paris & Jacobs, 1984; Paris & Oka, 1986) for improving students' reading comprehension.

The Effect of Metacognitive Reading Intervention

Researchers have attempted to examine systematically the effectiveness of metacognitive intervention in reading instruction. Some have found this intervention to be effective (Haller et al., 1988; Rosenshine, Meister, & Chapman, 1996) whereas others have not (Duffy et al., 1986; Jacobs & Paris, 1987). All of the summarized findings in the metacognitive intervention literature can be categorized according to two types of reviews, namely qualitative reviews and quantitative syntheses.

Qualitative reviews (Baker, 1989; Jacobs & Paris, 1987; Spires, 1990) do not inform the field about the average effect of metacognitive instruction. Although quantitative syntheses could answer the question of average effect, only one such synthesis (i.e., Haller, et al., 1988) has been conducted to evaluate specifically the effect of metacognition on reading comprehension. However, because that study was conducted 10 years ago, the synthesis did not include recent metacognitive intervention studies. Hence, a more up-to-date research synthesis, using improved meta-analysis techniques, was needed to accumulate new findings on metacognition. The present study was undertaken to serve that purpose. In addition, the researcher sought to answer questions that have not been answered in other reviews, by examining the relationship of the metacognitive intervention effect to training, including instructional time, small-group instruction, reading ability, and grade levels. These training characteristics are discussed in subsequent sections.

Training and Evaluation of Metacognitive Intervention

Observant readers may realize that training characteristics of metacognition frequently are confounded with nontraining characteristics such as measurement artifacts like teaching to the test and the use of nonstandardized tests. Haller et al. (1988) found an average effect of 0.71 standard deviation in the 20 metacognitive studies they synthesized. This effect (0.71), construed as "impressive" by Hattie et al. (1996, p.102), ranked second among five other meta-analyses on the intervention of study skills. The effect of metacognitive intervention on reading comprehension remains debatable without determining how much of the effect was contributed to by training characteristics, as opposed to measurement artifacts.

Research Questions on the Training of Metacognition

Q1. What is the relationship between training intensity and the effectiveness of metacognition?

The relationship between the effect of metacognitive instruction and the duration of intervention is an important issue because, all things being equal, nobody would object to providing a brief intervention to students if it was as effective as a year-long intervention. In their 1988 meta-analysis, Haller et al. investigated whether the effectiveness of metacognitive intervention was a function of duration of instruction. They concluded that 10 minutes or less instruction per lesson was insufficient. The researchers called for additional research to provide further clarification because some of the primary studies they used did not report the duration of intervention, and thus they were able to analyze only a subset of their primary studies for this *duration* variable. By including a larger pool of primary studies and by using a more detailed method to determine the importance of the time variable, the author expected that the effect of duration of intervention would be verified. The author used two variables (i.e., total number of intervention days and number of intervention days in a school week) to obtain more information about the effect of the duration of an intervention program.

Q2. How does the use of reading groups influence the effect of metacognitive training?

A substantial body of research, in reciporical teaching (Gilory & Moore, 1988; Lysynchuk, Pressley, & Vye, 1990; Palincsar & Brown, 1984; Palinscar, Brown, & Martin, 1987; Peterson, 1992) has indicated that significant gains in students' reading ability can be brought about through providing overt instruction, modeling, practice, and feedback. Some researchers (Slavin, 1983a, 1983b; Webb, 1985) have found that peer interaction, or cooperative learning, provides students with an opportunity to take responsibility for one another's achievement as well as their own. Students gain in achievement by taking turns elaborating their understanding of the skills. Reciprocal teaching and cooperative learning (peer interaction) require a substantial amount of time and interaction between students and teachers as well as among students themselves. Without interacting with peers, students cannot benefit from these modes of instruction. Assigning students to large groups, on the other hand, defeats the purpose of providing an opportunity for cooperative learning, as there is very little time and opportunity for peer interaction.

Even though research has supported the notion that group learning facilitates reading comprehension, little research has been done on the effect of metacognitive training under this mode of instruction. With a meta-analysis to investigate this issue, one could address questions such as "How well does the use of reading groups (collaboration) facilitate a metacognitive intervention, that requires individuals' abilities to control cognitive processes?"

Q3. To what extent can metacognitive training improve reading performance for different students (poor readers, students with learning disabilities, and students with no learning disabilities)?

Garner (1987, p.105) pointed out that strategy-training studies are invaluable for distinct reasons. Concerning the practical reasons for training students to use metacognitive strategies, Garner noted that it is important to investigate the extent to which these strategies can help poor readers improve their reading performance on academically fundamental tasks. An implication of Garner's statement is that poor readers could have been disadvantaged because they do not have the essential reading skills to perform on basic tasks, but that this phenomenon can be changed through metacognitive intervention.

Based on the premise that an effective intervention should be effective for a diversity of students, in the current meta-analysis, the author examined the claim that metacognition can improve the reading performance of students with learning disabilities, with no learning disabilities, and with low reading levels (poor readers). Students with learning disabilities and those who read at low reading levels are referred to as remedial students throughout the current meta-analysis. Students with no learning disabilities are referred to as nonremedial students.

Q4. To what extent does metacognitive training improve the reading performance of students in different grade levels?

Knowing the grade level at which metacognitive training can improve reading comprehension appears to be an important factor for making plans to embed this intervention in the regular school curriculum, because teachers and administrators can allocate their resources accordingly. Garner (1987) stated that "Younger children (particularly those in kindergarten or in grades 1 or 2) know substantially less than older children about themselves, the tasks they face, and the strategies they employ in the areas of memory, reading, and attention" (p.35). Haller et al. (1988) concluded that metacognition was more effective for seventh and eighth graders than for students in lower grades. Haller et al. did not report how much more effective metacognition would be for students in higher grades. In the current study, the author set out to reexamine that matter.

Research Questions on the Evaluation of Metacognition Intervention

Q5. What is the effect of "teaching to the test?"

An intervention program associated with a positive effect is not necessarily an effective program unless the positive effect is attributable to the program's treatment characteristics instead of other nontreatment characteristics (i.e., program characteristics) such as the presence of a teaching-to-the-test-effect. Teaching to the test is counterproductive to the intended goals of metacognition. Garner (1987) asserted that "teachers must present strategies as applicable to texts and tasks in more than one content domain" (p.134). If students are taught too specifically to the content and/or context of a test, they might not tend to generalize the strategy to a broader domain of knowledge, even within the same content area.

In the measurement literature, for example, Mehrens (1984) stated that teaching to the test could happen when teachers teach to specific questions on a test or to specific objectives. Taking Mehrens' definition of teaching to the test one step further, one could deduce that teachers are not likely to be able to teach to the test if they do not know what items or specific objectives will be on the test. Using this logic, the researcher hypothesized that, in the context of reading intervention studies, teaching to the test is likely to happen when experimenters have a dual role in the studies (i.e., experimenters being the test writers and the instructors).

Q6. How would the use of nonstandardized tests influence the metacognition effect?

In addition to the role of the instructors, the selection of measures is critical in evaluating the effectiveness of intervention programs. It is well documented in the literature that the results of reading comprehension interventions depend on the selection of outcome measures. Specifically, researchers such as Blaha (1979), Brady (1990), Cohen (1983), Dermody (1988), Haller et al. (1988), Jacobs and Paris (1987), Lysynchuk, et al. (1990), Rosenshine et al. (1996), Taylor and Frye (1992), and Walker and Schaffarzick (1974) have found that positive effects occur most frequently on nonstandardized tests. However, these researchers did not examine the interrelationship among metacognitive intervention, use of nonstandardized tests, and other variables (e.g., students' grade level and ability level) that could have an interaction effect with the use of standardized tests. Armbruster (1984) and Rosenshine et al. (1996) pointed out that standardized and nonstandardized tests differ in format and the knowledge required to answer the questions, and these differences might interact with students' ages and ability levels. The author of this study examined the net influence of the nonstandardized test effect after controlling for students' ability levels and ages.

METHODS

Literature Retrieval

Forty-three primary studies were selected for analysis. These studies met two criteria. They: (a) provided sufficient information for conducting a meta-analysis (i.e., means and standard deviations for the treatment group and for the control group) and (b) were designed to deliver metacognitive instruction.

Two approaches were used to select these 43 studies. Using the first approach, which White (1994) called *references in review papers written by others*, 23 primary studies were located. These 23 primary studies came from two review articles (Haller et al., 1988; Lysynchuk, Pressley, d'Ailly, Smith, & Cake, 1989) that were found using the Educational Resources Information Center (ERIC) electronic database.

Using Paris and Jacobs' (1987) taxonomy of metacognition, a well-recognized framework providing an operational definition for metacognition (Schraw & Moshman, 1995), the author selected 12 of the 38 journal articles reviewed by Lysynchuk et al. (1989). These 12 studies provided sufficient information and were judged to be related to metacognitive intervention. Another 11 relevant studies, including dissertations, presented papers, and journal articles came from the second review article, Haller et al. (1988), which was a meta-analysis of metacognitive interventions for reading comprehension. See the Theoretical Framework section for the differences between the current paper and Haller et al. (1988).

Requests for references in review papers written by others were not always successful. A potential list of 64 primary studies in metacognition was inaccessible because the author was unable to obtain the reference list for articles summarized in a comprehensive review on the history of the National Reading Conference (NRC) (i.e., Baldwin, et al., 1992). These reviewers conducted a global analysis of 2,139 articles published in the *Journal of Reading Behavior* and the NRC yearbook. Future researchers should continue to pursue the references.

Through the second approach, an approach to update and expand the search, 20 additional primary studies were found. White (1994) called this approach *computer search of abstract databases* in which the author used a keyword search on the ERIC database (1982 - 6/1996). Appendix A shows the keywords used in this search. The primary studies found using this approach matched the two criteria described earlier in this section.

Some specific metacognitive skills taught in the 43 primary studies included Text Summarization, Text Reinspection (look-back), Drawing Inferences From Text, and Monitoring and Resolving Text Comprehension Obstacles. The author and a group of researchers coded these primary studies. The section on Coding Procedure and Inter-Rater Reliability describes the coding process and results.

Outcome Measures (Dependent Variables) and Moderators (Independent Variables)

Reading Measures

Synthesizing nonreading measures would yield invalid results for the evaluation of an reading intervention program. If nonreading tests (e.g., motivation and affect) and reading tests were treated as a single outcome measure, one could not disentangle the intervention effect on reading comprehension from that on other constructs. To ensure that the current meta-analysis synthesized the effect of metacognitive intervention on reading comprehension, only reading-comprehension outcome measures were included in the meta-analysis. Using the classification developed by Harris (1990), the author selected 123 reading-comprehension outcome measures. These measures belonged to one of the two major categories (i.e., product measures and process measures) defined by Harris (1990). Two types of **product measures** are: (a) retelling and (b) using questions and answers. The questions-and-answers paradigm has three variations: aided recall; unaided recall, and true/false items. Four types of **process measures** are cloze tests, miscue analysis, think-aloud tests, and eye-movement tests. No studies in the meta-analysis used eye-movement tests as reading measures. Of the 123 reading comprehension measures, 32 were standardized tests and 91 were nonstandardized tests. Table 1 shows the standardized reading tests used in the primary studies.

Moderators

In addition to the outcome variables (reading measures) and independent variables directly related to the research questions, three variables also were analyzed to control for the quality of primary studies so that valid inferences could be made with regard to the effect of metacognitive intervention. These variables were *random assignment*, *design of the primary studies* (i.e., posttest-only control group design and pretest-posttest control group design), and *Hawthorne effect*. A background variable, *school location*, serving as a general purpose variable, to examine the sociodemographic status of the student participants, also was analyzed.

Coding Procedure and Inter-Rater Reliability

Training

The author coded all of the primary studies. Fifteen primary studies were randomly selected and double coded by 10 volunteers in one 90-minute session. These 10 coders were members of a meta-analysis research group including nine doctoral students and one faculty member. In the first 45 minutes, all coders practiced coding on one anchor study. Coders and the author discussed any ambiguity as they went through the practice. Minor changes in the labeling of the codes were made after the mock coding, and these changes were applied to the real coding of the 15 primary studies. See Appendices B and C for the coding instrument and the coding instruction, respectively. Due to time limits, no dissertations were assigned to coders. After the practice, each coder was randomly assigned to code a different study. As coding time varied by the length of articles and individual differences, five of the 10 coders each coded two studies and the rest each coded one study.

Rater Reliability Measures

Percentage of agreement between the author and the original codes of the additional coders was obtained for all 12 variables. The codings of all voluntary coders were treated as if they had come from a single coder. The percentage of agreement for every variable is presented in Appendix D. A high inference variable (i.e., Hawthorne effect) showed low interrater reliability (percentage of agreement = 43%). A follow-up interrater reliability index, Cohen's Kappa κ (Crocker & Algina, 1986), was also calculated for this variable and indicated that the reliability (κ = .018) was close to a random guessing level ($\kappa = 0$). Thus, this variable was excluded from subsequent analyses.

Meta-Analysis and Effect Size Computation

Glass (1976) employed a quantitative research synthesis technique, labeled meta-analysis, for summarizing research studies. Since then, meta-analysis has developed rapidly. The *standardized-mean-difference* effect size (Hedges' g) is appropriate when primary studies report means and standard deviations for a control group and for a treatment group. In this paper, the author used Hedges' *gs*. Other types of effect sizes, such as correlations (r) and proportions (e.g., Cohen's h), could also be used. Rosenthal (1994, pp. 231-244) provided a concise review of different types of effect sizes. Even for the same type of effect-size measure, various formulas exist for different purposes. For instance, Becker (1988) discussed the conception and provided formulas for synthesizing mean-change measures, which are used in one of the most common experimental designs, namely the pretest/posttest design (Becker, 1988; Campbell & Stanley, 1963). How to obtain standardized-mean-difference effect sizes from two common experimental designs (i.e., pretest/posttest-control-group design and posttest-only-control-group design) is described in the following two paragraphs.

For the pretest/posttest control-group design, the estimated effect size is defined as

$$g = \frac{(\hat{\mu}_{post\_treatment} - \hat{\mu}_{post\_control}) - (\hat{\mu}_{pre\_treatment} - \hat{\mu}_{pre\_control})}{\hat{\sigma}_{post\_pooled}}. \quad (1)$$

BEST COPY AVAILABLE

For the posttest-only-control-group design, the population mean difference between the treatment group and the control group in the pretest is assumed to be zero. The effect size is estimated by

$$g' = \frac{(\hat{\mu}_{post\_treatment} - \hat{\mu}_{post\_control})}{\hat{\sigma}_{post\_pooled}} \quad (2)$$

where for both types of research design

$$\hat{\sigma}_{post\_pooled} = \sqrt{\frac{(N_{post\_treatment}-1)\hat{\sigma}^2_{post\_treatment} + (N_{post\_control}-1)\hat{\sigma}^2_{post\_control}}{N_{post\_treatment} + N_{post\_control} - 2}} \quad (3)$$

The effect size *g* is a biased estimate of the population effect size. The unbiased estimator (Hedges, 1981) is d = c(m)g, where m = ($n_E + n_C$ - 2) and c(m) is approximated by 1 - 3/(4m - 1). Note that n_E and n_C are the sample size of the treatment group and control group, respectively.

The effect sizes (*ds*) obtained in each study are then treated as the dependent variable in the generalized least square (GLS) regression approach, and are predicted by moderator variables of interest. The essential underlying theory for GLS, discussed, for instance, by Seber (1977, p. 60) and Raudenbush et al., (1988), is summarized in Appendix E.

Regression Analysis for Multiple Dependent Outcomes

Metacognitive intervention studies frequently employ more than one reading measure to evaluate the intervention effect. These reading measures, however, are correlated and for this reason, the effect could be overestimated if the correlations were not adjusted accordingly. Given the recent proliferation of meta-analytical techniques, researchers have devised methods applicable for analyzing primary studies that used multiple outcome measures (i.e., multiple outcome measures used for measuring the same group of subjects). However, there is no overarching conclusion to this issue. Chiu (1997) reviewed methods for meta-analyzing studies with multiple outcomes and suggested that the GLS regression method be used for reading comprehension studies, provided that a sensitivity analysis was conducted.

Using the same primary studies as were analyzed in the current meta-analysis, Chiu (1997) found that treating correlated outcomes (r = .80) as if they were uncorrelated (r = 0) would overestimate both the effects (regression coefficients) and their precision (standard errors). He also concluded that, for studies in which correlations were not reported for the dependent multiple measures, a substitute of .60 would be a reasonable approximation when applying the GLS regression method. In the current study, this medium-size correlation (i.e., r = .60) was used as a substitute for the unreported correlations.

Fourteen GLS regressions were analyzed. Eleven were used to examine the extent to which each moderator contributed to the metacognitive reading intervention effect. Three other regressions were employed to examine the unique contribution of each moderator while holding constant the other moderators.

RESULTS

Summary of Primary Studies Included in the Meta-Analysis

Forty-three studies with 123 effect sizes were analyzed. A total of 3,475 students participated in these 43 studies, with an average of 81 students in each study. The unbiased average effect size was 0.67. The distribution of these effect sizes is shown in Figure 1. The pool of the primary studies came from a variety of sources, including journal articles, dissertations, and unpublished manuscripts (see Table 2). Table 3 shows the descriptive statistics for the study-level variables analyzed in the meta-analysis. The primary studies included student participants from second grade through college level. In 35 studies, students were selected from only one grade level, and in eight studies students were selected from multiple grade levels. Regarding the reading comprehension outcome measures, 24 studies used only nonstandardized tests and 19 used one or more standardized tests.

Approximately two-thirds ($\underline{n}$ = 27) of the 43 studies reported more than one outcome measure. On average, 2.86 outcomes were reported in each of the primary studies. The median and mode were 1.5 and 2, respectively. One study had six outcomes on each of the two groups of student participants who were provided metacognitive intervention. Consequently, this study contributed the largest number of outcome measures (i.e., 12 outcomes), among the 43 studies, to the current meta-analysis.

The Effectiveness of Metacognitive Reading Intervention

The effectiveness of metacognitive reading intervention is contingent on the outcome measure used in the program. When a nonstandardized test was used, the effect was significantly higher than when a standardized test was used. The metacognitive intervention effect was 0.24 ($z = 5.44$, $p < .001$) when standardized tests were used (see Table 4, Model 2). However, the effect was elevated to 0.24 + 0.37 = 0.61 when nonstandardized tests were used. This nonstandardized test effect was still significant, even when other factors were held constant. The *Final Model* showed that the effect size measured by nonstandardized tests and standardized tests could have a difference of .52 standard deviation (see Table 4, Model 14).

The *Final Model* also showed that when researchers or collaborators delivered instruction to the students, the average effect size was 0.24 standard deviations higher than that with regular classroom-teacher instruction ($z = 2.39$, $p = .009$), all other factors being equal. Therefore, these special instructors were likely to be able to teach to the test. Putting together the two pieces of information (i.e., the nonstandardized test effect and the instructor effect), one would conclude that the intervention had a significantly higher effect size (i.e., 0.52 + 0.24 = 0.76) when researchers taught the students and used a nonstandardized test. Even though this research showed that the instructor effect and nonstandardized test effect were significant, it did not prove that teaching to the test happened in metacognitive intervention. Unless instructors' intentions are measured, one cannot show that instructors did teach to the test.

The *Final Model* excluded the moderator *dura_prg* (duration of program) because of its high correlation with *random* (point-biserial correlation was -.70). *Dura_prg* was dropped from the *Interim* model, instead of *random*, because it made only a negligible contribution to the regression model (-0.003, $z = -2.43$, $p = .008$) and, in contrast, the moderator *random* had a relatively high coefficient (-.59, $z = -3.22$, $p < .001$). The multicolinearity between *drua_prg* and *random* was probably due to the fact that reading intervention programs were usually implemented during school days in regular classrooms and it was difficult to randomly assign students to a treatment group for longer-term programs.

In addition to the program variables (i.e., *random, mem_type,* and *instors1*), the intervention variables were also significant. The negative coefficient for *dura_int* indicated that less intensive programs were more effective than intensive programs. The effect size was reduced by 0.07 ($z = -2.94$, $p = 0.002$) for every treatment day given to students within the same week. Consistent with the notion that collaborative learning could facilitate metacognitive behaviors, the results indicated that metacognitive intervention had a larger effect (*smallgrp* = 0.30, $z = 2.86$, $p = 0.002$) in small-group settings. Although the test statistic was marginally significant (*remedial* = 0.15, $z = 1.60$, $p = .055$), metacognitive intervention seemed to work better for low-ability students or any students who were diagnosed as remedial students. The *Final Model* also indicated that metacognitive intervention was more effective (*grade5* = 0.21, $z = 2.43$, $p = 0.008$) when it was given to students in fifth grade or higher. This result was also consistent with those found by Haller et al. (1988), that metacognitive reading intervention required cognitive abilities that young children might not have developed or acquired.

CONCLUSION

The author found that primary studies associated with certain program characteristics had a larger effect than those that did not have such associations. Hence, it is clear that one must examine program characteristics (coined measurement artifacts in the current meta-analysis) when investigating the intervention's effectiveness. More specifically, it was found that studies associated with characteristics such as the absence of random assignment, the selection of nonstandardized outcome measures, and the presence of instructors' dual role would yield a favorable decision regarding the implementation of metacognitive reading programs.

With respect to the training effect, metacognitive reading intervention is particularly effective in small-group settings for students in fifth grade or higher. This paper also indicated that remedial students seem to benefit from metacognition. It also indicated that reading programs that have spanned a long period of time are just as effective as those that cover only a short period of time, all other treatment characteristics and program characteristic being equal. Moreover, less condensed programs are more effective than intensive programs, all characteristics being equal.

In this paper, the author did not examine why metacognitive intervention worked (i.e., no casual relationship was found) even though this paper has indicated that it was effective and identified some correlates of its effectiveness. Learning researchers might want to continue to explore this question. In addition, the author investigated metacognition in the absence of other related constructs in learning (e.g., motivation and self-regulation). Researchers may extend this paper by incorporating metacognition with these other constructs; Boekaerts (1995) and Zimmerman (1995), who discussed the construct of metacognition with motivation theories and self-regulation, respectively, provided a useful overview.

Methodologically, the author used the same substitute correlation across studies for all dependent outcome measures. If the correlations among a group of dependent measures or for a particular study vary dramatically, the conclusion discussed above might not hold. However, based on the assumption that measures of the same construct should exhibit high convergent construct validity, the correlations among reading measures should not vary to an extent that would alter completely the preceding conclusions. To examine this assumption, future construct validity research should be conducted (for construct validity research, see, for example, Anastasi, 1988; Brown, 1983; Cronbach & Meehl, 1955; Fiske, 1987; Nunnally & Bernstein, 1994), and the multitrait-multimethod technique (e.g., Crocker & Algina, 1986; Nunnally & Bernstein, 1994) could be used.

An alternative to carrying out construct validity research is to conduct computer simulated sensitivity analyses, using a unique substitute correlation for each pair of dependent measures. In simulation studies, one could determine the exact effect of unreported correlations. Computer simulations are especially suitable when it is difficult to obtain a solution analytically (algebraically). Nunnally and Bernstein (1994) provided an introduction to computer simulation, and Harwell (1992, 1995) provided a more in-depth discussion.

TABLES

Table 1: Standardized Reading Measures Used in the Primary Studies

1. California Achievement Test
2. Davis Reading Test
3. Gates-McGinitie Reading Test
4. Iowa Test of Basic Skills
5. La Prueba Spanish Reading Test
6. Metropolitan Achievement Test
7. Nelson-Denny Reading Test
8. Progressive Achievement Test
9. Stanford Achievement Test
10. Stanford Diagnostic Reading Comprehension
11. Test of Reading Comprehension

Table 2: Summary of Sources for Primary Studies Included in the Meta-Analysis

	Frequency	Percent
1 Bilingual Research Journal	1	2.3
2 Cognition and Instruction	2	4.7
3 Contemporary Educational Psychology	2	4.7
4 Dissertation	2	4.7
5 Educational Research Quarterly	1	2.3
6 Elementary School Journal	1	2.3
7 ERIC Document	8	18.6
8 Journal of Educational Psychology	3	7.0
9 Journal of Educational Research	1	2.3
10 Journal of Reading	1	2.3
11 Journal of Reading Behavior	3	7.0
12 Journal of Research in Reading	1	2.3
13 Learning Disability Quarterly	1	2.3
14 Modern Language Journal	1	2.3
15 Psychology in the Schools	1	2.3
16 Reading Research and Instruction	2	4.7
17 Reading Research Quarterly	9	20.9
18 Reading Teacher	1	2.3
19 Research and Teaching in Developmental Education	2	4.7
Total	43	100.0

Table 3: Descriptive Statistics for Primary Studies Included in the Meta-analysis (Study Level Information)

Variable Names	Variables Labels	N	Min	Max	Mean[a]	S.D.
STUD_NU	Number of effect sizes (i.e., measures) per study	43	1	12	2.86	2.36
MEM_TYPE	Number of studies that used ONLY nonstandardized tests	43	0	1	.56	.50
INSTORS1	Did the researchers(s) provide instruction to students?	42	0	1	.48	.51
RANDOM	Was random assignment employed in the study?	43	0	1	.86	.35
URBAN1	Was the school located in an urban area?	43	0	1	.23	.43
SUBURBA1	Was the school located in a suburban area?	43	0	1	.28	.45
RURAL1	Was the school located in a rural area?	43	0	1	.16	.37
LOCATUK1	Was the school location unknown?	43	0	1	.33	.47
DURA_PRG	Duration of the entire training program (days)	41	1	180	44.87	49.05
DURA_INT	Number of intervention days per week (i.e., 5 school days in a week)	42	0.4	5	3.159	1.594
ONESTUDT	Was the instruction on a one-to-one basis?	43	0	1	.23	.43
SMALLGRP	Was the instruction on a small-group basis?	43	0	1	.44	.50
LARGEGRP	Was the instruction on a large-group/classroom basis?	43	0	1	.33	.47
REMEDIAL	Did the students have reading problems?	43	0	1	.35	.48
PUB_YR	Year of publication of the study	43	1979	1995	1987	4.34
Valid N (listwise)		40				

a. Footnote: The mean of dichotomous variables (i.e., min = 0 & max = 1) multiplied by 100 equals the percentage of studies that were associated with the presence of the corresponding characteristics.

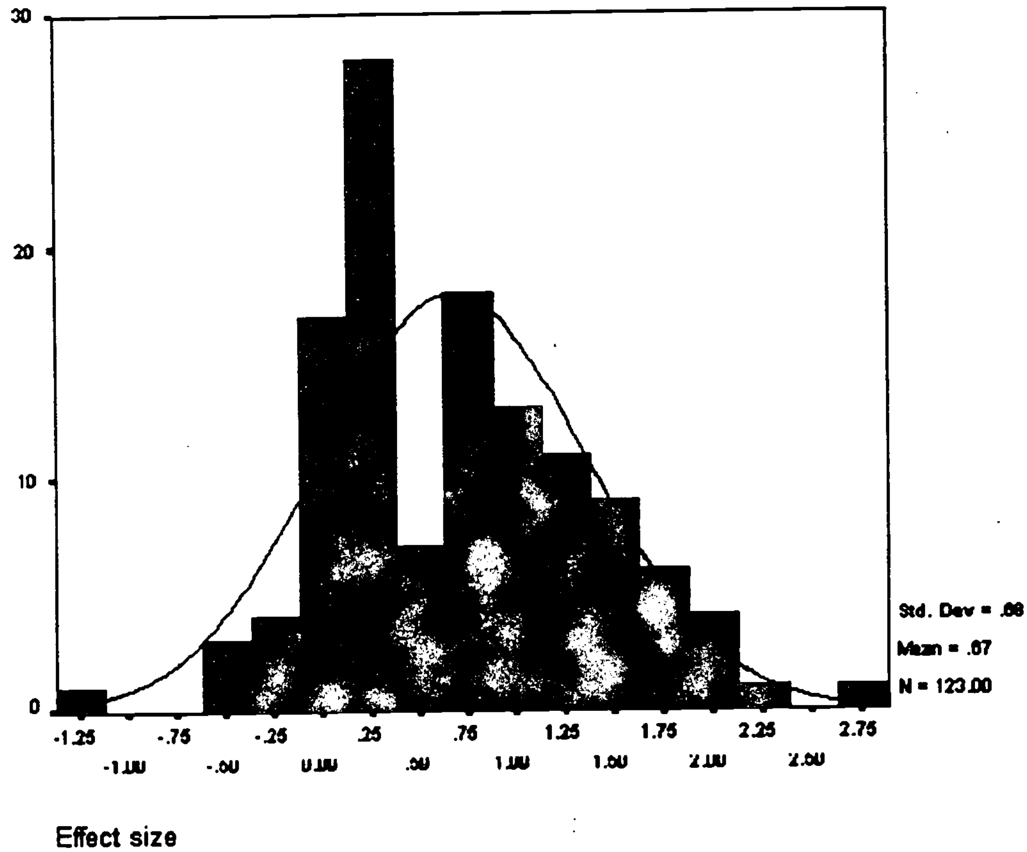
Table 4: Regression Coefficients and Corresponding Standard Errors

Research Questions/ Regression Models	Independent Variables	Coefficients (SE)
1. Unconditional model	Intercept	0.4** (0.035)
2. Nonstandardized effect	Intercept, mem_type	0.24** (0.044), 0.37** (0.059)
3. Teaching to the test	Intercept, instors1	0.37** (0.043), 0.12 (0.074)
4. Research design	Intercept, res_degn	0.5** (0.06), -0.14* (0.073)
5. Random assignment	Intercept, random	0.33** (0.073), 0.09 (0.083)
6. Program length	Intercept, dura_prg	0.58** (0.054), 0 (0.001)
7. Treatment intensity	Intercept, dura_int	0.42** (0.073), -0.01 (0.021)
8. Group size	Intercept, onestudt & smallgrp	0.43** (0.066), -0.03 (0.113), -0.03 (0.081)
9. Student ability	Intercept, remedial	0.39** (0.039), 0.08 (0.086)
10. Grade level	Intercept, grade5 (college level included)	0.27** (0.051), 0.25** (0.07)
11. School location	Intercept, urban1, suburba1, rural1	0.56** (0.081), -0.21* (0.12), -0.2* (0.094), -0.13 (0.125)
12. Combined Model (included all above variables)	Intercept, mem_type instors1, res_degn, random, dura_prg, dura_int, onestudt, Smallgrp, remedial, grade5(college level included), urban1 Suburba1, rural1	0.81* (0.33), 0.45** (0.085), 0.25* (0.111), 0.11 (0.101), -0.62* (0.191), -0.003* (0.001), -0.11** (0.032), -0.18 (0.138), 0.29* (0.112), 0.24* (0.119), 0.13 (0.097), 0 (0.144), 0.08 (0.126), 0.07 (0.137),
13. Interim Model (excluded insignificant variables from the Combined Model)	Intercept, mem_type, instors1, random, dura_prg, dura_int, Onestudt, smallgrp, Remedial, grade5(college level included)	0.89* (0.3), 0.43** (0.082), 0.21 (0.1), -0.59** (0.176), -0.003* (0.001), -0.1** (0.028), -0.22* (0.123), 0.31* (0.106), 0.19* (0.097), 0.12 (0.096),
14. Final Model (excluded 'dura_prg' from the Interim Model)	Intercept, mem_type, instors1, random, dura_int, onestudt, Smallgrp, remedial, grade5(college level included)	0.26* (0.150), 0.52** (0.073), 0.24* (0.099), -0.28* (0.120), -0.07* (0.027), -0.17 (0.122), 0.30* (0.106), 0.15* (0.096), 0.21* (0.088)

Note: All models were based on 114 cases. Nine cases were excluded because of missing data. These cases are 21, 22, 38, 92-95, 114, and 115. The single-starred coefficients and the double-starred coefficients were significant at $p < .05$ and at $p < .001$, respectively.

FIGURES

Figure 1: Distribution of Unbiased Effect Sizes



APPENDICES

Appendix A: Keyword Search Used in ERIC

(Metacognitive or Metacognition) and PY=19xx where xx = 88 - 96	n = 1106
Reading Comprehension and (instruction or intervention) and effect* and meta*	n = 144

Appendix B: Coding Instrument

Coder's Initial: ____________________ **Article title:** ____________________

Date of Coding: ____________________ **Article ID:** ____________________
(Number appears on the top right corner of the article)

Instruction: For each of the following variables, put down the page number of where you find the answer. Do not leave variables unchecked. You are encouraged to make inferences for information that is not reported in the article. When making an inference, please use a brief description to document your reasoning.

Section A: Study Identification

- Year of Publication: __________

Section B: Characteristics of Setting

1. Where is the school located? Page: ________

    1. Urban
    2. Suburban
    3. Rural
    -9. Not reported

2. What is the size of the instructional group? Page: ________

    1. One student (i.e., individual basis)
    2. Small groups (i.e., 2 to 10 students, classroom of 10 or fewer students included)
    3. Large groups (i.e., more than 10 students, classroom of more than 10 students included)

3. Who provided instruction to the students? Page: ________

    0. Non-researchers (e.g., classroom teachers and other teaching personnel)
    1. Rsearchers and collaborators (i.e., including researchers who were also teachers)
    2. Others. Please specify: ________

Section C: Subject Characteristics

4. Were the students selected in the study because they had reading problems?
Page: ________

    0. No
    1. Yes, the study was a remedial program.

5. What is/ are the grade level(s) of the student participants? Page: __________

1	2	3	4	5	6	
7	8	9	10	11	12	college level
middle school			junior high school			high school

Section D: Training Characteristics

6. Were students/classrooms randomly assigned to treatment groups?
Page: __________

   0. No.
   1. Yes, students were randomly assigned.
   2. Yes, classrooms were randomly assigned.

7. Did the same instructor(s) teach both the treatment group and the control group?
Page: __________

   0. No.
   1. Yes.
   2. Not applicable, because classrooms were randomly assigned to treatment groups.
   3. Not reported.

8. Based on what was reported in the study, do you think the control group believed they were receiving a treatment?
Page: __________

   0. No, I do not think so.
   1. Yes, I think so.

9. How many training sessions were given to the students? < reported inferred >

   Number of sessions: ________ Page: __________

10. How long (in minutes) did each training session last? < reported inferred >

   Duration of a session: ________ Page: __________

11. How long did the entire training program last? < reported inferred >

   Number of: Page: __________
   months ______
   weeks ______
   days ______

12. On average, how many days in a week (i.e., five school days) did the instructions take place?

Number of days: ______ < reported inferred >

Page: ________

<<The End>>

Appendix C: Coding Instructions

Section B: Characteristics of Setting

1. **School location** – is usually reported. If it is not reported, then circle *not reported*. Do not make any inferences using the authors' affiliation because researchers could conduct their research in schools far away from their affiliations.

2. **Size of instructional groups** – is not always reported explicitly in an article. You will have to put pieces together, usually from the method section. One pitfall that you should beware of -- do not make inferences of the size of an instructional group from the Ns reported for the treatment groups. Researchers might put treatment subjects into small instructional groups but report only the total number of subjects in the treatment group as a whole. Another issue you may find in a study is that the size of an instructional group could change over time. If this happens, it usually goes from a large group to a smaller group. For this instance, you would consider the size of the instruction group to be *small groups*.

3. **Instructors** – a variable used to identify who provided instruction to student participants. The objective is to identify whether the experimenters or their collaborators provided instruction to students. Circle *researchers* if experimenters or collaborators gave instruction. On occasions, the instructors would have a dual role that they were both the researchers and classroom teachers. In this case, select *researchers*. The category *non-researchers* includes any instructors who were classroom teachers of the student participants and other teaching personnel. Please note that some studies used computer systems to provide online instruction. Human instructors were present only to provide minimal instruction and technical support. In this case, you would circle the third option, *others*, and put down *computerized instruction*.

Section C: Subject Characteristics

4. **Subject selection criterion** – a variable used to capture whether the subjects were having reading problems, reading below a certain grade level, or had a learning disability. Instructional programs provided for these students were considered remedial programs. Although studies might use different language to describe student participants, *subject selection criterion* is a fairly straightforward variable because researchers always describe their subjects in the abstract of the article or make their titles explicit enough to catch readers' attention. The following are sample titles for remedial programs:

- Fostering Comprehension Monitoring in Below Average Readers Through Self-Instruction Training.
- An Instructional Study: Improving the Inferential Comprehension of *Good and Poor* Fourth-Grade Readers.
- Comprehension Monitoring: Detection and Identification of Test Inconsistencies by LD (Learning Disabled) and Normal Students.

5. **Grade level of students** – a *circle all that applies* variable. Note that some studies might not report the grade level of students. They might just report whether the students were in *middle school, junior high school, or high school*. In this case, you don't have to make an inference to the grade levels. Simply circle any one of the three choices.

Section D: Training Characteristics

6. **Random assignment** – two common ways that random assignment is used in intervention studies (i.e., random assignment to student level and random assignment to classroom level). Both types should be considered as random assignment.

7. **Counterbalance of instructors** – a variable that you may not even have information to make inference to. If that's the case, circle *Not reported*. Based on my coding, 21% of the 43 studies did not report and provided no information for this variable.

8. **Hawthorne effect** – a high-inference variable for which you will make an inference based on what was given to the control group. Do not leave this variable unchecked. In some studies, researchers reported how they had tried to avoid the Hawthorne effect by giving students tasks to work on so that students believed they were receiving a treatment. In this instance, you would consider that the Hawthorne effect does not exist. If the control group were being told to do some *busy work* (e.g., reading a book with no instruction provided), you would consider that the Hawthorne effect was likely to happen. That is, it is unlikely that the subjects believed they were receiving a treatment.

Duration and intensity of treatment – a set of four variables.

9. **Number of training sessions**
10. **Duration of each session**
11. **Duration of the entire program**
12. **Number of sessions per week**

The above variables were reported in a wide range of ways. For some studies, you would have to gather the pieces from different sections of the article, or even have to make an inference from your own experience. For other studies, you could find all of the information reported in a sentence or two. Excerpt 1 in the following illustrates a situation in which you need not make any inference. Excerpt 2 is an example in which you need to make an inference.

Excerpt 1.

Each group of students received three 1-hour training sessions. Each treatment was carried out on the same day of the week for 3 weeks – for example, training for the four groups in the SQ condition took place on 3 consecutive Mondays. (Nolan, 1991)

The above excerpt is used to illustrate how you would fill in the following four variables.

1. Number of sessions: ____3____ <reported inferred>
2. Duration of a session: ___60____ minutes <reported inferred>
3. Entire training program lasted for: <reported inferred>

   ____0____ months

   ____3____ weeks

   ____0____ days

4. Number of days in a school week did instructions took place:

   ____1____ days <reported inferred>

Excerpt 2.

Both groups were pulled out of their regular English classes for three weeks to receive "special instruction. The schema group met on Mondays and Wednesdays and the traditional group on Tuesdays and Thursdays. (Singer & Donlan, 1982)

The above excerpt is used to illustrate how you would fill in the following four variables. Note that duration of instruction that took place in a regular school day was assumed to be 50 minutes.

1. Number of sessions: ____6____ <reported inferred >
2. Duration of a session: ___50____ minutes < reported inferred>
3. Entire training program lasted for: <reported inferred >

   ____0____ months

   ____3____ weeks

   ____0____ days

4. Number of days in a school week (i.e., 5 days) that instruction took place:

   ____2____ days <reported inferred >

Appendix D: Interrater Reliability

Variables	% Agreement	Valid Responses
Schl_loc	93%	15
Size_int	80%	15
Instors1	92%	13
Remedial	93%	14
Grade	93%	15
Random	86%	14
EptrSame	79%	14
Hawthorn	43%	14
Nu_sessn	71%	14
Dura_sen	93%	14
Dura_prg	73%	15
Dura_int	85%	13

Variables Labels:

Schl_loc: school location;
Size_int: size of instruction group;
Instors1: whether or not the instructors were also the experimenters;
Remedial: whether or not student participants had reading problems;
Grade: student grade level;
Random: whether or not random assignment was used;
EptrSame: whether or not the same experimenters provided instruction to both the treatment and the control group;
Hawthorn: whether or not control group subjects believed they were receiving a treatment;
Nu_sessn: number of instruction sessions;
Dura_sen: duration of an instruction session (in minutes);
Dura_prg: duration of the entire reading program (in days);
Dura_int: average number of days in a week that instruction was provided to students.

Appendix E: Formulas for the Generalized Least Square Regression Coefficients and Associated Standard Errors

The dependency caused by multiple measures is accounted for by the estimated variance-covariance matrix (S) of the effect size, where S is a block diagonal matrix with the first block containing the variance-covariance matrix of the effect sizes in the first primary study and the last block containing the variance-covariance matrix of effect sizes in the last primary study. Within each block, the variances and covariances of the dependent effect sizes are modeled (e.g., Gleser & Olkin, 1994, p. 348, equations 22-20 and 22-21), respectively, by

$$\hat{\sigma}_{jj} = \frac{1}{n_E} + \frac{1+\frac{1}{2}d_j^2}{n_C}, \; j = 1, \ldots, p,$$

$$\hat{\sigma}_{jj^*} = \left(\frac{1}{n_E} + \frac{1}{n_C}\right) r_{jj^*} + \frac{\frac{1}{2} d_j d_{j^*} r_{jj^*}^2}{n_C}, \; j^* \neq j. \quad (4)$$

The correlations between two dependent effect sizes, r_{jj^*}, are imputed by the substitute correlation (i.e., .6).

The effect sizes and their corresponding variances and covariances are used to estimate the regression parameters, standard error of the parameters, and associated probability values. In the following formulas, X and d represent the design matrix of a model and the vector of effect sizes, respectively. S is the estimated block diagonal variance-covariance matrix. The following matrix algebra formulas provide estimates of the predictors and other information needed for statistical tests.

- Estimates of the predictors are: $\hat{\beta} = (X'S^{-1}X)^{-1}X'S^{-1}d$ (5)
- Estimated variance-covariance matrix of the predictors is : $\hat{\beta} = (X'S^{-1}X)^{-1}$ (6)
- Estimated standard error $\hat{\sigma}_{hh}$ of any predictor is the square root of the *h*th diagonal element of $\hat{\beta}$ (7)

To obtain the individual test of the predictors, one would test the ratio of the estimated coefficient against the corresponding standard error. The ratio would have a z-distribution. More specifically, the test statistic is $Z = \hat{\beta}_h / \hat{\sigma}_{hh}$, where $\hat{\beta}_h$ is the *h*th element of the estimated parameter vector $\hat{\beta}$ (see Equation 5) and $\hat{\sigma}_{hh}$ (see Equation 7) is the standard error of the estimated parameter.

References [1]

+Anastasi, A. (1988). Psychological testing (5th ed.). New York: Macmillan.

Armbruster, B. B. (1984). The problem of "inconsiderate" text. In G. G. Duffy, L. R. Roehler, & Mason (Eds.), Comprehension instruction: Perspectives and suggestions (pp. 202-217). New York: Longman.

Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. Educational Psychology Review, 1(1), 3-38.

Baldwin, R. S., Readence, J. E., Schumm, J. S., Konopak, J. P., Konopak, B. C., & Klingner, J. K. (1992, December). Forty years of NRC publications: 1952-1991. Journal of Reading Behavior, 24(4), 505-532.

+Becker, B. J. (1988). Synthesizing standardized mean-change measures. British Journal of Mathematical and Statistical Psychology, 41, 251-278.

Billingsley, B. S., & Wildman, T. M. (1990). Facilitating reading comprehension in learning disabled students: Metacognitive goals and instructional strategies. Remedial and Special Education (RASE), 11(2), 18-31.

Blaha, B. A. (1979). The effects of answering self-generating questions on reading. Unpublished doctoral dissertation, Boston University.

Boekaerts, M. (1995). Self-regulated learning: Bridging the gap between metacognition and metamotivation theories. Educational Psychologist, 30(4), 195-200.

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

+Brown, F. G. Principles of educational and psychological testing (3rd ed.). New York: Holt, Rinehart and Winston.

+Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi-experimental design for research. Boston: Houghton Mifflin.

+Chiu, C. W. T. (1997). Synthesizing multiple outcome measures with missing data: A sensitivity analysis on the effect of metacognitive reading intervention. Unpublished Apprenticeship, Department of Counseling, Educational Psychology, and Special Education Michigan State University, E. Lansing.

Cohen, R. (1983). Students generate questions as an aid to reading comprehension. Reading Teacher, 36, 770-775.

[1] References tagged with a '+' are methodological references.

+Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52(281-302).

+Crocker, L., & Algina J. (1986). Introduction to classical and modern test theory. New York: Holt.

Dermody, M. (1988, February 17-20). Metacognitive strategies for development of reading comprehension for younger children. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, New Orleans, LA.

Duffy, G., Roehler, L., Meloth, M., Vavrus, L., Book, C., Putnam, J., & Wesselman, R. (1986). The relationship between explicit verbal explanation during reading skill instruction and student awareness and achievement: A study of reading teacher effects. Reading Research Quarterly, 21, 237-252.

+Fiske, D. W. (1987). Construct invalidity comes from method effect. Educational and Psychological Measurement, 47(2), 285-307.

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? Human Development, 14, 272-278.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition, pp. 3-34. Hillsdale, NJ: Lawrence Erlbaum.

Forrest-Pressley, D. L., & Waller, T. G. (1984). Cognition, metacognition, and reading. New York: Springer-Verlag.

Garner, R. (1987). Metacognition and reading comprehension. Norwood, NJ: Ablex Publishing.

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

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

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

Harris, A. J. (1990). How to increase reading ability: A guide to developmental and remedial methods (9th ed.). New York: Longman.

+Harwell, M. R. (1992). Summarizing Monte Carlo results in methodlogical research. Journal of Educational Statistics, 17 (4), 297-313.

+Harwell, M. R. (1995). An empirical study of the Hedges (1982) Homogeneity Test. ERIC document ED392822.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. Review of Educational Research, 66(2), 99-136.

+Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6(2), 107-8.

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

Lysynchuk, L., Pressley, M., d'Ailly, H., Smith, M., & Cake, H. (1989). A methodological analysis of experimental studies of comprehension strategy instruction. Journal of Reading Quarterly, 24 (4).

Lysynchuk, L., Pressley, M., & Vye, G. (1990). Reciprocal instruction improves reading comprehension performance in poor grade school comprehenders. Elementary School Journal, 40, 471-484.

Mehrens, W. (1984). National tests and local curriculum: Match or mismatch? Educational Measurement: Issues and Practice, 3(3), 9-15.

+Naunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw-Hill.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension fostering and monitoring activities. Cognition and Instruction, 1, 117-175.

Palinscar, A. S., Brown, A. L., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. Educational Psychologist, 22, 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, 76, 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., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. Developmental Review, 6, 25-56.

Peterson, S. (1992). The effects of reciprocal teaching and attributional training on reading comprehension. Unpublished master's thesis, University of Auckland.

+Raudenbush, S. W., Becker, B. J., & Kalaian, H. (1988). Modeling multivariate effect sizes. Psychological Bulletin, 103(1), 111-120.

Rosenshine, B., Meister, C., & Chapman, S. (1996) Teaching students to generate questions: A review of the intervention studies. Review of Educational Research, 66(2), 181-221.

+Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (1st ed., pp. 231-244). New York: Russell Sage.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. Educational Psychology Review, 7(4), 351-371.

+Seber, G. A. F. (1977). Linear regression analysis. New York: Wiley.

Slavin, R. E. (1983a). Cooperative learning. New York: Longman.

Slavin, R. E. (1983b). When does cooperative learning increase student achievement? Psychological Bulletin, 94, 429-445.

Snowman, J. (1994). Learning tactics and strategies. In G. D. Phye & T. Andre (Eds.), Cognitive classroom learning: Understanding, thinking and problem solving (pp. 243-276). New York: Academic Press.

Spires, H. (1990). Metacognition and reading: Implications for instruction. Reading, 24(3), 151-56.

Taylor, B. M., & Frye, B. J. (1992). Comprehension strategy instruction in the intermediate grades. Reading Research and Instruction, 92, 39-48.

Walker, D. F., & Schaffarzick, J. (1974). Comparing curricula. Review of Educational Research, 1, 83-111.

Webb, N. (1985). Student interaction and learning in small groups: A research summary. In R. E. Slavin, S. Kagan, R. Hertz-Lazarowitx, C. Webb, & R. Schmuck (Ed.), Learning to cooperate, cooperating to learn. New York: Plenum.

+White, H. D. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (1st ed., pp. 41-55). New York: Russell Sage.

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition. Educational Psychologist, 30(4), 217-222.

Primary Studies Evaluated

Andre, M. E. D. A., & Anderson, T. H. (1978). The development and evaluation of a self-questioning study technique. Reading Research Quarterly, 14(4), 605-23.

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

Bean, W. T., & Steenwyk, F. L. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. Journal of Reading Behavior, 16, 297-306.

Ben-Barka, A. C. (1984). Effects of induced schemata upon intermediate and advanced high school ESOL students' reading comprehension of selected expository passages.

Benito, Y. M. (1993). The effect of instruction in question-answer relationships and metacognition on social studies comprehension. Journal of Research in Reading, 16(1), 20-29.

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

Carr, E. M., Dewitz, P., & Patberg, J. P. (1983). The effect of inference training on children's comprehension of expository text. Journal of Reading Behavior, 15(3), 1-18.

Carr, M., & Thompson, H. (1995). Brief metacognitive intervention and interest as predictors of memory for text. ERIC document ED 387 791.

Carrasquillo, A., & Nunez, D. (1988). Computer-assisted metacognitive strategies and the reading comprehension skills of ESL elementary school students. ERIC document ED 301 838.

Chan, L. K. S., Cole, P. G., & Barfett, S. (1987). Comprehension monitoring: Detection and identification of text inconsistencies by LD and normal students. Learning Disability Quarterly, 10(2), 114-24.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. Journal of Educational Psychology, 78(4), 256-62.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M. S., Vavrus, L. G., Wesselman, R., Putnam, J., & Bassiri, D. (1987). Effects of explaining the reasoning associated with using reading strategies. Reading Research Quarterly, 22(3), 347-68.

Franklin, M. R., Roach, P. B., & Clary, E. J. (1993). Overcoming the reading comprehension barriers of expository texts. Educational Research Quarterly, 16(1), 5-14.

Gersten, R.(1989). Guided practice: Use of low-cost networking. ERIC document ED 310 784.

Golden, N., Gersten, R., & Woodward, J. (1990). Effectiveness of guided practice during remedial reading instruction: An application of computer-managed instruction. Elementary School Journal, 90(3), 291-304.

Gordon, C. J. (1980). The effects of instruction in metacomprehension and inferencing on children's comprehension abilities. Unpublished doctoral dissertation, University of Minnesota (University Microfilms International No. 8019528).

Hansen, J. (1981). The effects of inference training and practice on young children's reading comprehension. Reading Research Quarterly, 16(3), 391-417.

Hansen, J., & Pearson, P. D. (1983). An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. Journal of Educational Psychology, 75(6), 821-29.

Hodge, E. A. (1993). The effects of metacognitive training on the reading comprehension and vocabulary of at-risk college students. Research and Teaching in Developmental Education, 10(1), 31-42.

Kern, R. G. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. Modern Language Journal, 73(2), 135-49.

Lonberger, R. B. (1989). The effects of training in a self-generated learning strategy on the prose-processing abilities of fourth and sixth graders. ERIC document ED 310 374.

MacDonald, J. D. (1986). Self-generated questions and reading recall: Does training help? Contemporary Educational Psychology, 11(3), 290-304.

Magliano, J. P., Little, L. D., & Graesser, A. C. (1993). The impact of comprehension instruction on the calibration of comprehension. Reading Research and Instruction, 32(3), 49-63.

Meloth, M. S., & Deering, P. D. (1992). Effects of two cooperative conditions on peer-group discussions, reading comprehension, and metacognition. Contemporary Educational Psychology, 17(2), 175-93.

Miller, G. E., Giovenco, A., & Rentiers, K. A. (1987). Fostering comprehension monitoring in below average readers through self-instruction training. Journal of Reading Behavior, 19(4), 379-94.

Muniz Swicegood, M. (1994-1995). The effects of metacognitive reading strategy training on the reading performance and student reading analysis strategies of third grade bilingual students. Bilingual Research Journal, 18(1-2), 83-97.

Nolan, T. E. (1991). Self-questioning and prediction: Combining metacognitive strategies. Journal of Reading, 35(2), 132-38.

Nolte, R. Y., & Singer, H. (1985). Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. Reading Teacher, 39(1), 24-31.

Palincsar, A. S. (1987). Collaborating for collaborative learning of text comprehension. ERIC document ED 285 123.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, 1, 117-175.

Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. Child Development, 55(6), 2083-2093.

Patching, W., Kameenui, E., Carnise, D., Gerstem, R., & Colvin, G. (1983). Direct instruction in critical reading skills. Reading Research Quarterly, 18(4), 406-418.

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.

Schunk, D. H., & Rice, J. M. (1986). Comprehension strategy importance: Effects on remedial readers. ERIC document ED 268 491.

Singer, H., & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. Reading Research Quarterly, 17(2), 166-86.

Sjostrom, C. L., & Hare, V. C. (1984). Teaching high school students to identify main ideas in expository text. Journal of Educational Research, 78(2), 114-18.

Smith, B., Hynd, C., & Valeri-Gold, M. (1990). Increasing reading test performance: Modeling metacognitive strategies to poor college readers. Research and Teaching in Developmental Education, 7(1), 5-11.

Stevens, R. J. (1986). The effects of strategy training on the identification of the main idea of expository passages. ERIC document ED 297 263.

Stevens, R. J., Madden, N. A., Slavin, R., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. Reading Research Quarterly, 22(4), 433-54.

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1989). A cooperative learning approach to elementary reading and writing instruction: Long-term effects. ERIC document ED 328 901.

Taylor, B. M., & Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. Reading Research Quarterly, 19(2), 134-46.

Weaver, P. A. (1979). Improving reading comprehension: Effects of sentence organization instruction. Reading Research Quarterly, 15(1), 129-46.

Westera, J., & Moore, D. W. (1995). Reciprocal teaching of reading comprehension in a New Zealand high school. Psychology in the Schools, 32(3), 225-32.

CS 013226

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: SYNTHESIZING METACOGNITIVE INTERVENTION: WHAT TRAINING CHA

Author(s): CHRISTOPHER W.T. CHIU

Corporate Source: A PAPER PRESENTED TO THE SYMPOSIUM "METACOGNITION: ASSESSMENT AND TRAINING", at THE ANNUAL MEETING of THE AMERICAN EDUCATION RESEARCH ASSOCIATION, APRIL 13-17 1998.

Publication Date: 1998 APRIL

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, i reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 [✓]	Level 2A [ ]	Level 2B [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here,→ please

Signature: [signature]

Printed Name/Position/Title: CHRIS W.T. CHIU

Organization/Address: Michigan State University 1428 Haslett Rd Apt 104 E. Lansing MI 48823

Telephone: (517)337-7486

FAX:

E-Mail Address: CHIUWING@PILOT.MSU.EDU

Date: 4/30/98

(over)

ERIC®
Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

Tel: (800) 464-3742
(301) 405-7449
FAX: (301) 405-8134
ericae@ericae.net
http://ericae.net

March 20, 1998

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a printed copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at http://ericae.net.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (424)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:
AERA 1998/ERIC Acquisitions
University of Maryland
1129 Shriver Laboratory
College Park, MD 20742

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://aera.net). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

---

[1]If you are an AERA chair or discussant, please save this form for future use.