BRO05424

Improving Standardized Reading Comprehension: The Role of Question-Answering

Gail Brown, Herbert W. Marsh, Rhonda G. Craven, and Mary Cassar
SELF Research Centre*, *University of Western Sydney, Australia

Abstract

This paper provides empirical evidence that effective instruction in question-answering leads to statistically significant improvements in reading comprehension, when compared to regular classroom reading instruction. The presentation reports both the features of intervention materials and the differences in reading instruction between a treatment and control group that contributed to differences in posttest treatment group performance. The study involved a quasi-experimental, pretest-posttest design that targeted students enrolled in regular Year 5 classrooms across three schools. There were no statistically significant pretest differences between the treatment groups. Classroom teachers implemented the intervention with their classes over a ten week period. Comparisons were made between students who completed their regular classroom reading program and students who completed the intervention. Statistical analyses used multilevel modelling to ensure that adjustments were made for potential differences at the treatment group level and at the class level. Posttest comparisons on both a standardised reading comprehension measure and researcher-devised question-answering measures significantly favoured the intervention group. This presentation outlines the theoretical foundation and methodology for effective classroom instruction in question-answering. The potential future applications of this instructional technology to a range of complex cognitive skills are discussed.

The impact of reading comprehension on our daily lives has never been as crucial as in our modern society today. Individuals use literacy skills to communicate relationships between complex concepts and knowledge. Considerable literature has documented researchers on the state of reading (Kamil, Mosenthal, Pearson, & Barr, 2000; National Institute of Child Health and Human Development [NICHHD], 2000a) and reading comprehension (Muth, 1990; Pearson & Johnson, 1978; Pressley & Afflerbach, 1995). Research reviews have focussed specifically on reading comprehension and its instruction (Dole, Duffy, Roehler, & Pearson, 1991; Fielding & Pearson, 1994; Pearson & Fielding, 1991; Pressley, Brown, El-Dinary & Afflerbach, 1995; Rosenshine & Meister, 1994) and on results for specific student populations (Gersten, Fuchs, Williams, & Baker, 2001; Mastropieri & Scruggs, 1997; Weisberg, 1988). However, effective classroom instructional programs for reading comprehension are yet to be identified.

Reading is a complex, cognitive process: "…a whole complex system of skills and knowledge… knowledge and activities in visually recognising individual printed words are useless in and of themselves…" (Adams, 1990, p.3). Of particular relevance to this study were Adams additional comments that such decoding processes should be "guided and received by complementary knowledge and activities of language comprehension" (page 3). She implicitly supported the "simple view of reading" which outlined that reading was basically the product of decoding and comprehension processes (Hoover & Gough, 1990, p. 127; Hoover & Tunmer, 1993).

Taking this view further, the National Reading Panel identified four key components of reading skills: Phonemic skills, vocabulary, reading fluency and comprehension (National Institute of Child Health and Human Development [NICHHD], 2000a). Identification of these four key components was primarily based on research on beginning reading and the prevention of difficulties in learning to read (Abbott, Walton, & Greenwood, 2002; Denton, Vaughn & Fletcher, 2003; Elliott, Lee & Tollafson, 2001; Good III, Simmons & Kameenui, 2001; Kaminski & Good III, 1998; Snow, Burns, & Griffin, 1998). More recently, empirical data and descriptive analyses reflecting these four components has outlined different types of readers based on patterns of strength and weakness across reading accuracy, reading fluency, vocabulary knowledge and question-answering (Valencia & Buly, 2004). In addition, a complex picture of neurological functioning relevant to reading has emerged (Johnson, Hetzel & Collins, 2002). In the current study, these four components of reading were reflected in the measures used to determine the efficacy of the intervention program. These included standardised reading comprehension and vocabulary, oral reading fluency and written question-answering measures.

More specific to this study, researchers have called for instructional reforms in reading comprehension for decades (Ares & Peercy, 2003; Biemiller, 1994; Durkin, 1978-9; Schmidt, Rozendal, & Greenman, 2002; Simons, 1971; Thurlow, Ysseldyke, Wotruba, & Algozzine, 1993). However, two key limitations have persisted. Firstly, researchers have often failed to identify effective reading comprehension teaching strategies in sufficient detail to serve as an instructional program for classroom teachers. Instead, general methods of implementation have been suggested using terms such as "explicit instruction" (Pearson & Dole, 1987, p. 151), "thinking aloud" (Kucan & Beck, 1996, p. 259) and "direct instruction" (Carnine, Silbert, & Kameenui, 1997, p.1). This has met with considerable difficulty by teachers in attempting to translate these recommendations into instructional programs for use in classrooms.

Also, researchers have only loosely defined the comprehension curriculum in terms of specific comprehension (Dole et al., 1991; Fielding & Pearson, 1994;

Guszak, 1967; McNeil, 1987; Pearson & Fielding, 1991; Pressley, El-Dinary et al., 1992; Rosenshine & Meister, 1994). The National Reading Panel reviewed extant research for the purpose of improving classroom instruction in America (NICHHD, 2000a). Instruction in seven categories of specific reading comprehension skills, including question-answering instruction were reviewed. However, recommendations emanating from this report remained limited to general teaching strategies for classroom teachers.

There has been little consideration of the selection of teaching examples and the crucial role of student materials for improving classroom instruction in reading comprehension: that is, a specific instructional program for classrooms (Gersten et al., 2001). Historically, it would seem that teachers have been provided with suggestions for general methods for teaching comprehension and still have not been provided with either a clearly defined instructional program for the teaching of specific reading comprehension skills nor with appropriate classroom materials.

Reading comprehension is a cognitive process by nature. Recent theoretical advances, using information processing models, offer some promise for improving the efficacy of instructional interventions in reading comprehension research (Bransford, Brown, & Cocking, 2000; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Donovan, Bransford, & Pellegrino, 2001; Gordon, Hendrick, & Johnson, 2001; Shavelson & Towne, 2002). Information processing models utilise analogies between computer systems and human cognition. These models provide a theoretical basis for detailed analysis and simulation of complex cognitive tasks, including those found in classrooms and workplaces (Baddeley, Aggleton, & Conway, 2001; Kintsch, 1998; Miyake & Shah, 1999b; Newell, 1990; van Merrienboer & Paas, 2003). For example information processing models have providing insights into the cognitive processes used in decoding (Coltheart et al., 2001).

La Berge and Samuels (1974) applied the foundational concepts of automaticity and capacity limitations to reading. However, their model predominantly outlined the decoding processes of letter sounds. The model was of limited application to the current study where the focus was on question-answering as one reading comprehension skill. To date, there does not appear to have been a specific application of information processing models to the design of classroom materials that was focussed on how to answer questions.

Information processing models define two broad types of knowledge: declarative facts and procedural knowledge (Aitkenhead & Slack, 1985; Anderson, 1993; Hasselbring, Goin, & Bransford, 1988; Sieck & Yates, 2001; Sorace, Heycock, & Shillcock, 1999). Declarative facts are stored and retrieved more accurately and effortlessly depending on the strength of the relationship between a stimulus and a response (Hasselbring et al., 1988) or on the number of opportunities for practice with a particular stimulus (Logan, 1988). In the present study, declarative knowledge includes the meanings of words in questions and the types of questions taught. Procedural knowledge is defined as knowledge of sequences of steps in a strategy (Howell & Nolet, 2000; Pellegrino & Goldman, 1987). In the current study, procedural knowledge includes the strategy steps used for question-answering.

Over time and with practice, declarative and procedural knowledge practised and transformed into complex skill levels of expert performance (Bransford et al., 2000; De Corte, 2003; De Corte, Verschaffel, Enwistle & van Merrienboer, 2003; Engelmann & Carnine, 1982). Skill development is a gradual process that takes time and involves changes to efficient strategies (Anderson, 1982, 2002; Goldman, Mertz, & Pellegrino, 1989; Strayer & Kramer, 1994). The development of such expertise can

take a period of ten years or longer (Ericsson, Krampe, & Tesch-Romer, 1993). Everyday examples of such complex skills might include playing musical instruments, driving a car, playing a sport, and reading (Bransford, et al., 2000; Chaffin & Imreh, 2002; Ericsson, et al., 1993; Langan-Fox, Armstrong, Balvin & Anglim, 2002; Proctor & Dutta, 1995).

Attention and working memory are two broad constructs in information processing models that may impact on skill development. The design of the current intervention program first focuses on selecting examples with specific features that direct attention to features critical to concept learning (Engelmann, 1980; Engelmann & Carnine, 1982; Howell & Nolet, 2000; Thorley, 1987; van Merrienboer & Paas, 2003). The intervention program then gradually increases the difficulty of examples and their instructional context across lessons in order to take account of the working memory limitations in the completion of question-answering (Sweller, van Merrienboer, & Paas, 1998; van Merrienboer & Paas, 2003). Researchers have been able to decrease cognitive load by simplifying tasks and increasing task difficulty over time. Sweller and his colleagues have emphasised the importance of completing simple "part-task practice" as part of limiting the effects of working memory (van Merrienboer & Paas, 2003, p. 11). Completed examples have been shown to reduce working memory limitations and provide support to learners (van Merrienboer & Paas, 2003). These features are incorporated into the intervention program in the current study.

Question-answering is both a common indicator of reading comprehension and integral to our daily lives. Research has confirmed the importance of question-answering to classroom functioning, and specifically to reading comprehension (Andre, 1987; Armbruster, 1992; Beck, McKeown, Hamilton, & Kucan, 1997; Cazden, 1988; Guszak, 1967; Rickards, 1979; Weedman & Weedman, 2001). As such, it could be readily taught within a classroom instructional program. Despite this wealth of discussion and research, effective instructional programs for question-answering have not been evident. Durkin (1978-9) reported classroom teaching practices that predominantly involved repeated teacher assessments rather than instruction on how to comprehend the question.

By clearly linking questions with answers using passages of text, a small body of research has provided some insights for how students might approach the task of answering questions (Pearson & Johnson, 1978; Raphael, 1982). Pearson and Johnson's (1978) taxonomy of question-answer relations was based on reading theories that viewed text reading as an interactive process involving the text and the reader. Raphael's (1982) interpretation of Pearson and Johnson's (1978) taxonomy was utilised for the present study. This interpretation involved "Right There", "Think and Search" and "On My Own" question-answer relationships (p. 188).

Selection of Raphael's (1982) interpretation was based on age-appropriateness of the language for the participants in the study, use of terminology that indicated processing steps within the definitions, and the ease of translation of the "Right There" question definition for teaching particular examples that used one sentence in the text (see Appendix A, for definitions as used in the intervention). Studies by Raphael and her colleagues (Raphael, 1982, 1984; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985) had reported improvements in researcher developed measures but no corresponding improvements in standardised reading comprehension. In reading comprehension strategy instruction, some "powerful learning environments" (van Merrienboer & Paas, 2003, p.3) have documented significant gains in student performance, notably excluding standardised reading comprehension

measures (De Corte, Verschaffel & Van de Ven, 2001; De Corte, et al., 2003).

Previous reading comprehension research, including question-answering research, has been threatened by serious methodological flaws (Lysynchuk, Pressley, d'Ailly, Smith & Cake, 1989). Lysynchuk et al. (1989) reported weaknesses including questionable validity and reliability, limited empirical data, small sample sizes, specific types of participants and insufficient details of methodology to enable replication. The quasi-experimental design of the current study, reporting valid, reliable measures with inter-rater reliability and integrity of implementation data, addresses many methodological weaknesses of previous studies.

The primary purpose of the current study is to determine the effectiveness of a theoretically designed question-answering program to enhance standardised reading comprehension, and question-answering performance of Year 5 students using complex, statistical analyses. A secondary purpose of the current study is to develop an empirically validated question-answering program that can be readily implemented by classroom teachers. The specific hypotheses were that students who completed the question-answering intervention will demonstrate statistically significantly higher scores on a standardised measure of reading comprehension P.A.T.Comprehension) (Australian Council for Educational Research, 1986) and on measures of written question-answering on a narrative and a factual passage than Year 5 students who completed regular classroom reading programs.

Method

Design

The research design of the current study was a quasi-experimental pretest posttest design with intact year 5 classes. An experimental group of classes (n = 6) received the intervention program while a control group of classes (n = 4) continued with their regular classroom reading program. Year 5 classroom teachers in each school volunteered for the study and nominated whether they wished their class to be allocated to the experimental or the control group, within the constraint that there needed to be at least one class in each treatment group at each school. Hence, assignment to treatment groups was not determined by the researcher. This procedure resulted in a total of 167 students (92 males or 55%, 75 females or 45%) in the intervention group and 100 students (52 males or 52%, 48 females or 48%) in the control group. The mean age of the experimental group was 10 years 2 months while the mean age for the control group was 10 years 1.7 months. Pretest differences between the two treatment groups on outcome measures were not statistically significant and are reported in the Results section.

Participants

Participants were predominantly middle class suburban school students (n = 288) in Year 5 classes (n = 10) who attended three schools in metropolitan Sydney. A total of 288 students were enrolled in ten Year 5 classes in three schools, with five classes from School 1, three classes from School 2 and two classes from School 3. Twenty one students were excluded from data analyses. Reasons for exclusion included leaving the school, students absent for more than one week and special education students on individual programs provided outside the regular classroom.

The final sample comprised 267 students, 92.7% of the enrolled students, which included 144 males and 123 females. The ten classes had an average of almost 27 students in each class (sd = 2.2), with class sizes ranging between 22 and 31 students. The mean chronological age of these students was 10 years 2 months (sd = 4.9 months) with class averages ranging from 10 years, 1 month to 10 years 5 months. The ten classes included, on average, 14 boys and 12 girls in each class of almost 27.

Intervention Treatment

The intervention treatment comprised classroom teacher implementation of 30 lessons of student materials designed to teach students how to write answers to questions. Classroom teachers implemented the intervention to their whole class, during their regular reading time, at least three times per week. While a lesson time of no longer than 45 minutes was recommended, treatment integrity data showed that lessons implemented by experimental classroom teachers had an average of 48 minutes duration, with a range from 35 minutes to 75 minutes. A detailed description of the intervention has been reported elsewhere (Brown, 2004).

Implementation integrity data were collected using classroom observations and completed intervention materials. Student workbooks were collected regularly from the experimental treatment group and teachers were provided feedback on the number of examples completed independently by all students in each workbook. Integrity of implementation of the instructional program was ensured by these data. Summary data for work completed documented that, on average, students completed 225 answers to 57 passages, that comprised 90% of the 247 questions included in the intervention program (Brown, 2004). Informal classroom observations confirmed implementation occurred three times each week.

Control Group Treatment

All control teachers completed their usual classroom program in reading as if the research had not been occurring in their school. Control classes also completed pretesting and posttesting at the same time and in the same way as experimental classes. Classroom teachers selected six students from each control class. Work samples of all reading activities were copied by the researcher for these six students The six students were selected by the teachers as very competent, average competence and struggling readers in each control class. Descriptions of these work samples that comprised the control group reading programs were analyzed to determine the nature of the control treatment and reported elsewhere (Brown, 2004). Analyses of these work samples documented that control teachers presented a number of comprehension activities over the 30 lessons and a wide range of responses were documented across control classes.

Reading Measures

The reading measures used in the current study included a standardised reading comprehension measure and curriculum based reading comprehension measures in written question-answering form for the narrative and factual passages. Additionally, standardised reading vocabulary and oral reading fluency measures were used and are reported elsewhere (Brown, 2004).

Standardised Reading Comprehension Measure

One standardised measure "Progressive Achievement Tests in Reading: Comprehension" (Australian Council for Educational Research, 1986, p. 1) was used to measure reading comprehension. Form B was selected for the current investigation. Students silently read a series of short passages and responded to written multiple choice reading comprehension questions for each passage. This test included eight prose passages (between 200 and 300 words long), with two narrative passages, two descriptive passages and four expository passages. The Year 5 test comprised 41 multiple choice 21 factual and 20 inferential questions as defined in the manual. At the time this study was conducted the PAT Reading Comprehension was the most appropriate measure of reading comprehension for group administration at the time as it provided valid and reliable normative data.

Written Question-Answering Measures

Reading comprehension was also measured by two sets of written answers to questions about two text passages, one narrative and one factual. The narrative passage selected, entitled "Tropical" (see Appendix B), previously had been used as part of the Year 3 state-wide testing program (NSW Department of Education and Training, 1996). A passage designed as challenging for Year 3 students was deliberately selected for the current investigation in order to ensure that the passage could be readily decoded by Year 5 students participating in the present study. In consequence, decoding skills were likely to have minimal impact on written answers to questions. The passage, titled "Whales", was selected as the factual passage (see Appendix B) based on previous research examining "think aloud" protocols with Year 4 students (Kucan & Beck, 1996, p.259).

The readability of both passages was calculated using Flesch Reading Ease Score (Neibauer, 1998) and found to be 82 for the narrative passage and 74 for the factual passage. A score within the range of 70 - 80 was claimed to be graded fairly easily in relation to reading difficulty (Neibauer, 1998). Both passages were considered to be at an appropriate level of difficulty for Year 5 students and at a similar level.

For each passage, a set of questions of three types defined in previous research (Raphael, 1982) was written. The narrative passage had four questions written for each question type. The factual passage had ten questions written. These included three questions answered both directly from the passage and using background knowledge, and four questions answered using more than one sentence from the passage.

Written student responses were marked by the researcher using a standard marking guide. A trained research assistant marked a randomly selected sample of 53 student written answers (20% of the sample) against the marking guide. Inter-rater agreement for narrative answers and factual answers was calculated at 91.9% and 91.5% respectively. The passages and questions were considered to be typical of the content that might be used in Year 5 classrooms. The standardised administration and scoring of the answers, supported by inter-rater reliability data, ensured the reliability of the scores for written question-answering for all participants.

Data Analyses

The current study used multilevel modelling for most of the statistical analyses (Bryk & Raudenbush, 1992; Goldstein, 1995). The complex nature of the variables and data required multilevel modelling in order to take into account both the hierarchical nature of the sample data and the number of interdependent variables. Statistical analyses used the Multilevel Modelling Software, MlwiN, (Rasbash, Browne, Goldstein, Yang, Plewis, Healy, et al., 2002). Multilevel analyses used coefficients from regression equations to provide measures of effects, somewhat similar to *t* tests with statistical adjustments that controlled for data organised in successively larger groups (e.g., classes within schools). The complexity of the statistical models was elaborated at each stage of analysis by introducing treatment group and related measures variables. Models 1 & 2 examined pretest performance while Models 3, 4 & 5 examined posttest performance for treatment groups. Model 4 controlled for the effects of related dependent variables (reading vocabulary and reading fluency) and Model 5 controlled for the covariance effects of pretest performance on each dependent variable on posttest performance on that same dependent variable.

Results

Pretest Performance

Results for the simplest multilevel analysis indicated that there were no statistically significant differences between classes on any of the dependent variables at pretest. The lack of difference between classes was evident from the non-significant standard t values for the six dependent variables were all non-significant (all ps < 0.05) (Brown, 2004). Hence, there were no significant differences between the 10 classes on pretest data collected prior to the start of the intervention. Model 2 analyses tested more explicitly the assumption that there were no systematic pretest differences between the experimental and control groups, taking into account the hierarchical nature of the data. Standard t scores were calculated by comparing the fixed effects of the dichotomous experimental grouping variable with the corresponding standard error term. Consistent with the results of Model 1, there were no statistically significant differences between the experimental and control groups since all six t values corresponding to the six dependent variables were non-significant (all ps < 0.05). In summary, there were no statistically significant pretest differences between the experimental group and the control group on any of the three dependent variables.

Posttest Performance

Standardised Reading Comprehension Effects

Preliminary results indicated mean pretest reading comprehension performances for the experimental group and the control group of 59.3 percentiles and 62.9 percentiles, respectively, which were not significantly different (see Figure 1). Posttest mean reading comprehension performances for the experimental and control groups were 73.0 percentiles and 64.3 percentiles, respectively. Therefore, there was an increase in mean reading comprehension by 13.7 percentiles for the experimental group, compared to an increase of 1.4 percentiles for the control group. Additional multilevel analyses evaluated the statistical significance of differences between the two groups, controlling for pretest differences and taking into account the multilevel nature of the data. Model 3 and Model 4 confirmed the highly significant effect of the treatment.

Model 5 evaluated whether the difference between the experimental and control groups on posttest reading comprehension varied as a function of pretest levels of reading comprehension by introducing a term that represented the interaction between treatment group and pretest reading comprehension. A statistically significant standard t score was found for the interaction effect between treatment group and pretest reading comprehension performance ($t = 2.7$, $p < 0.01$). Students with lower pretest reading comprehension performance showed greater improvements in reading comprehension in response to the treatment than students with higher pretest reading comprehension.

Narrative Question-Answering

Mean posttest performance for the experimental group was 9.87 answers compared to 8.8 answers correct for the control group (see Figure 2). The increase in mean group performance from pretest to posttest was 1.41 answers correct for the experimental group compared to an increase in mean performance of 0.37 answers correct for the control group. Statistical significance of posttest differences, using multilevel analyses, reported a statistically significant effect ($t = 4.803$, $p < 0.001$). Model 4 confirmed the statistically significant effect for treatment group ($t = 6.392$, $p < 0.05$). Significant effects for pretest narrative answers and reading comprehension were reported ($t = 8.66$, $p < 0.05$ and $t = 2.0$, $p < 0.05$). The effects of reading vocabulary and narrative fluency were not significant.

A statistically significant interaction effect between treatment group and pretest narrative answers was evident in Model 5 ($t = 4.25$, $p < 0.05$). As for reading comprehension, the interaction between group and pretest narrative answers showed that, at lower pretest levels, student posttest performance was higher for the experimental treatment group than for the control treatment group.

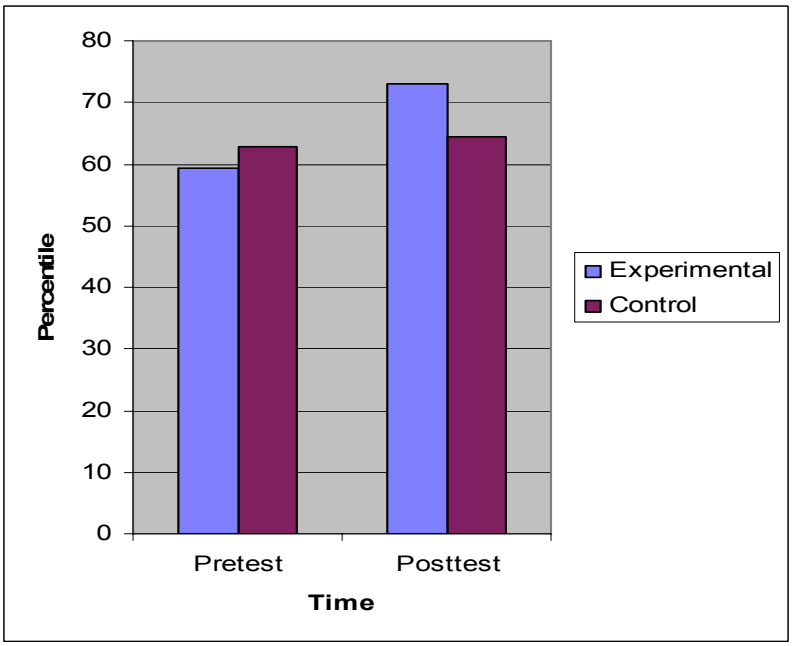

Figure 1. Mean treatment group pretest and posttest scores for standardised reading comprehension

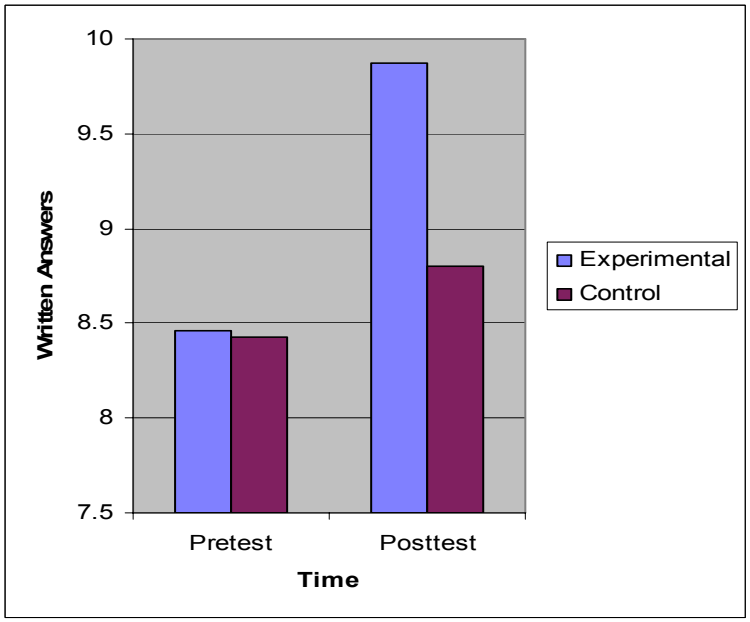

Figure 2. Mean treatment group pretest and posttest scores for narrative question-answering

Factual Question-Answering

Posttest mean performance for the two treatment groups reported a difference of 0.67 answers (see Figure 3). The posttest experimental group mean was 7.09 answers correct and posttest control group mean was 6.42 answers correct. Treatment group effects from all three multilevel models were statistically significant based on coefficients of the terms in the multilevel equations The introduction of the pretest variables into the multilevel equation in Model 4 confirmed significant effects for three of the four pretest variables, namely factual written answers, factual reading fluency and reading vocabulary. Pretest reading comprehension did not have a significant effect on posttest factual written answers, once the effects of the other pretest variables were controlled. Model 5 confirmed significant effects for treatment group and these three pretest variables. In addition, the interaction between treatment group was calculated as a non-significant interaction effect (t =1.7, p > .05).

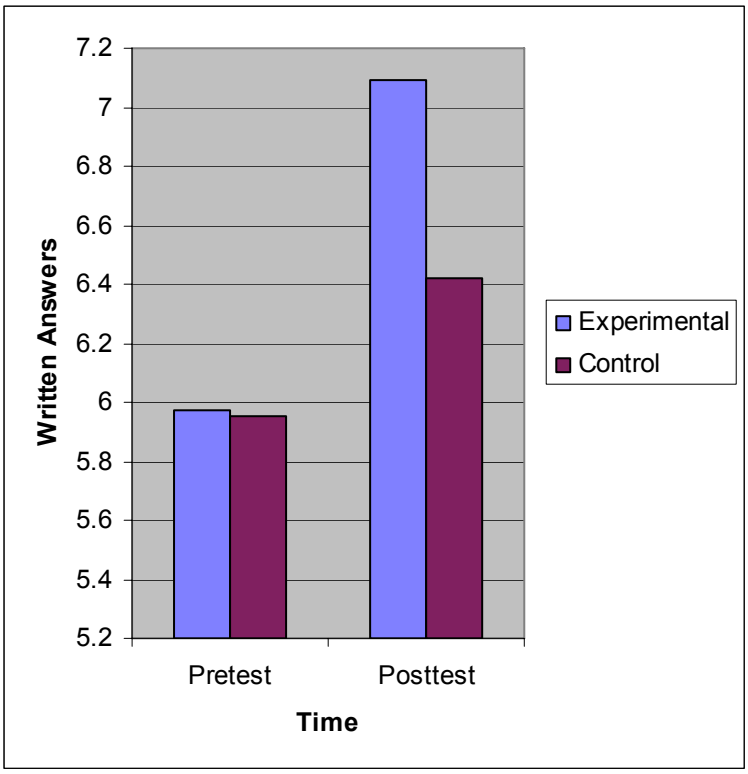

Figure 3. Mean treatment group pretest and posttest scores for factual question-answering

Discussion, Limitations and Conclusion

The ultimate goal of all reading comprehension instruction is to improve the performance of all students on standardised reading comprehension measures. To date, previous research has not documented such improvements for question-answering interventions (Benito, Foley, Lewis & Prescott, 1993; Ezell, Kohler, Jarzynka, & Strain, 1992; Graham, 1995; Graham & Wong, 1993; Raphael, 1982; Raphael, 1984, 1986; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985). In addition, methodological weaknesses and lack of sophisticated statistical methods have plagued previous research (Lysynchuk et al., 1989).

The current study found that question-answering instruction impacted positively and significantly on reading comprehension. The current study was

designed to avoid the pitfalls of previous research by using a sound research design, valid and reliable instruments and strong statistical tools. Previous results and advances in information processing models suggested that the development of a question-answering intervention based upon the best available theory might lead to a statistically significant improvement in standardised reading comprehension scores. The scientific design of the current study ruled out many competing causes of reading comprehension improvements and, therefore, supported the link between the intervention and performance improvements in the experimental group.

Therefore, statistically significant improvements in reading comprehension performance is attributed to the instructional design used for the question-answering intervention. Previous suggestions from strategy instruction research (De Corte et al., 2001; NICHHD, 2000a), and from question-answering interventions (Benito et al., 1993; Ezell, et al., 1992; Graham, 1995; Graham & Wong, 1993; Raphael, 1982, 1984, 1986; Raphael & McKinney, 1983; Raphael & Wonnacott, 1985) provided an incomplete analysis of the information and cognitive processing for designing classroom instructional programs. One weakness of earlier research was the lack of application of information processing models to the design of question-answering materials for classrooms. In applying information processing models to question-answering instruction, two broad principles were included: A detailed set of strategy steps for question-answering, and selection of controlled teaching examples for use with each strategy. These principles were not reported in previous studies of question-answering instruction, nor in the control programs in the current study, and are unique features of the current intervention.

In addition, the demand for research-based and effective classroom reading instruction for all students has never been stronger (De Corte, 2003; van Merrienboer & Paas, 2003). Recommendations for effective reading instruction have included directions in both decoding skills and in reading comprehension skills (Denton, et al., 2003; Johnson, et al., 2002; NICHHD, 2000b). While question-answering instruction has been included in these recommendations, we are not aware of any other classroom instructional program in question-answering that has empirically documented statistically significant improvements in standardised reading comprehension. The current study has reported such improvements and thereby offers an exciting and promising beginning for the development of effective intervention for the regular classroom. Integrity of implementation data, along with anecdotal teacher reports, supported the usefulness and appeal of the intervention in the current study for classroom teachers.

Statistically significant intervention effects were limited to reading comprehension measures. The research design included classes within the same school and year, and documented high levels of research control that suggested that posttest differences were a result of differences between the instructional programs, rather than other variables. Therefore, the results of the present study strongly suggest that these differences in the intervention materials used in experimental classes resulted in significant differences in posttest performance favouring those students.

In both experimental and control classes, whole class lessons predominated in classroom reading instruction. Classroom teachers provided all participants in their class with the same materials or activities and participants were provided with examples completed by the teacher. Control teachers presented one or two examples of concepts taught. The features of the intervention materials resulted in teacher modelling that involved presentation of a larger number of worked examples than in control classes. Across all classes, on average, similar lesson times were documented.

Hence, posttest performance differences were more likely attributed to differences in the features of the instructional programs.

Differences were documented between the two treatment groups in relation to the type of classroom reading instruction. The experimental group was presented with intervention materials that focussed explicitly and directly on written question-answering. Rather than focussing on a single comprehension skill, control teachers presented their classes with a wide variety of comprehension skills that included question-answering, advertisements, written chapter summaries, letter writing, descriptions, maze activities, cloze passages, story maps, plot profiles, directions, retelling, vocabulary instruction, listening comprehension, illustrations, poetry, research, grammar, character and cause and effect sequences. Consequently, there were fewer opportunities to practise any single comprehension skill to the same level as experimental participants had practised written question-answering.

These data confirmed the lack of interference in control classes during the study and ruled out possible threats to external validity due to experimental arrangements, such as Hawthorne effects (Tuckman, 1999), and the compensatory rivalry of John Henry Effects (Gall, Borg, & Gall, 1996). The lack of instruction specific to written question-answering in control classrooms confirmed "experimental treatment diffusion" was not a threat to internal validity (Gall et al., 1996, p. 471).

Strengths, Limitations, Conclusions and Future Directions

The current study examined a theory-based question-answering intervention with efficacy that, for the first time, extended to significant improvements on standardised reading measures for Year 5 students. A major strength of the current investigation was the establishment of an intervention that was effective in improving reading comprehension without a long period of teacher training. This was achieved through the reliance on the selection and sequencing of the teaching examples presented in the materials. The teaching examples not only established and controlled appropriate participant responses, but may have also provided scaffolding for classroom teachers where their knowledge of comprehension instruction may have been lacking. Hence, an important component of the current intervention was its presentation by classroom teachers, without the intrusion of scripted lesson presentations. This increases the external validity of the intervention and the likelihood of teacher acceptance of the current intervention in the longer term.

A clear change in the analyses used in the present study is the introduction of multilevel modelling approaches to evaluate the statistical significance of the effects of a question-answering intervention. Multilevel modelling has been applied to comprehension strategy instruction (DeCorte et al., 2001). However, in previous question-answering intervention studies, the multilevel, hierarchical nature of classroom data (i.e., students nested within classes) has been ignored. This failure of the existing question-answering research introduces a potentially large bias in tests of statistical significance in the direction of reporting differences to be statistically significant when they are not. The present investigation is one of the first to apply multilevel modelling to an intervention study, and the primary focus of statistical analyses was on treatment group effects. However, the current study highlighted the use of sophisticated multilevel modelling statistics for determining intervention efficacy. Potential applications of multilevel modelling procedures might examine other reading comprehension interventions and more detailed effects of specific classroom variables. No doubt this will be of interest to researchers in the future.

Conclusions from the current study were limited to Year 5 students enrolled in regular classrooms in Sydney metropolitan schools. Whether similar differences in

posttest performance would have resulted for students in other grades or in other schools was not investigated. In addition, conclusions were limited to the reading measures used in the current study that reflected the question-answering, reading fluency, standardised reading comprehension and reading vocabulary skills sampled in the measures that were used. Whether similar performance differences would have been reported on alternative standardised measures, sampling cloze passage, retelling or other comprehension skills, remains unknown.

Posttest performance differences were limited to measures tested on completion of the intervention program. Therefore, long term maintenance of skills and knowledge was not documented. Effects of additional practice were documented particularly for students with low reading skills through integrity of implementation data and interaction effects in reading comprehension and narrative question-answering. Such additional practice cannot be ruled out as a cause of interaction effects, and of performance differences between the treatment groups.

The question of which single component or combination of components of the intervention was more effective than other components cannot be determined in the current investigation. The intervention focussed on a synthesis of knowledge and cognitive processing that led to a complex set of materials and teaching strategies and changed the classroom environments during reading comprehension lessons. As with De Corte et al. (2001), there was no way to determine which specific components of the new classroom environment were effective. In addition, the purpose of the current study did not include examination of the complex interrelations between the reading measures used in reading comprehension, reading vocabulary, question-answering and reading fluency. However, none of these limitations detract from the power of the intervention for improving student performance in the reading comprehension measures used in the current study.

In summary, the current investigation documents the first reported statistically significant improvements in a standardised reading comprehension measure resulting from question-answering instruction. The importance of the design of classroom interventions based on a sound theoretical foundation in information processing models has been supported. Rather than a sole focus on general strategy instruction, the current investigation strongly supports the need for specially designed classroom materials that will firstly, foster the generalised use of effective reading comprehension strategies by all students, and secondly, be received and implemented by classroom teachers. Finally, the theoretical principles outlined in the current investigation, provide clear direction for textbook authors that could have implications for classroom practice in reading comprehension.

References

Abbott, M., Walton, C., & Greenwood, C. (2002). Research to practice: Phonemic awareness in kindergarten and first grade. *Teaching Exceptional Children, 34* (4), 20-26.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. University of Illinois: Centre for the Study of Reading.

Aitkenhead, A. M., & Slack, J. M. E. (1985). *Issues in cognitive modelling*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Andre, T. (1987). Questioning and learning from reading. *Questioning Exchange, 1*(1), 47-86.

Ares, N. M., & Peercy, M. M. (2003). Constructing literacy: how goals, activity systems, and text shape classroom practice. *Journal of Literacy Research, 35*, 633-662.

Armbruster, B. B. (1992). On answering questions. *The Reading Teacher, 45*, 724-725.

Australian Council for Educational Research. (1986). *Progressive achievement tests in reading teachers handbook* (2nd ed.). Hawthorn, Victoria: Australian Council for Educational Research.

Baddeley, A., Aggleton, J. P., & Conway, M. A. (2001). *Episodic memory: New directions in research*. Oxford, England: University Press.

Beck, I. L., McKeown, M. G., Hamilton, R. L., & Kucan, L. (1997). *Questioning the author*. Newark, DE: International Reading Association.

Benito, Y. M., Foley, C. L., Lewis, C. D., & Prescott, P. (1993). The effect of question-answer relationships and metacognition on social studies comprehension. *Journal of Research in Reading, 16*, 20-29.

Biemiller, A. (1994). Some observations on beginning reading instruction. *Educational Psychologist, 29*, 203-209.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*. Washington, D.C.: National Academy Press.

Brown, G. (2004). The efficacy of question-answering instruction for improving Year 5 reading comprehension, Unpublished Doctoral Dissertation, University of Western Sydney, New South Wales.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Mahwah, NJ: Erlbaum.

Carnine, D. W., Silbert, J., & Kameenui, E. J. (1997). *Direct instruction reading*. Upper Saddle River, NJ: Merrill.

Cazden, C. B. (1988). Classroom discourse and student learning. In C. B. Cazden (Ed.), *Classroom discourse: The language of teaching and learning* (pp. 99-120). New Hampshire, NJ: Heinemann.

Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science, 13*(4), 342-349.

Coltheart, M., Rastle, K., Perry, P., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

De Corte, E. (2003). Designing learning environments that foster the productive use of acquired knowledge and skills. In E. De Corte, L. Verschaffel, N. Entwistle & J. van Merrienboer (Eds.), *Powerful learning environments: Unravelling*

basic components and dimensions (pp. 21-34). Oxford, England, UK: Elsevier Science.

De Corte, E., Verschaffel, L., Entwistle, N., & Van Merrienboer, J. (2003). *Powerful learning environments: Unravelling basic components and dimensions*. Oxford, England: Elsevier Science.

DeCorte, E., Verschaffel, L., & VanDeVen, A. (2001). Improving text comprehension strategies in upper primary children: A design experiment. *The British Journal of Educational Psychology, 71*, 531-559.

Denton, C. A., Vaughn, S., & Fletcher, J. M. (2003). Bringing research-based practice in reading intervention to scale. *Learning Disabilities Research & Practice, 18*, 201-212.

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*, 239-264.

Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (2001). *How people learn: Bridging research and practice*. Washington, D.C.: National Academy Press.

Durkin, D. (1978-9). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*, 481-538.

Elliott, J., Lee, S. W., & Tollefson, N. (2001). A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills-Modified. *School Psychology Review, 30*, 33-49.

Engelmann, S. (1980). Toward the design of faultless instruction: The theoretical basis of concept analysis. *Educational Technology, 20,* 2, 28-36.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications*. New York: Irvington.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Ezell, H. K., Kohler, F. W., Jarzynka, M., & Strain, P. S. (1992). Use of peer-assisted procedures to teach QAR reading comprehension strategies to third grade children. *Education and Treatment of Children, 15*(3), 205-227.

Fielding, L. G., & Pearson, P. D. (1994). Synthesis of research reading comprehension: What works. *Educational Leadership, 51*, 62-68.

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. White Plains, NY: Longman.

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*, 279-320.

Goldman, S. R., Mertz, D. L., & Pellegrino, J. W. (1989). Individual differences in extended practice functions and solution strategies for basic addition facts. *Journal of Educational Psychology, 81*, 481-496.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Arnold.

Good III, R., Simmons, D. C., & Kameenui, E. J. (2001). The importance and decision-making utility of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13*, 425-430.

Graham, L. (1995). *The 3H strategy: Improving poor readers' comprehension of content area materials.* Paper presented at the National Conference of the

Australian Association of Special Education, Darwin, Australia.

Graham, L., & Wong, B. Y. L. (1993). Comparing two models of teaching a question-answering strategy for enhancing reading comprehension: didactic and self-instructional training. *Journal of Learning Disabilities, 26*, 270-279.

Guszak, F. J. (1967). Teacher questioning and reading. *The Reading Teacher, 21*, 237-246.

Hasselbring, T. S., Goin, L.I., & Bransford, J. D. (1988). Developing math automaticity in learning handicapped children: the role of computerised drill and practice. *Focus on Exceptional Children, 20*, 1-7.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An interdisciplinary Journal, 2*, 127-160.

Hoover, W. A., & Tunmer, W. E. (1993). The components of reading. In G. B. Thompson, W. E. Tunmer & T. Nicholson (Eds.), *Reading acquisition processes* (pp. 1-19). Great Britain: WBC Print Ltd.

Howell, K. W., & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision making* (3rd ed.). Belmont, CA: Thomas Learning.

Johnson, E. L., Hetzel, J., & Collins, S. (2002). Reading by design: Evolutionary psychology and the neuropsychology of reading. *Journal of Psychology and Theology, 30*, 3-25.

Kamil, M. L., Mosenthal, P. B., Pearson, P. D., & Barr, R. (2000). *Handbook of reading research: Volume III*. Mahwah, NJ: Erlbaum.

Kaminski, R. A., & Good III, R. H. (1998). Assessing early literacy skills in a problem-solving model: Dynamic indicators of basic early literacy skills. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 113-142). New York: Guilford.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.

Kucan, L., & Beck, I. L. (1996). Four fourth graders thinking aloud: An investigation of genre effects. *Journal of Literacy Research, 28*, 259-287.

La Berge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Langan-Fox, J., Armstrong, K., Balvin, N., & Anglim, J. (2002). Process in Skill Acquisition: Motivation, Interruptions, Memory, Affective States, and Metacognition. *Australian Psychologist, 37*, 104-117.

Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review, 95*, 492-527.

Lysynchuk, L. M., Pressley, M., d'Ailly, M., Smith, M., & Cake, H. (1989). A methodological analysis of experimental studies of reading comprehension strategy instruction. *Reading Research Quarterly, 24*, 458-470.

Mastropieri, M. A., & Scruggs, T. E. (1997). Best practices in promoting reading comprehension in students with learning disabilities. *Remedial and Special Education, 18*, 197-213.

McNeil, J. D. (1987). *Reading comprehension: New directions for classroom practice*. Chicago: Scott Foresman.

Miyake, A., & Shah, P. (Eds.). (1999b). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge, England: Cambridge University Press.

Muth, K. D. (1990). *Children's comprehension of text: Research into practice*. Newark, DE: International Reading Association.

National Institute of Child Health and Human Development. (2000a). *Report of the*

National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, D.C.: U.S. Government Printing Office.

National Institute of Child Health and Human Development. (2000b). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington D.C.: U.S. Government Printing Office.

Neibauer, A. (1998). *Corel Wordperfect Suite 8 Professional: The official guide*. Berkeley, CA: McGraw-Hill.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

NSW Department of Education. (1996). *Basic Skills Statewide Test*. Sydney, Australia: NSW Department of Education.

Pearson, P. D., & Dole, J. A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *The Elementary School Journal, 88*, 151-165.

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research: Volume II* (pp. 815-861). White Plains, NY: Longman.

Pearson, P. D., & Johnson. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart & Winston.

Pellegrino, J. W., & Goldman, S. R. (1987). Information processing and elementary mathematics. *Journal of Learning Disabilities, 20*, 23-32.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

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*, 215-224.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance*. Beverly Hills, CA: Sage.

Raphael, T. (1982). Question-answering strategies for children. *The Reading Teacher, 36*, 186-190.

Raphael, T. E. (1984). Teaching learners about sources of information for answering comprehension questions. *Journal of Reading, 27*, 303-311.

Raphael, T. E. (1986). Teaching question answer relationships, revisited. *The Reading Teacher, 39*, 516-522.

Raphael, T. E., & McKinney. (1983). An examination of fifth- and eighth-grade children's question-answering behavior: An instructional study in metacognition. *Journal of Reading Behavior, 15*, 67-86.

Raphael, T. E., & Wonnacott, C. A. (1985). Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly, 20*, 282-296.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., et al. (2002). *A user's guide to MLwiN: Version 2.1d for use with MLwiN1.10*. London: Centre for Multilevel Modelling, Institute of Education, University of London.

Rickards, J. P. (1979). Adjunct postquestions in text: A critical review of methods and processes. *Review of Educational Research, 49*, 181-196.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*, 479-530.

Schmidt, R. J., Rozendal, M. S., & Greenman, G. G. (2002). Reading instruction in the inclusion classroom. *Remedial and Special Education, 23*,130-140.

Shavelson, R. J., & Towne, L. E. (2002). *Scientific research in education*. Washington, D.C.: National Academy Press.

Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.

Sieck, W. R., & Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1003-1021.

Simons, H. D. (1971). Reading comprehension: The need for a new perspective. *Reading Research Quarterly, 6*, 338-363.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, D.C.: National Academy Press.

Sorace, A., Heycock, C., & Shillcock, R. (1999). *Language acquisition : Knowledge representation and processing*. Oxford, England: Elsevier Science.

Strayer, D. L., & Kramer, A. F. (1994). Strategies and automaticity: 1. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 318-341.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Thorley, B. J. (1987). *Learning and special education: Collected works*. Sydney, Australia: Macquarie University.

Thurlow, M. L., Ysseldyke, J. E., Wotruba, J. W., & Algozzine, B. (1993). Instruction in special education classrooms under varying student-teacher ratios. *Elementary School Journal, 93*, 305-320.

Tuckman, B. W. (1999). *Conducting educational research* (5th ed.). Orlando, GA: Harcourt Brace College.

Valencia, S.W. & Buley, M.R. (2004). Behind test scores: What struggling readers really need. The Reading Teacher, 57, 520-531.

van Merrienboer, J. J. G., & Paas, F. (2003). Powerful learning and the many faces of instructional design: Toward a framework for the design of powerful learning environments. In E. De Corte, L. Verschaffel, N. Entwistle & J. van Merrienboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 3-20). Oxford, England: Elsevier Science.

Weedman, D. L., & Weedman, M. C. (2001). When questions are the answer. *Principal Leadership, 2*(2), 42-46.

Weisberg, R. (1988). 1980s: A change in focus of reading comprehension research: A review of reading/learning disabilities research based on an interactive reading model. *Learning Disability Quarterly, 11*, 149-159.

APPENDICES

Appendix A Question-Answering Framework

SOURCE 1: RIGHT THERE QUESTIONS Right There Questions are those where the answer is **right there in one sentence.** There is only **one correct answer** to a Right There Question.
SOURCE 2: THINK & SEARCH QUESTIONS Think & Search Questions are where the answer is in **more than one sentence** or the answer uses **the story & your thinking together** to give a complete answer There might be **more than one** complete and correct answer to a Think & Search question.
SOURCE 3: ON MY OWN QUESTIONS On My Own Questions are where the answer requires you to think of the answer **on your own** - there are **few** (sometimes no) clues in the story. You need to **think about what you already know about the topic and the story**, and write an answer that **fits in with both** the story and what you already knew. There can be **more than one** answer to On My Own questions

Appendix B: Passages used for Narrative and Factual Question-Answering

Passage 1: Tropical

Tropical Paradise Friday PM

Keith sprinted out of the hardware store, paint cans thumping together in his school bag.

The clock on the war memorial across the street said eight minutes past eleven. Keith stared. Then he remembered it had been wrong ever since a coconut had hit it in a cyclone.

He looked at his watch. Nineteen minutes to four. Two hours and twenty-four minutes left.

He should just make it.

As long as Mum and Dad didn=t see him.

Keith decided he=d better not risk going too close to the shop so he ran across the road, through the fringe of palm trees and onto the beach. He ran along the soft sand, trying to look like a tourist out for a jog with a couple of tins of paint in a school bag.

He glanced through the palm trees at the shop. Mum and dad were both behind the counter but neither of them was looking in his direction. They were looking at each other. Dad was saying something to Mum, pointing at her with a piece of fish, and Mum was saying something back, waving the chip scoop at him.

Even at that distance, Keith could see that Dad=s mouth was droopier than a palm frond and that Mum=s forehead had more furrows in it than we sand when the sea is choppy.

Keith=s stomach knotted even tighter. Another argument. Poor things. Stuck in a fish-and-chip shop all day in this heat. Anyone=d get a bit irritable standing over a fryer all day with this poxy sun pounding down nonstop.

The trouble with tropical paradises, thought Keith, as he ran along the beach, is that there=s too much good weather.

Passage 2: Whales
WHALES

There are about ninety kinds of whales in the world. Scientists have divided them into two main groups: toothed whales and baleen whales.

Toothed whales have teeth and feed mostly on fish and squid. They have only one blowhole and are closely related to dolphins and porpoises.

The sperm whale is the only giant among the toothed whales. It is the animal that comes to mind when most people think of a whale. A sperm whale has a huge, squarish head, small eyes, and a thick lower jaw. The male grow to about sixty feet long and weighs as much as fifty tons. The female is smaller, reaching only forty feet and weighing less than twenty tons.

A sperm whale=s main food is squid, which it catches and swallows whole. A sperm whale is not a very fast swimmer, but it is a champion diver. It dives to depths of a mile in search of giant squid and can stay underwater for more than an hour.

There are smaller and less familiar kinds of toothed whales. The narwhal is a leopard-spotted whale about fifteen feet long. It is sometimes called the unicorn whale, because the male narwhal has a single tusk. The tusk is actually a ten foot long front tooth that grows through the upper lip and sticks straight out. No one knows for sure how the narwhal uses its tusk. Narwhals live along the edge of the sea in the Arctic.

Perhaps the best known of the toothed whales is the killer whale. That=s because there are killer whales that perform in marine parks around the country. A killer whale is actually the largest member of the dolphin family. A male can grow to over thirty feet and weigh nine tons.

Orcas are found in all the world=s oceans, from the poles to the tropics. They hunt for food in herds called pods. Orcas eat fish, squid, and penguins, as well as seals, sea lions, and other sea mammals, including even the largest whales. Yet they usually appear gentle in captivity, and there is no record that an orca has ever caused a human death.