Critical Review: The use of reciprocal teaching to improve reading comprehension of both normal-learning and learning disabled individuals in the reading to learn stage.

Park, H.
M.Cl.Sc. (SLP) Candidate
School of Communication Sciences and Disorders, U.W.O.

This critical review examines whether implementing a reciprocal teaching program will improve the reading comprehension of participants in the reading to learn stage. Study designs include: a systematic review, experimental, and quasi-experimental designs. Overall, research supports that reciprocal teaching can improve the reading comprehension of participants in the reading to learn stage, including adults and learning disabled participants.

Introduction

In the past 28 years there has been a surge in empirical evidence on the positive association between metacognitive knowledge and reading comprehension (Kelly, Moore, & Tuck, 1994). Metacognition is "the ability to plan, organize, and reflect on our own cognitive strategies" (Paul, 2001, p. 536). Comparatively, comprehension monitoring strategies consist of "evaluating the issues or failure of the meaning-making process [(comprehension)] and selecting strategies to remedy comprehension problems" (Irwin, 2007, p. 125).

Palincsar and Brown (1984) discuss a teaching method which aims to equip subjects with techniques to increase the use of metacognition and reading comprehension strategies. Their program is called reciprocal teaching (RT). RT is an "instructional technique in which reading comprehension is viewed as a problem-solving activity in which thinking is promoted while reading" (Glaser, 1990, p. 30).

Palincsar and Brown (1984) explain that comprehension (assuming adequate decoding ability) is the product of "(1) considerate texts, (2) the compatibility of the reader's knowledge and text content, and (3) the active strategies the reader employs to enhance understanding and retention, and to circumvent comprehension failures" (p. 118). This illustrates that metacognitive factors affect reading comprehension, because while one reads, s/he must attend to his/her comprehension of the material. Palincsar and Brown (1984) elected to study the reading strategy element of comprehension. Through a literature review of traditional reading education and theoretical treatments, the authors developed four "comprehension-fostering and comprehension-

monitoring" strategies: summarizing, questioning, clarifying and predicting (p. 121). Summarizing helps one to monitor his/her progress by finding and retaining information. Questioning text meaning leads to active monitoring of ones own comprehension. Clarification and prediction improve comprehension and help one to monitor his/her own comprehension.

Hart and Speece (1998) explain "the goal of [RT] is to improve students' skill in independently comprehending text" (p. 671). RT has two major features. (1) Instruction and practice of the four comprehension-fostering strategies. During this stage, the teacher models these strategies, and may use a 'think-aloud model', illustrating why they are used as well as the mental processes involved in their use (Alfassi, 2004, p. 172). (2) Dialogue between the teacher and student acts as a vehicle for learning and practicing these four strategies (Rosenshine & Meister, 1994). It also acts to focus on "planning, implementing and evaluating the strategies during the discussion of text" (Hart & Speece, 1998, p. 671). Dialogue allows novices to learn from the contributors of more capable peers and it exposes the learner to various points of view which may clarify his/her initial understanding (Alfassi, 2004). Heterogeneous grouping by age or reading ability may maximize the value of RT by providing effective peer models for poor comprehenders. Great emphasis is also placed on encouraging students to provide instructional support for each other. Traditionally the direct instruction can be done with the entire class, and the dialogue is carried out within smaller class groups (Alfassi, 2004).

RT is considered better than explicit teaching or instructional methods alone, in which transfer or generalization effects are rarely found

(Palincsar & Brown, 1984). RT takes into consideration Vygotsky's developmental theory and the benefits of expert scaffolding and proleptic teaching. Firstly, Vygotsky describes a child's "zone of proximal development" as the difference between what a child can accomplish unaided and what s/he can accomplish with the help of a more knowledgeable person (Palincsar & Brown, 1984, p. 123). "Vygotsky believed that a great deal of development was mediated by expert scaffolding" (Palincsar & Brown, 1984, p. 123). Expert scaffolding plays a significant role in RT. In "the early stages of RT, the instructor assumes the major responsibility for instruction by explicitly modeling the four strategies. After the initial stage…the students take turns leading the group dialogue and practicing the strategies on other sections of text. At that stage, the teacher becomes a mediator who provides guidance and feedback tailored to the needs of the current dialogue leader and his or her respondents" (Palincsar & Brown, 1989, as cited by Alfassi, 2004, p.172). As the students use the four strategies more competently, the teacher begins to diminish his/her scaffolded assistance.

Lastly, "proleptic means "in anticipation of competence," and in the context of instruction refers to situations where a novice is encouraged to participate in a group activity before she is able to perform unaided, the social context supporting the individual's efforts" (Palincsar & Brown, 1984, p. 123). This means the novice performs a simple task while s/he observes and learns from an expert model.

Many studies, including Palincsar and Brown's (1984), have found that young readers and poor readers do not use effective strategies for monitoring and constructing meaning from text. Poor readers include individuals who have language disorders which is an "impaired comprehension and/or use of spoken, written and/or other symbol systems" (Larson, 2003, p. 1). However, "experimental studies have also clearly shown that any student can be taught these higher order skills, and that significant gains in students' reading comprehension may be brought about through such explicit metacognitive instruction" (Haller, Child, & Walberg, 1988, as cited by Kelly et al., 1994, p. 53). RT has been proven to successfully increase ones metacognition and reading comprehension (Palincsar & Brown, 1984).

Objectives

The primary objective of this paper is to critically evaluate existing literature regarding the influence of RT on improving the reading comprehension of individuals in the reading to learn stage, including young children and adults as well as normal-learning and learning disabled individuals. The secondary objective is to propose evidence-based recommendations for future research and implications for the use of reciprocal teaching programs.

Methods

Search Strategy

Computerized databases, including Cochrane Library (current), Eric Plus Text, PsycINFO, PubMed, MEDLINE--OVID, ProQuest Education Journals (ProQuest Education complete) and EBSCO Host Research Databases were searched using the following strategies: Search 1: ((reciprocal reading) OR (reciprocal teaching)) AND ((reading comprehension) OR (comprehension)). Search 2: ((reciprocal teaching) OR (reciprocal reading)) AND (metacognition) AND (reading) AND (comprehension)

The search was limited to articles published in English between 1993 and 2007. Reference lists of articles were hand searched for further relevant studies.

Selection Criteria

Research studies selected for inclusion in this critical review paper were required to examine the effects of reciprocal teaching on the comprehension of individuals in the reading to learn stage. Therefore, participants in RT programs were required to be older than eight years, or beyond grade three. Studies were to include expert scaffolding, as well as, Palincsar and Brown's four strategies to improve comprehension.

Data Collection

Results of the literature search yielded the following six studies: one systematic review, five experimental/ quasi-experimental controlled trials.

Results

The following studies are ranked in order of credibility.

Rosenshine and Meister (1994) conducted a systematic review of studies to determine the overall effects of RT interventions. The reviewed literature included articles obtained through ERIC and Dissertation Abstracts International databases, as well as programs from the annual meetings of the American Educational Research Association. Articles were required to meet stringent inclusion criteria and had to consist of comparable experimental and control groups. Each study was rated in quality. The median effect size was .32 when standardized tests were used to measure comprehension and .88 when experimenter developed tests were used. Effect size by type of student and type of test was also measured. Students good in decoding but poor in comprehension obtained an effect size of .29 on standardized tests and .88 on experimenter developed tests. Below average students obtained an effect size of .08 on standardized tests and 1.15 on experimenter developed tests. The authors noted effective results by means of experimenter-developed outcome measures, regardless of grade level, number of instructional sessions, class size, number of strategies taught, or whether a teacher or experimenter provided the instruction. However, they did find a significant result by type of student interaction, which may have been due to type of measure used. It was noted that experimenter developed tests were used more often with participants who were good at decoding but poor in reading comprehension. These tests revealed significant results. If experimenter generated tests were shown to produce more significant results in general, then the type of test chosen to assess comprehension could have skewed the test results. This suggests that different results may have been obtained if different tests were used on the aforementioned population. The type of test measure could have skewed the results because experimenter developed tests produced more significant results than standardized tests, as the former is easier in nature. Another weakness was the inclusion of unpublished data and non-peer reviewed data. These studies may not have been published for methodological reasons which may have skewed the review's results.

Alfassi (2004) carried out two studies, study one was utilized in the present review. Alfassi (2004) hypothesized that students exposed to RT incorporated within a language arts class would show greater improvements in reading comprehension than would students exposed to only traditional methods of literacy instruction and immersion. A comparative study was carried out. Two equivalent mainstream freshman classes of good readers were randomly assigned to a condition, with an experimental group (RT) consisting of 29 subjects, and a control group (traditional literacy instruction) of 20 participants. Equivalent teachers taught the lessons and received six hours of training. Intervention lasted 20 days and was outlined well, however, fidelity of treatment was not measured. Both groups were assessed pre-, throughout, and post- intervention and maintenance testing was completed. Experimenter developed comprehension questions were used and rated independently, generating a chronbach's alpha of .71 to .85. Participants were also assessed using a standardized test. No effect size was given. A MANCOVA was carried out with post testing, revealing a significant difference favouring the experimental group on reading assessments and standardized measures. The experimental group significantly improved, both experimenter developed and standardized testing showed significant changes between pre- and post-testing. Therefore, the educational benefits of incorporating RT into the English Language arts curriculum were verified.

Lovett, Borden, Warren-Chaplin, Lacerenza, DeLuca, & Giovinazzo (1996) conducted a controlled comparison of two different approaches, Text Content and Structure Program (TCS) and RT, to train text comprehension skills of a group of adolescents (grade 7/8) with multifaceted reading disabilities. Matched pairs participants were randomly assigned to a program, teacher, and instructional test. There were two experimental groups with 16 participants and a control group with 14 participants. All participants (33 boys and 13 girls) were poor readers (<25th percentile) and 37% were deficient in word identification. The 25 day intervention was led by special education teachers who taught pairs of children matched for reading level. The intervention was not well explained and fidelity of treatment was not reported. Pre- and post-testing occurred using both experimenter developed and standardized tests. However, no assessments occurred during the study to measure progress. Assessment of both *taught to* (TT) and *not taught to* (NTT) materials occurred. Three types of comprehension measures were considered

separately using a MANCOVA, ANCOVA, and univariate ANCOVA with post hoc Tukey tests. Results indicated that program effects were large for the RT group and attributable to post test superiority of RT trained group on both TT and NTT texts. Effect sizes of RT on both TT and NTT texts ranged from medium to large on the 4 strategies. Transfer effects for the 4 strategies were evident.

Lederer (2000) conducted a study illustrating the effectiveness of RT on text comprehension in social studies classes. He found that the use of scaffolding approaches, such as RT, in general educational classrooms can educate both regular and LD students. This experimental study had control groups but lacked randomization which may bias effects and reduce reproducibility. Participants were students in inclusive classrooms, grades 4-6, with approximately 5/22 LD students per classroom. During the 17 day intervention the researcher served as the principle teacher and class participants were split into several working groups. The intervention was not described well and fidelity of treatment is unknown. Further, the way in which the leader led the sessions was not analyzed, therefore, there could have been significant differences in instruction between groups, which may have skewed the results. Nonstandardized experimenter developed comprehension assessments were conducted pre- , throughout and post-intervention and 20% of the assessments were scored independently by two different raters. Interrater agreement was 94.55%. A MANOVA and Tukey's post hoc testing found a significant change in reading comprehension for treatment and as a function of grade. Also a significant improvement was seen in the experimental group's ability to answer questions, generate questions, and compose summaries. A specific comparison between LD's in experimental and control groups was carried out using an ANOVA. No significant difference was found in the ability to generate questions, but a significant difference was found in the ability to compose summaries. An independent-test indicated that the experimental classes at the 4th and 5th grade performed higher, but not significantly higher than the control groups. The 4th and 6th grade experimental classes demonstrated significant comprehension gains 30 days post intervention. Lederer explained possible teacher confounds for grade five. Medium to large effect sizes were found for answering questions, generating questions, and

creating summaries. A problem with this study was that it only assessed the use of answering and generating questions and summarizing.

Hart & Speece (1998) examined the effects of RT on reading comprehension in post secondary developmental reading programs. They conducted a quasiexperimental, nonequivocal control group design with pre- and post-testing. The intervention was led by the researchers, which may have biased the results. The participants were below average readers in community college. Pre- and post-test scores were obtained from standardized and nonstandardized testing. Prior to test scoring, high interrater reliability of .80 was established. After scoring was completed, 25% of the assessments were randomly selected to check reliability, revealing a coefficient of .80. Various statistical analyses were carried out consisting of a Wilcoxon-Mann-Whitney test, MANCOVA, ANOVA, and Tukey's honest difference procedure. The analyses revealed a significant difference favouring the experimental groups on post-test scores on all four strategies with an effect size of .30. Post test scores revealed a significant difference between scores on all four strategies for poor readers, and on questioning for better readers and a significant difference between pre- and post-testing on all strategies. This study had a high treatment fidelity index of 95% and 93%. A weakness was that this study had a few drop outs and the authors did not indicate what was done to account for the missing data.

Kelly, Moore and Tuck (1994) assessed the effects of RT with the four strategies on poor reading comprehenders in a regular classroom setting. A multiple baselines across group design with between phase comparisons was utilized. Participants were proposively chosen by academic delays. Twelve participants in the experimental group (received RT) were in two parallel standard three and four classrooms. The attention only comparison group consisted of six students in one class who were performing at average or above average levels. Comprehension probes occurred daily, however, the three groups did not receive identical testing throughout the sessions. Therefore, between group comparisons may have been limited. The authors did not address this issue. As well the sample size was small, therefore, the results may not be replicable. No effect size was reported. A

strength of this study was that participants received frequent testing and interrater reliability of assessments revealed a mean of 98% agreement. Fidelity was addressed by examining audiotape recordings of six baseline and 18 intervention sessions with experimental group one for all instances where one of the strategies was used. However fidelity index was unknown.

Recommendations

Several concerns were apparent regarding the design and methodology of the research reviewed. It is recommended that future research focus on the following to provide higher levels of evidence: randomly assign participants to groups; outline intervention clearly to increase reproducibility; assess and document implementation to improve fidelity; assess quality of leader-student dialogues because of the importance of scaffolding; assess post-test measures on all four strategies; blind raters who are analyzing the intervention and assessments; and test for transfer and generalizability.

It is hypothesized that if RT is implemented by a Speech-Language Pathologist (S-LP), participant comprehension may increase. This is because the success of RT depends somewhat on a teacher's ability to provide an "*on-line diagnosis* that will guide her own level of participation, a level of participation that is finely tuned to the student's changing cognitive status" (Palincsar & Brown, 1984, p. 169). S-LP's have considerable training and experience in remediation for learning disabilities, comprehension processes, comprehension breakdowns, and scaffolding. Therefore, S-LP's may be well-attuned to the student's needs and posses the ability to competently scaffold accordingly.

Conclusion

In general, the studies included in this review are relatively strong. Minor weaknesses were evident in the lack of proper randomization and in describing the intervention, which created unknown treatment fidelity. According to Dollaghan (2007), a well known researcher and teacher of evidence-based practice, the information in this analysis is overall compelling for both validity and importance. According to

the Oxford Centre for Evidence-Based Medicine (2001), the study designs lend to a high level of evidence (levels 1 and 2), which strongly supports the hypothesis (grade A/B). Therefore, the present research findings suggest that implementing the use of RT in the classroom setting or in small groups can increase the reading comprehension of participants of various ages who are either normal-learners or learning disabled.

References

Alfassi, M. (2004). Reading to Learn: Effects of Combined Strategy Instruction on High School Students. *Journal of Educational Research*. *97*(4), 171-184. Retrieved November 14, 2007, from EBSCO Host Research Databases.

Billingsley, B.S. (1993). Strategies to Facilitate Reading Comprehension In students with Learning Disabilities. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 9, 263-27.*

Brown, A.L., & Campione, J. (1986). Psychological theory and the study of learning disabilities. *America Psychologist*, 14, 1059-1068.

Dollaghan, C. A. (2007). The Handbook for Evidence-Based Practice in Communication Disorders. Baltimore: Paul H. Brookes Publishing Co.

Glaser, R, (1990). The reemergence of learning theory within instruction at research. *American Psychologist*. 45, 29-39.

Hart, E. R., & Speece, D. L. (1998). Reciprocal Teaching Goes to College: Effects for Postsecondary Students at Risk for Academic Failure. *Journal of Educational Psychology*. *90*(4), 670-681. Retrieved October 23, 2007, from Psych INFO database.

Irwin, J. W. (2007). *Teaching Reading Comprehension Processes Third Edition*. Boston. Pearson Ed. Inc.

Kelly, M., Moore, D. W., & Tuck, B. F. (1994). Reciprocal Teaching in a Regular Primary School Classroom. *Journal of Educational Research*. *88*(1) 53-61. Retrieved November 14, 2007, from EBSCO Host Research Databases.

Larson, B. & McKinley, N. (2003). C*ommunication solutions for Older Student: Assessment and Intervention Strategies.* Eau Claire, WI: Thinking Publications.

Lederer, J.M. (2000). Reciprocal Teaching of Social Studies in Inclusive Elementary Classrooms. *Journal of Learning Disabilities. 33* (1), 91-106. Retrieved on October 23, 2007, from MEDLINE--OVID database.

Levels of Evidence and Grades of Recommendation. *Oxford Centre for Evidence-Based Medicine.* Retrieved May 2001, from http://www.cebm.net/levels_of_evidence.

Lovett, M. W., Borden, S. L., Warren-Chaplin, P. M., Lacerenza, L., DeLuca, T., & Giovinazzo, R. (1996). Text Comprehension Training for Disabled Readers: An Evaluation of Reciprocal Teaching and Text Analysis Training Programs. *Brain and Language.* 54, 447-480. Retried October 23, 2007, from MEDLINE--OVID database.

Palincsar, A. S., & Brown, A, L. (1984). Reciprocal Teaching of Comprehension Fostering and Comprehension-Monitoring Activities. Cognition and Instruction, *1*(2), 117-175.

Paul, R. (2001). *Language disorders from infancy to adolescence: Assessment and intervention.* St. Louis: Mosby.

Rosenshine, B., & Meister, C. (1994). Reciprocal Teaching: A Review of the Research. *Review of Educational Research. 64*(4), 479-530. Retrieved November 1, 2007, from Eric Plus Text database.