The Impact of *Thinking Reader* Software Program on Grade 6 Reading Vocabulary, Comprehension, Strategies, and Motivation

**AUTHORS:** Marjorie Chinen, Lindsay Fryer, Kathryn Drummond, Teresa Garcia Duncan, and H. Ray Miller

**ABSTRACT**
This paper presents final results from the *Thinking Reader* study, a randomized controlled trial (RCT) sponsored by the Institute of Education Sciences (IES). *Thinking Reader* is a software program intended to improve the reading comprehension of middle-schoolers by having students read novels on computers and respond to prompts intended to teach students different strategies for understanding text. *Thinking Reader* provides tools that teachers can use to monitor student progress and provide feedback. The study includes sixth-graders and assesses the impact of the program on reading vocabulary, reading comprehension, reading strategies, and motivation to read. The presentation will describe the study’s design, implementation, analysis, and results.
BACKGROUND
Improving adolescent literacy is a critical step to improving adolescent academic achievement, especially helping students meet the increased literacy demands in grade 4 and beyond.

Recent policy reports emphasize the need to build students’ reading vocabulary and comprehension skills to meet the increased literacy demands that begin in grade 4 (for example, Carnegie Council on Advancing Adolescent Literacy, 2010; Meltzer, Smith, & Clark, 2001). Experts who drafted the Common Core Standards for English Language Arts emphasize that students must show a steadily increasing ability to discern more from text to become successful readers (National Governors Association & Council of Chief State School Officers, 2010).

According to the reading 2009 National Assessment of Educational Progress (NAEP), 68% of all eighth-grade students read below the proficient level. Among low-income students, more (84%) score below this threshold. Comprehension difficulties are common among adolescent readers. Because of these problems, many students struggle to learn from advanced texts and may disengage from reading entirely (National Governors Association, 2005; RAND Reading Study Group, 2002). Certain teaching and learning approaches have been shown empirically to improve middle and high school students’ reading achievement, including direct teaching of vocabulary and teaching a combination of comprehension strategies (Kamil, 2003). The current evaluation examines the effectiveness of Thinking Reader®, which uses a software program to directly teach vocabulary and evidence-based reading comprehension strategies, supported by specific teacher practices (Tom Snyder Productions 2006a, b).

The direct evidence base for Thinking Reader is limited. The current study is the first randomized field trial to test its effectiveness. A previous quasi-experimental study evaluated a pilot version of the software for a sample of 102 struggling middle school readers (average age, 12.5). After controlling for pretest scores on reading vocabulary and comprehension, the study found that students in the Thinking Reader intervention demonstrated significantly greater gains in comprehension on the Gates-MacGinitie Reading Tests (GMRT; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 1999; Dalton et al., 2002). The effect size was 0.29, approximately half a grade level of reading achievement gain. That evaluation differed from the current work in several ways. First, a program of traditional (noncomputer-based) reciprocal teaching methods was delivered to the control group. Second, because the software was a pilot program, it had fewer features—less scaffolding and fewer assessments. Finally, the analysis focused only on struggling readers who had performed at the 25th percentile or lower on the pretest.

PURPOSE AND RESEARCH QUESTIONS
This evaluation of the impact of Thinking Reader use focused on two confirmatory research questions on the effect of the program on two measures of students’ reading achievement:

1. What is the effect of Thinking Reader on students’ reading vocabulary?
2. What is the effect of Thinking Reader on students’ reading comprehension?

---

1 This was a deliberate decision to make the intervention and control conditions as comparable as possible so that the main difference between the groups was whether or not Thinking Reader was used.
The study also examined the effect of *Thinking Reader* on two ancillary questions:

3. What is the effect of *Thinking Reader* on students’ use of reading comprehension strategies?

4. What is the effect of *Thinking Reader* on students’ motivation to read?

**SETTING AND PARTICIPANTS**

The target population for this study was grade 6 teachers and students in high-need schools in three states that had two or more reading/ELA teachers in grade 6 during the 2008/09 school year. Grade 6 was selected because the transition from elementary to middle school is the time when reading for content-area comprehension becomes increasingly important for academic success. High-need schools were defined as having more than 33% of students eligible for free or reduced-price lunch (based on the 2006/07 school data available at the time of recruitment in spring 2008). The cutpoint of greater than 33% was used to ensure a sufficiently large sampling frame of schools that could be considered economically disadvantaged. Additionally, eligible schools had to have the technology to implement the intervention in a whole-group setting, with a computer for each student.

**INTERVENTION**

*Thinking Reader* is a software program intended to improve the comprehension skills of middle-school students. The program incorporates elements commonly outlined as key components of effective adolescent literacy instruction in policy reports. Such elements include: instruction in comprehension strategies, attention to motivation and self-directed learning, ongoing formative assessment and inclusion of technology as an instructional tool. With *Thinking Reader*, students read novels on computers and respond to a variety of prompts intended to support students as they learn different strategies for understanding text. Teachers can customize the amount of support and guidance each student receives within the program. *Thinking Reader* provides tools that teachers can use to monitor student progress and provide feedback.

Intervention teachers were provided with three *Thinking Reader* digital novels to read with their students during the 2008–2009 academic year. These novels are commonly read in Grades 5-9. Intervention teachers were asked to participate in professional development (two 6-hour group sessions and three individual coaching sessions totaling between 7.5 and 8.5 hours) to learn how to use the *Thinking Reader* software. They were asked to incorporate *Thinking Reader* into their regular English language arts or reading instruction. Teachers were to use the software with their students between 110 and 165 minutes per week during the time in which a novel was being covered. Trainers estimated that each novel would take 4 to 6 weeks to complete. Trainers also suggested that one novel be covered in fall, one in winter, and one in spring.

**RESEARCH DESIGN**

The study is a multisite cluster randomized controlled trial that randomly assigned grade 6 teachers to intervention or control groups within schools. The study involved reading/English
language arts (ELA) teachers in 32 elementary and middle schools in 16 districts located in three states within the REL-NEI region. The study sample consists of 92 teachers and 2,407 students. Within each school, teachers and their intact classes were randomly assigned to intervention or control groups after students had been scheduled into classes using the typical school procedures. The random assignment produced two groups that did not have statistically significant differences on teacher characteristics, or on student preintervention achievement measures.

Control teachers used the typical curriculum the school used in the absence of the intervention (i.e., business as usual). These students engaged in the activities their usual reading/English language arts curriculum and instructional program entailed (e.g., reading short stories, reading newspaper and magazine articles, reading novels).

DATA COLLECTION AND ANALYSIS

Each school was visited twice during the study year to collect outcome measures that would address the primary and ancillary research questions. The baseline data collection (pretesting) took place September-October 2008 after random assignment had taken place and the follow-up data collection took place during May-June 2009.

At both pre- and posttest, the study team administered the reading vocabulary and comprehension subtests from the Gates-MacGinitie Reading Test (GMRT), a standardized, multiple-choice test (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 1999). The two ancillary outcomes — students’ use of comprehension strategies and student motivation to read — were measured using two existing self-report instruments, the Metacognitive Awareness of Reading Strategies Inventory (Mokhtari & Reichard, 2002) and the Motivation for Reading Questionnaire (Wigfield & Guthrie, 1997). Data from all measures given at the beginning and end of the year were used to estimate the impact of the intervention.

Other student data gathered include variables such as student gender, birthdate, ethnicity, whether each student was considered an English language learner and/or had an individualized education plan.

We distributed a teacher questionnaire that asked about years of teaching experience, educational attainment, and certifications/endorsements held. Additionally, the study team collected data on school characteristics from state education agency websites.

Implementation of Thinking Reader Program

Two data sources were used to measure if the Thinking Reader program was implemented as intended.

First, to determine and document the nature of differences in instruction provided to intervention and control classrooms and as one way to examine fidelity of Thinking Reader use, we conducted structured observations two times in a sample of study classrooms. Two teachers – one treatment and one control – were observed at each school, for a total of 64 teachers. The study team adapted the Center for the Improvement of Early Reading Achievement (CIERA)
Second, the Thinking Reader software collects information about students’ program use, including time logged in for each novel and the last passage that each student read. Time logged in was aggregated to the teacher level to gauge the level of dosage: total time using the program and per novel. Though the program does not record students’ time of day or date per log-in, the team used the dates of first and last log-in to calculate the average weeks per book to measure dosage intensity. Program data were also used to calculate the number of novels each student initiated and completed. These data provide information on the duration and adherence to (or divergence from) program developers’ “prescribed services” (Dane & Schneider, 1998). In this case, the recommendation for use was approximately 120 minutes per week for 4-6 weeks for each of the three novels.

**Analytical Strategy**

The analytic strategy for assessing the impacts of Thinking Reader was to compare the achievement outcomes for students whose teachers were randomly assigned to the intervention condition or to a business-as-usual control condition. The average outcome in the control group represents an estimate of the achievement level that would have been observed in the intervention group if it had not used Thinking Reader. The difference in outcomes between the intervention and control conditions provides an unbiased estimate of the impact of the intervention.

Given the research design and hierarchical data structure (i.e., students are nested within teachers who in turn are nested within schools), we used a three-level hierarchical linear model to estimate the impact of Thinking Reader on the outcomes of interest (Raudenbush & Bryk, 2002). To improve the precision of the impact estimates, students’ pretest scores, English language learner status, and special education status were included as covariates at Level 1; teacher education and years of teaching experience were included at Level 2; and school poverty and school size were included at Level 3. In addition, the study team explored whether the intervention effect for each outcome varied or was relatively stable and homogeneous across schools. Effect sizes were computed to show the magnitude of the effect of Thinking Reader, expressed in standard deviation units. Further, to account for the two comparisons being performed for the two primary outcomes (reading vocabulary and reading comprehension), and to avoid spurious positive findings, the team used a multiple comparison adjustment (a Bonferroni correction), dividing the critical p-value (alpha) in half.

We tested the robustness of the results by conducting several analyses to determine how sensitive the impact estimates were to the assumptions that were made. The team checked whether the results were sensitive to listwise deletion (by reanalyzing the data using two multiple imputation models); to the specification of the treatment coefficient as a random-effect (by reanalyzing the data using a fixed-effect treatment coefficient); to the specification of a three-level model (by reanalyzing the data using a two-level hierarchical model where schools were added as fixed...
effects); to the number of classes per teacher\(^2\) (by reanalyzing the data after randomly selecting a single classroom for the teachers who taught multiple sections); and to the keeping in the sample those students whose original teachers left their school at the beginning of the school year.

RESULTS AND CONCLUSIONS
The study final report is currently under peer review, so results cannot be included in this proposal. The complete final report is expected to be released by the end of fall 2010, at which point data can be shared publicly.

\(^2\) Because of the mix of different school configurations in our sample, we had 62 teachers who taught only one classroom, and 30 teachers who taught multiple classrooms. We ran our impact models using the entire sample of teachers.
References


