Running head: READ NATURALLY: RCT

Empirical Evaluation of Read Naturally Effects: A

Randomized Control Trial (RCT)

Theodore J. Christ

Jennifer Davie

University of Minnesota

Abstract

Early reading development requires that component skills are established, including

accurate and fluent decoding and word identification skills. There is substantial evidence

that modeling by a teacher (e.g., reading along with a teacher or a recording) and

repeated practice (e.g., repeated reading, flash card practice) can improve both reading

accuracy and fluency. This study evaluated the effectiveness of a commercially available

computer-based intervention that is designed to improve reading accuracy and fluency.

The Read Naturally Software Edition (SE) is an intervention that is delivered by a

computer. It requires minimal teacher time and minimal supervision. A randomized

control trial was conducted across six schools in four districts with 109 low performing

third grade students. Students were randomly assigned to groups within each classroom.

Those students assigned to the experimental condition were exposed to the intervention

for approximately 20 min per day for 10 weeks. Results provide evidence of statistically

significant differences and small-to-moderate effects on multiple standardized measures

of reading accuracy and fluency.

Empirical Evaluation of Read Naturally Effects: A Randomized Control Trial (RCT)

The availability of effective interventions to promote skill development among struggling students is a critical issue to educators. Statutory provisions within both No Child Left Behind (NCLB, 2001) and the Individuals with Disabilities Improvement Act (IDEIA, 2004) emphasize accountability for effective instructional practices along with early intervention and prevention services. The conceptual development of problem solving (Deno, 2002, 2005) and Response to Intervention (RtI; National Association of State Directors of Education, 2005; Speece, Case, & Molloy, 2003; Vaughn & Linan-Thompson, 2003) provide the framework for the more effective use of resources and a multi-tiered approach to service delivery that spans general and special education. There are generally three tiers of service, which include Tier 1 core services, Tier 2 supplemental services, and Tier 3 intensive services. Tier 1 services provide support to all students through core instruction and behavior supports. Tier 2 services provide additional support to a subgroup of students (15%) through supplemental supports, which are typically provided at the group level to optimize resources. Tier 3 services are provided to a small minority of students (5%). Services are individualized and intense to address the most resistant and severe academic and behavior problems within schools. High quality core supports and supplemental supports can function to prevent and remediate most problems. Tier 2 supplemental services are often described as standard protocol interventions because they are often evidence-based interventions that effectively remediate a common problem for most students. A research-base is necessary to inform Tier 2 and Tier 3 services, especially those services that are designed to remediate high incidence problems, such as deficit levels of reading achievement.

The National Reading Panel (NRP; 2000) concluded that there are five major components that contribute to reading, which are phonemic awareness, phonics, fluency, vocabulary, and comprehension. Reading fluency is the primary focus of this paper, as it is a critical skill that enables reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001) and is a strong predictor of later reading achievement (Good, Simmons, & Kame'enui, 2001; Hintze & Silberglitt, 2005; Stage & Jacobsen, 2001). Reading fluency is typically defined by the rate of accurate reading with appropriate expression (National Reading Panel, 2000). The importance of reading fluency is established across the predominant theories of early reading development (Adams, 1994; Chall, 1983; Ehri, 1995; LaBerge & Samuels, 1974; Stanovich, 1984). Theories converge to provide that individuals generally develop familiarity with letter-sound correspondences, orthographic features, and sight words as they progress to become automatic word readers. Automaticity is generally considered a critical component of early reading development, because readers allocate a greater proportion of their cognitive resources to the other critical aspects of reading once decoding and word recognition is automated. That is, if individuals expend less effort decoding and recognizing words, then they have the cognitive resources available to better attend to and monitor the meaning of text (Stanovich, 1984).

National estimates from the National Assessment of Educational Progress suggest that approximately 44% of fourth grade students in the United States function with low reading fluency within grade-level materials (Pinnell, Pikulski, Wixon, Campbell, Gough, & Beatty, 1995). Oral reading fluency can be measured using reading speed calculations to estimate the number of words read correctly per minute (WRCM), which can be compared to national norms (Hasbrouck & Tindal, 2006, 1992) and

benchmarks (Good et al., 2001; Hintze & Silberglitt, 2005; Stage & Jacobsen, 2001). In fact, oral reading fluency correlates highly with reading comprehension ($r = .91$; Fuchs, Fuchs, Hosp, & Jenkins, 2001), and evidence suggests that there may be a causal link between fluency training and increases in comprehension (Reutzel & Hollingsworth, 1993). The level of fluency development and its critical place in early reading achievement converge to establish reading fluency instruction and intervention as a critical issue to address.

Beginning readers develop accuracy and fluency skills when they are provided opportunities to interact with print. Critical features of effective fluency instruction include repeated exposure to words, opportunities for practice, and opportunities to receive feedback (National Reading Panel, 2000). Evidence suggests that increased opportunities to read (practice) contributes to the development of reading fluency (Stanovich, 1986). The NRP focused primarily on two instructional approaches related to fluency development. Those were guided oral reading and independent silent reading. The panel concluded that there was strong evidence for guided oral reading to promote reading fluency and reading achievement.

Kuhn and Stahl (2003) identified a variety of instructional approaches that are related to automaticity and reading fluency. Those included repeated reading, assisted reading (also known as reading with a model), isolated word practice, text segmenting, and choral responding. Their review generally supported the conclusion that research provides support for any of a variety of fluency-based approaches to reading instruction. The critical components were practice and support through repetition or modeling. Low resource and easy to carry out intervention strategies that employ those components, which include repeated readings (Rashotte & Torgesen, 1985) or listening-while-reading

(Daly & Martens, 1994), can also improve the frequency of implementation in and out of classrooms. That is, reading fluency interventions are found to be effective and often easy and efficient to implement. Progress in fluency development can be monitored over time to examine the efficacy of any interventions used. It is beneficial for interventions to include performance feedback along with praise and encouragement; and, if performance data are tracked and graphed, then both teachers and students can monitor and evaluate progress toward goals (Fuchs & Fuchs, 1986a, 1986b; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Fuchs, Fuchs, Hamlett, & Whinnery, 1991).

A repeated reading intervention requires students to read the same passage multiple times until they reach an acceptable level of fluency (Samuels, 1979). According to a metaanalysis in the NRP report (2000), guided repeated oral reading procedures are effective in improving reading accuracy ($ES = .55$), reading fluency ($ES = .44$), and reading comprehension ($ES = .35$). When stories are read repeatedly, reading speed increases on new stories that have high word overlap with the repeated story (Rashotte & Torgesen, 1985). In a variation of repeated reading, one study included 20 target words in each of five short stories that were read on subsequent days under different conditions (Reitsma, 1988). Of these conditions, guided reading with sustained feedback and independent reading with self-selected computer feedback produced substantially greater improvements in reading rate when compared to the control group. These results indicate that the repetitive reading of words in different contexts contributes to fluency gains that generalize across texts (Dowhower, 1987; Samuels, 1979).

Listening-while-reading (also known as reading with a model) is another effective intervention approach to promote reading fluency and generalized reading

acheivment. The procedure provides repeated opportunities for the student to observe

fluent reading and follow along (Skinner, Logan, Robinson, & Robinson, 1997). This

intervention can be used one-on-one or through taped/computerized models of reading

and can be combined with repeated reading. Results of research provide evidence for

improved rate and accuracy (Rasinski, 1990); although other research yields negligible

results for modeling alone without explicit practice (Reisma, 1988). The majority of

research that examined modeling and listening-while-reading required students to follow

along silently rather than reading aloud, making it difficult to determine if students were

indeed following along with the text. It is possible, or even likley, that the negligible

results coincide with instances when students fail to read along, which effectively

negates components related to repeated positive practice, exposure, and feedback that are

associated with the listening-while-reading intervention.

Read Naturally is an intervention system designed to implement three research-

based strategies: repeated reading, reading with a model, and progress monitoring with

feedback to promote student reading achievement. The primary skills that are targeted

are reading accuracy and fluency. In addition, vocabulary and reading for meaning are

emphasized through introductions to key words and comprehension questions. The

following description is provided on the Read Naturally website (2009):

> The Read Naturally [Software Edition; SE] provides engaging nonfiction
> material … appropriate for a student's age and skill level … Using teacher
> modeling and repeated reading and the appropriate curriculum, students can
> improve their fluency, and depending upon the [SE] Series, develop phonics
> skills, improve vocabulary, and promote comprehension.

> The Software Edition automates the placement process, assigns an appropriate
> reading goal, and allows you to customize the program by accelerating or
> skipping steps for advanced readers. It also creates detailed reports that track
> student progress and indicates areas that need additional coaching.

> SE guides students through the steps of the program with on-screen prompts, synchronizes professional audio recordings of the stories with on-screen text, provides vocabulary-building features, pictures, animations, and sound, and automatically calculates reading and quiz answer scores

The website also includes a description of reports, studies, case examples, and the evidence-base for the Read Naturally instruction program.

Prior research on the Read Naturally intervention system provides evidence that elementary and middle school students using the program generally approximate typical or ambitious rates of achievement of oral reading skills (Hasbrouck, Ihnot, & Rogers, 1999). To date, no randomized control trials of the Read Naturally program have been conducted. The What Works Clearinghouse (2007) reviewed 14 studies of Read Naturally and identified two studies in their evaluation of the intervention. One was an unpublished master's thesis that evaluated intervention effects on a small sample ($N =$ 15) of first graders who were randomly assigned to intervention conditions (Mesa, 2004) and the other is an unpublished dissertation on a modest sample ($N = 94$) of second graders who were matched across control and intervention conditions (Hancock, 2002). Neither of these studies provides sufficient evidence of intervention integrity and the latter study (Hancock, 2002) clearly did not implement all components of the intervention; moreover, the purpose of the latter study was not to evaluate the effects of the Read Naturally intervention system.

Purpose

The purpose of this study was to evaluate the effect of the Read Naturally SE intervention on skill development over a brief period of implementation (10 weeks). Skills in reading accuracy, reading fluency, and reading comprehension were used to index and evaluate intervention effects within a randomized control trial. Although the

intervention is designed for use in any tier of service, the primary focus on this study

was to evaluate Read Naturally SE as a Tier 2 intervention.

Method

Participants and Setting

A total of 109 third grade students (60 male, 49 female) from six schools in four

Midwestern suburban school districts participated in the study. The schools had not

previously used the Read Naturally program. Of the students, 10% were receiving

special education services, 23% were English Language Learners (ELL), and 60%

received free or reduced lunch. A total of 42% of the students were White, 28% were

African American, 23% were Hispanic, 6% were Asian American, and 1% were

American Indian. All participants performed at or below the 40th percentile on a measure

of oral reading fluency administered in the fall of third grade (DIBELS or AIMSweb), as

well as at or below the 40th percentile on a measure of reading comprehension

administered at the end of second grade (the Measures of Academic Progress test).

Students who qualified for the study, and whose parents gave permission for them to

participate, were randomly assigned within their classrooms to the control group (54

students) or the intervention group (55 students). Control group students continued to

experience the reading instruction typical of their classroom with no supplemental

fluency instruction. Estimates within group for both verbal ability (PPVT-4) and

processing ability (CTOPP) are presented in Table 1.

Measures

Pre-test measures of ability. Both the Peabody Picture Vocabulary Test (PPVT-

4) and subtests from the Comprehensive Test of Phonological Processing (CTOPP) were

administered at pre-test (December 11, 12 & 15 -19, 2008) as measures of verbal and

processing ability. Standard scores were derived for vocabulary knowledge on the

PPVT-4 and for both Rapid Digit Naming and Rapid Letter Naming on the CTOPP.

Results are presented in Table 1.

Pre-/post-measures of achievement. Prior to the intervention, all participants

were administered a battery of tests (December 11, 12 & 15 -19, 2008) to assess their

reading skills (fluency, accuracy, and comprehension). The battery of tests was repeated

after the intervention was completed (March 16-20, 2009). Reading fluency was

assessed with three curriculum-based measurement of reading (CBM-R) passages from

the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) third grade test package

(Good & Kaminski, 2002). The raw-score median of the three passages was used for

analysis. Reading fluency was also assessed with two additional measures, The Test of

Word Reading Efficiency (TOWRE; $M = 100$, $SD = 15$) and the Gray Oral Reading

Tests – Fourth Edition (GORT-4; $M = 10$, $SD = 3$). Standard scores were used for

analysis. Standard scores for reading accuracy were obtained from the GORT-4 ($M = 10$,

$SD = 3$) and the Word Identification subtest from the Woodcock Reading Mastery Test –

Revised (WRMT-R; $M = 100$, $SD = 15$). Standard scores for reading comprehension

were obtained from the GORT-4 ($M = 10$, $SD = 3$) and the Word and Passage

Comprehension subtests from the WRMT-R ($M = 100$, $SD = 15$). CBM-R was also

administered at a midpoint in the intervention (after approximately 5 weeks of

intervention).

Usability. The Usage Rating Profile – Intervention (URP-I; Chafouleas, Riley-

Tillman, Briesch, & Chanese, 2008) scale was used to evaluate acceptability and

usability. Ratings are recorded on a 6 point Likert scale with anchors at each point,

which are strongly disagree, disagree, slightly disagree, slightly agree, agree, strongly

agree. There are four factors with acceptable levels of internal reliability (α). Those

factors are Acceptability (.96), Understanding (.90), Feasibility (.85), and Systems

Support (.84).

   Intervention integrity. Read Naturally intervention integrity checklists were used

to assess and evaluate the components of intervention. The company publishes both a

teacher checklist (Appendix A) and a student checklist (Appendix B), which identify the

relevant components and procedures.

Conditions

   The intervention was implemented for approximately 10 weeks (January 5, 2009

to March 13, 2009) with assessments administered approximately 2 weeks prior to

implementation and one week after implementation. During the class time designated for

the Read Naturally intervention, students in the control condition engaged in nonreading

related activities. Students assigned to the intervention group were scheduled to work

with the Read Naturally software for 30 minutes a day, five days a week for 10 weeks.

The time of day when the intervention took place varied by classroom, but was always

scheduled so that no student missed typical reading instruction to receive the Read

Naturally intervention. The interventions took place either in a computer lab or in other

schoolrooms that were otherwise unoccupied. Instructional groups consisted of no more

than six students with one teacher supervising. Read Naturally utilizes fluency and

comprehension-building strategies that are designed to accelerate reading achievement.

Active components include curriculum matching, teacher modeling, repeated reading,

feedback, and progress monitoring. The intervention protocol was initially developed

and distributed in 1991; and it is used by many schools across the country.

Experimental (intervention) condition. Once students are individually placed at

their readability levels and their individual goals are set, the daily procedure for the Read

Naturally Software Edition (SE) is as follows: 1. *Select a Story*: Students click a picture

to select a story to read at their appropriate level; 2. *Read Along to the Key Words*:

Students click on each of the vocabulary words to hear them pronounced and to learn

what they mean. Students should read along as the words are pronounced and read the

student-friendly definitions along with the narrator; 3. *Write a Prediction*: Students use

the story illustration, key words, and title to write one or more sentences about what they

think the story will say about the topic; 4. *Cold Timing*: The students are timed reading

the unpracticed story aloud for one minute. Students are instructed to click difficult

words (words they stop in front of, skip, or stumble on) as they read. The SE subtracts

the number of errors, determined by the number of words the student clicks on, from the

total number of words read to calculate their cold timing or baseline score. A graph

displays their cold timing score in blue; 5. *Read Along to the Story*: The student reads

along quietly while the SE plays a recording of the story. A highlighted reading guide

automatically moves along to each sentence of the story as the student reads along with

the synchronized audio recording. Subvocalizing along with the narrator on the

recording ensures that students actually read along, and it helps them learn proper

pronunciation, expression, and phrasing. This component is designed to build word

recognition and word reading accuracy in preparation for subsequent reading practice.

The SE includes three recordings appropriate to the readability level of the selected

story, each one slightly faster than the previous one. Teachers can customize the read-

along step for each student by adjusting the number of read alongs and selecting the

reading speeds; 6. *Practice Reading the Story*: Students practice reading the story

independently without the recording multiple times until they can read at the

predetermined goal rate. A highlighted reading guide helps students focus on one line at

a time, and students may click on a difficult word to hear the word pronounced during

practice. The SE times each practice and indicates when students are ready to pass the

timed reading by allowing them to move on to the next step; 7. *Answer the Questions*:

Students answer five to nine multiple choice and open-ended questions about the story.

For most of the levels, the questions follow a specific pattern, allowing for the teacher to

detect when a student has trouble with particular types of questions. The SE corrects the

multiple-choice questions. The teacher corrects the open-ended questions during the Pass

step (step 9); 8. *Retell the Story*: In the sequenced levels, students retell or write what

they learned from the story. Teachers can specify how much time is allowed for the

retell step. If time is limited or students have inadequate keyboarding skills, teachers

may choose to have students retell the story orally or skip this step; 9. *Pass Timing*:

Students read the story aloud to the teacher in order to demonstrate that they can read the

story at the goal rate. To pass a story, students must reach their goal, make no more than

three errors, read with good expression, and answer the questions correctly. The teacher

and student view the results together. If the student passes, a graph shows their pass

timing score as a red bar above their cold timing score (the blue bar). If they do not pass,

the teacher can assign some remedial work (e.g., further practice) and then test them

again; 10. *Repeat*: Once students pass a story, they repeat these steps with a new story.

Selection and training of teachers

The four teachers were recommended by employees of the school districts

involved in the study to guide implementation across the six sites. Three out of the four

teachers had recently retired from the school districts. These teachers received a stipend

for their participation in this study. The fourth teacher was on school staff and used

available time in her schedule to implement the intervention. This school received

additional levels of the Read Naturally software as compensation for the teacher's time.

All four of the implementation teachers attended a six-hour training session in

December 2008. This is the typical initial training provided by Read Naturally

consultants when the program is implemented in elementary school settings. The

training included lecture and hands-on practice with the software. All the teachers were

given the Read Naturally SE Self-Study Course to take home at the end of the training.

This allowed for additional practice with the SE materials. Teachers were offered

compensation for up to two additional hours to review the self-study course on their

own.

Fidelity of instruction

Trained observers conducted observations of intervention fidelity every two

weeks to assess the integrity with which the intervention was implemented. Researchers

used two fidelity checklists that were developed by Read Naturally to evaluate teacher

implementation and student implementation. In addition to the information gathered by

the observers, the software documented the time spent on the program and the way

students progressed through the program.

<div align="center">Results</div>

Data were screened for missing values and accuracy of data entry. Data met the

necessary assumptions for planned analysis. Descriptive statistics for all assessments are

presented in Table 1.

The PPVT-4 was used to measure verbal ability prior to intervention. The mean

(*SD*) level of performance on the PPVT-4 for students in the control and experimental

conditions was 95 (10) and 94 (12) respectively, which was not a statistically significant

difference (α = .05), $F(1, 104)$ = .277, p = .60. Performances among students in the

control and experimental groups were generally equivalent on both the CTOPP Rapid

Digit Naming and Rapid Letter Naming, which approximated 9 (2) to 10 (2). Differences

were not statistically significant (p > .05). These results support the assumption that the

verbal and processing ability of student groups were substantially equivalent.

As described in the method section, three measures of fluency (CBM-R, GORT-

4 Fluency, and TOWRE Fluency), two measures of accuracy (GORT-4 Accuracy,

WRMT-R Word Identification), and two measures of comprehension (GORT-4

Comprehension, WRMT-R Comprehension Composite Score) were administered both

pre- and post-intervention. GORT-4 results were reported and analyzed on a standard

scale with a mean of 10 and standard deviation of 2. WRMT-R results were reported and

analyzed on a standard scale with a mean of 100 and standard deviation of 15. CBM-R

results were reported and analyzed as raw score units. Descriptive statistics are presented

in Table 1.

Multivariate Analysis of Covariance (MANCOVA), using pre-test performance

as a covariate, was used to test for statistically significant differences (Wilks Criterion, α

= .05) on post-test performance for the multivariate combinations of the measures for

reading fluency (3), reading accuracy (2), and reading comprehension (2). Planned

Analyses of Covariance (ANCOVAs) were used to test for statistically significant

differences across the control and experimental conditions for all univariate measures if

statistically significant differences of multivariate combination were observed. Hedges'

g was computed as the "*covariate adjusted mean difference* divided by *unadjusted*

pooled within-group SD" (emphasis original, p. 37) and was used to estimate univariate

effect sizes per the recommended procedures within the *What Works Clearinghouse*

Procedures and Standards Handbook V2 (WWC, 2008). Those estimates can be

interpreted such that effects that approximate .20 are small, .50 are modest and .80 or

greater are strong (Cooper & Hedges, 1994).

The MANCOVA was statistically significant for the multivariate combinations

of reading fluency measures, $F(3, 98) = 4.701$, $p < .05$ (Table 2), and reading accuracy

measures, $F(2, 101) = 8.01$, $p < .05$ (Table 3). Results were not statistically significant

for the multivariate combination of comprehension measures ($p > .05$) with observed

power of .95.

Results of ANCOVA for fluency measures yielded statistically significant

differences in post-test performance on CBM-R and GORT-4 Fluency, but not for

TOWRE Fluency (Table 2). CBM-R adjusted post-test means among students in the

control and experimental conditions were 70 and 76 respectively, which was a

statistically significant difference and a small effect size, $F(1, 100) = 5.20$, $p < .05$,

Hedges' $g = .20$. GORT-4 Fluency adjusted post-test means within the control and

experimental conditions were 7.5 and 8.5 respectively, which was a statistically

significant difference and approximately a median effect size, $F(1, 100) = 12.77$, $p < .05$,

Hedges' $g = .41$. TOWRE Fluency adjusted post-test means within the control and

experimental conditions were 93.5 and 94.9 respectively, which was not a statistically

significant difference, $F(1, 100) = 1.02$, $p = .31$, Hedges' $g = .13$ with observed power of

.17.

Results of ANCOVA for accuracy measures yielded statistically significant

differences in post-test performance on the GORT-4 Accuracy, but not the WRMT-R

Word Identification (Table 3). GORT-4 Accuracy adjusted post-test means among

students in the control and experimental conditions were 7.2 and 8.5 respectively, which

was a statistically significant difference and a medium effect size, $F(1, 100) = 5.20, p <$

.05, Hedges' $g = 48$. Group differences were not statistically significant as measured by

WRMT-R Word Identification, $F(1, 100) = .104, p = .75$, Hedges' $g = .04$, with

observed power of .062.

Alternate Analysis: Rate of Growth on CBM-R

CBM-R measures are often administered repeatedly to estimate the rate of

growth for individuals and groups of students. The CBM-R was administered at three

points in time to facilitate analysis of slope, which were calculated to estimate the rate of

weekly gain in units of WRCM (Figure 1). The mean (*SD*) level of performances within

the control and experimental groups were 1.10 (1.00) and 1.53 (1.10) WRCM gain per

week, which was a statistically significant difference and a small effect size, $F(1, 99) =$

4.58, $p < .05$, Hedges' $g = .18$.

Analysis of Implementation Fidelity

The software timed the duration in which students engaged with the Read

Naturally software. Students in the experimental conditions spent a mean (*SD*) duration

of 998 (114) min interacting with the intervention software, which is approximately 100

(11) min per week or five 20 min sessions per week. That observed duration for time

spent engaged was one third less than the 30 min per day recommended duration that is

published within the Read Naturally treatment integrity checklist (Appendix A).

Students progressed through an average of 23 (6) models in 10 weeks and spent an

average of 47 (13) min per module.

Analysis of Usage Rating Profile

Responses for each individual item of the URP were rated on a six-point scale of Strongly Disagree (1) to Strongly Agree (6). Descriptive statistics were calculated for ratings on individual items and aggregate scores on each of the four factors (Table 4). The ratings (rounded to the nearest whole number) indicate that teachers "agreed" that the intervention was acceptable for use as a intervention and instructional strategy (Acceptability; $M = 5.26$, $SD = .38$), "strongly agreed" that the procedures and rationale were clear and understood (Understanding/Knowledge; $M = 5.50$, $SD = .35$), "agreed" that the procedures could be implemented with integrity (Feasibility; $M = 4.75$, $SD = .41$), and "slightly agreed" that support from administration and support staff is necessary to implement the procedures (Systems Support; $M = 3.50$, $SD = .27$).

Discussion

This study was designed as a randomized control trial of a popular computer-based intervention, which is a likely candidate for use in any of the service tiers. The primary focus of this study was to evaluate the Read Naturally SE for use to remediate and prevent reading difficulties, which is a Tier 2 application. The intervention program was designed to accelerate the acquisition of reading fluency skills, which is defined as accurate rate-based reading. This study examined the impact of a 10-week implementation of Read Naturally. Intervention effects were examined with measures of accuracy, fluency, and comprehension. Intervention implementation was examined with measures of both intervention integrity and teacher evaluations of usability. Most aspects of the intervention were implemented with high degrees of integrity (Appendix A & B) and teachers found the intervention to have clear procedures that are feasible to implement within schools. Ratings indicate a high degree of intervention acceptability among teachers; however, ratings did indicate that teachers perceive the need for some

system level support, which might include additional staff and administrative support. This might be related to computer and software access along with staff to supervise students in the computer rooms. Overall, the results of this study provide evidence to support the use of Read Naturally if accuracy or fluency is defined as an instructional goal.

Statistically significant differences were observed between the control (no treatment) and experimental (Read Naturally SE) conditions for the multivariate combination of both accuracy and fluency measures. There was not a statistically significant effect for the multivariate combination of comprehension measures, which was an expected outcome. That is, the procedures and instructional targets focus primarily on accuracy and fluency skills. Although improved levels of accuracy and fluency can enable improved levels of comprehension, researchers would expect comprehension gains over more extended periods of implementation and not necessarily after only 10 weeks of implementation.

Students in the Read Naturally condition outperformed students within the control conditions on all measures of accuracy and fluency, although not all differences were statistically significant. Statistically significant differences along with small-to-moderate effect sizes were observed for most, but not all, of the standardized measures of reading accuracy and reading fluency. That is, student performance on post-tests was approximately one quarter to one half of a standard deviation improved at post-test as compared to students in no treatment control.

Additional analysis was conducted to examine the rate of CBM-R growth (i.e., fluency gain) within the control and experimental conditions. The mean rates were 1.10 and 1.53 WRCM per week gains respectively, which coincides with a 39% improvement

in the academic gain among students in the experimental condition. That is, students in

the Read Naturally condition gained, on average, .40 WRCM more than those students

in the control condition. The magnitude of improvement should be compared to typical

growth and expected levels of achievement among third grade students. Published

estimates of typical growth in third grade approximate 1.2 WRCM among students in

general education and .58 WRCM among students in special education (Deno et al.,

2001). Students are expected to exceed 100 WRCM by the end of the third grade year if

they are to be successful on large scale statewide assessments. A sustained improvement

of .40 WRCM across the academic year is substantial. Assuming a typical growth rate of

approximately 1.20 WRCM per week, a 39% improvement would result in improved

growth to 1.68 WRCM per week, which is a cumulative gain of 14.4 more words read

correct across 36 weeks. The implications of this magnitude of accelerated growth are

substantial. Although the effect sizes associated with CBM-R group differences could be

interpreted as moderate if compared to published standard criterion (Cohen, 1988;

Cooper & Hedges, 1994), those criterion are presented with caveats that clearly establish

effect sizes must be interpreted in context. The small-to-moderate effect sizes observed

for measures of both accuracy and fluency would indeed translate to meaningful and

substantial gains in academic achievement. An acceleration of .40 WRCM per week, or

14.4 WRCM within an academic year, is meaningful improvement.

    Figure 1 illustrates that the CBM-R average for the experimental group was

below that of the control group at pre-test. After 10 weeks of intervention, the students in

the Read Naturally experimental condition outperformed those in the control group.

These results were not specific to CBM-R measures. After accounting for pre-test levels

of performance, the students in the experimental condition scored one to two standard

score units better on most standardized measures of fluency and accuracy after 10 weeks

of intervention (see adjusted post-test scores, Table 1). If extrapolated across the

academic year then that those gains would translate to approximately three to seven

standard score units of improved performance for those students within the Read

Naturally experimental conditions as compared to the control condition. Such

improvements would be substantial. The pattern of growth also provides preliminary

evidence that the rate of improvement might be greater after extended periods of

implementation. Visual analysis of Figure 1 indicates that the rate of growth improved in

the second half of the intervention period. Future research is necessary to replicate and

examine that phenomenon. It might be that students demonstrate a larger magnitude of

growth after they become accustomed to the intervention procedures, which were

unfamiliar early on in the study.

The observed gains for the experimental group were obtained with minimal

resources. In schools, teacher and personnel time is one of the most valuable resources.

The Read Naturally intervention was implemented with high levels of integrity and

positive ratings for acceptability/usability. Those results were observed within

conditions where there was minimal training and teacher support for implementation.

Those results provide evidence that the intervention can be scaled (i.e., used in practice).

Results support the conclusion that the intervention is both effective (to accelerate

accuracy and fluency) and feasible for use in practice.

Future Directions and Limitations

The results of this study generally support the conclusion that the Read Naturally

SE effects are more robust than indicated by the prior studies that were identified by

What Works Clearinghouse (2007) as having met design standards (Mesa, 2004) or met

design standards with reservations (Hancock, 2002). This study was designed and

implemented with the high standards for design to establish high internal and external

validity. Student participants were selected across four districts and six schools.

Participants were assigned at random within classrooms to either the control or

experimental condition. The student participants in this study were relatively low

achieving and the impact on typical achieving and high achieving students should be

examined in the future, especially in the lower grades with developing readers. It is

important to note that the Read Naturally intervention integrity checklist indicates that

students should be actively engaged with the software for 30 min on each occasion.

Although integrity was generally high, the estimated average duration for

implementation approximated 20 min. It is possible that the dosage was a full one third

below the recommended level, which might indicate that improved rates of growth could

be expected if integrity were improved. The software includes an integrity component so

schools and researchers can monitor the duration of time on task. Although data

collectors perceived 30 min implementation, the actual recordings by the software

indicate deficit durations. It is likely that transition times were included in the estimates

by data collectors, but more direct measures with software recorded actual time on task.

Although the primary purpose of this study was to examine a Tier 2 application,

it is likely that the intervention would be an effective component of Tier 1 and Tier 3.

Future research is necessary to examine the effects of typically developing readers,

especially those students in first and second grade who are establishing accurate and

fluent decoding and word identification skills. Future research is also necessary to

examine the effects for students receiving intensive supports in Tier 3.

The essential features of effective service delivery through problem solving and RtI are both data-based decision making and access to evidence-based practices. Research is necessary to both identify the effective components of intervention and contribute to the development and evaluation of multi-component interventions; especially those that are promoted as effective and have scalability characteristics. This will require high quality research carried out by independent researchers to evaluate the claims of those purveyors that distribute intervention systems.

Author Note

2. Additional information is available upon request, which includes integrity data,

    time spent on intervention, and teacher acceptability ratings. Alternate

    measurement outcomes, including raw scores and percentile ranks, are also

    available.

References

Adams, M. J. (1994). *Beginning to read: Thinking and learning about print*. Cambridge,

MA: MIT Press.

Chafouleas, S. M., Briesch, A. M., Riley-Tillman, T., & McCoach, D. (2009). Moving

beyond assessment of treatment acceptability: An examination of the factor

structure of the Usage Rating Profile--Intervention (URP-I). *School Psychology

Quarterly, 24*(1), 36-47.

Chafouleas, S. M., Riley-Tillman, T., Briesch, A. M., & Chanese, J. A. (2008).

Generating usable knowledge: Initial development and validation of the Usage

Rating Profile for Interventions (URP-I). *Canadian Journal of School

Psychology, 23*(2), 175-189.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).

Hillsdale, NJ: Lawrence Earlbaum.

Cooper, H., & Hedges, L. V. (1994). Research synthesis as a scientific enterprise. In H.

Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 3-14).

New York: Russell Sage Foundation.

Daly, E. J. III & Martens, B. K. (1994). A comparison of three interventions for

increasing oral reading performance: Application of the instructional hierarchy.

Journal of Applied Behavior Analysis, 25, 823-840.

Deno, S. L. (2002). Problem solving as best practices. In A. Thomas & J. Grimes (Eds.),

Best practices in school psychology IV. (pp. 37-56). Bethesda, MD: National

Association of School Psychologists.

Deno, S. L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment*

   for intervention: A problem-solving approach (pp. 10-42). New York: Guilford

   Press.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of*

   Research in Reading, 18, 116-125.

Fuchs, L. S., & Fuchs, D. (1986a). Curriculum-based assessment of progress toward

   long-term and short-term goals. *Journal of Special Education* (20), 69-82.

Fuchs, L. S., & Fuchs, D. (1986b). Effects of systematic formative evaluation: A meta-

   analysis. *Exceptional Children* (53), 3-14.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative

   evaluation of academic progress: How much growth can we expect. *School*

   Psychology Review, 22, 27-48.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Whinnery, K. (1991). Effects of goal line

   feedback on level, slope, and stability of performance within curriculum-based

   measurement. *Learning Disabilities Research and Practice, 6*, 66-74.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an

   indicator of reading competence: A theoretical, empirical, and historical analysis.

   Scientific Studies of Reading, 5(3), 239-256.

Gest, S. D. & Gest, J. M. (2005). Reading tutoring for students at academic and

   behavioral risk: Effects on time-on-task in the classroom. *Education and*

   Treatment of Children, 28, 25-47.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills*

   [6th ed.]. Retrieved November 2002, from http://dibels.uoregon.edu

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(3), 257-288.

Hancock, C.M. (2002). Accelerating reading trajectories: The effects of dynamic research-based instruction. *Dissertation Abstracts International, 63*(06), 2139A. (UMI No. 3055690).

Hasbrouck, J. E., Ihnot, C., and Rogers, G. (1999). Read Naturally: A strategy to increase oral reading fluency. *Reading Research Instruction, 39*(1), 27–37.

Hasbrouck, J., & Tindal, G.A. (2006, April). Oral Reading Fluency Norms: A Valuable Assessment Tool for Reading Teachers

Hasbrouck, J. E., and Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children, 24*(3), 41–44.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*(3), 372-386.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3-21.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, Vol. 6*(2), 293-323.

Mesa, C. L. (2004). *Effect of Read Naturally software on reading fluency and comprehension.* Unpublished master's thesis, Piedmont College, Demorest, GA.

National Reading Panel. (2000). *Teaching children to read: An evidence-based*

assessment of the scientific research literature on reading and its implications

for reading instruction. Washington, DC.

National Association of State Directors of Education (NASDE). (2005). *Response to*

intervention: Policy considerations and implementation. Alexandria, VA: Author

Pinnell, G. S., Pikulski, J. J., Wixon, K. K., Campbell, J. R., Gough, P. P., and Beatty, A.

S. (1995). Listening to children read aloud data from NAEP's Integral Reading

Performance Record CIRPR at grade 4. Report No. 23-FR-04 prepared by the

Educational Testing Service, Washington D.C.: Office of Educational Research

and Improvement, U.S. Department of Education.

Read Naturally. (2009). Website Description. Retrieved June 17, 2009, from

www.ReadNaturally.com

Rashotte, C.A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in

learning disabled children. *Reading Research Quarterly, 20*,   180-188.

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-

while-listening and independent reading with computer-based speech feedback.

Reading Research Quarterly, 23(2), 219–235.

Reutzel, D. R., and Hollingsworth, P. M. (1993). "Effects of fluency training on second

graders' reading comprehension." *Journal of Educational Research* 86 (6), 325–

331.

Samuels, S. J. (1979). The method of repeated readings. *The Reading Teacher, 32,* 403–

408.

Skinner, C. H., Logan, P., Robinson, S. L., and Robinson, D. H. (1997). Demonstration

> as a reading intervention for exceptional learners. *School Psychology Review,*
> *26*(3), 437–447.

Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education

> instruction as the first gate to learning disabilities identification. *Learning*
> *Disabilities Research & Practice, 18*(3), 147-156.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated

> performance-based assessment using oral reading fluency. *School Psychology*
> *Review, 30*(3), 407-419.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual

> differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4),
> 360–407.

Stanovich, K. E. (1984). The interactive-compensatory model of reading: A confluence

> of developmental, experimental, and educational psychology. *RASE: Remedial &*
> *Special Education, 5*(3), 11-19.

Vaughn, S., & Linan-Thompson, S. (2003). What is special about special education for

> students with learning disabilities? *Journal of Special Education, 37*(3), 140-147.

Table 1. Descriptive Statistics for Measures of Ability, Fluency, Accuracy, and Comprehension

Measure	Condition	N	Pre-Test M (SD)	Post-Test M (SD)	Adjusted Post-Test M (SD)
Verbal Ability					
PPVT-4	Cont	53	95 (10)		
	Exp	53	94 (12)		
	Total	106	95 (11)		
Automatic Naming Ability					
CTOPP- Rapid Letter Naming	Cont	52	10 (2)		
	Exp	53	9 (2)		
	Total	105	9 (2)		
CTOPP – Rapid Number Naming	Cont	53	10 (2)		
	Exp	53	10 (2)		
	Total	106	10 (2)		
Reading Fluency					
CBM-R – Words Correctly per Min	Cont	53	58 (24)	72 (25)	70 (12)
	Exp	53	55 (26)	74 (29)	76 (12)
	Total	106	56 (25)	73 (27)	
TOWRE – Word Reading Efficiency	Cont	53	91 (12)	94 (11)	93 (7)
	Exp	53	90 (11)	94 (10)	95 (7)
	Total	106	91 (11)	94 (11)	
GORT-4 – Fluency	Cont	53	8 (3)	8 (3)	7 (2)
	Exp	52	7 (3)	8 (3)	9 (2)
	Total	105	7 (3)	8 (3)	
Reading Accuracy					
GORT-4 – Accuracy	Cont	53	8 (3)	7 (3)	7 (2)
	Exp	52	8 (3)	9 (3)	9 (2)
	Total	105	8 (3)	8 (3)	
WRMT-R – Word Identification	Cont	52	98 (8)	98 (8)	98 (4)
	Exp	53	97 (7)	97 (7)	99 (4)
	Total	105	97 (8)	97 (8)	
Reading Comprehension					
GORT-4 – Comprehension	Cont	52	9 (2)	9 (2)	10 (2)
	Exp	53	9 (3)	9 (3)	10 (2)
	Total	105	9 (2)	9 (2)	
WRMT-R – Passage Comprehension	Cont	52	95 (7)	95 (7)	97 (4)
	Exp	53	95 (7)	95 (7)	96 (4)
	Total	105	95 (7)	95 (7)	

Note. Students were removed for analysis if data were missing (removed by test)

Table 2.
Tests of Statistical Significance for Fluency Measures

	MANCOVA	ANCOVA					
		CBM-R		TOWRE Fluency		GORT-4 Fluency	
	$F(3,98)$	$F(1,100)$	g	$F(1,100)$	g	$F(1,100)$	g
Conditions	4.701*	5.20*	.20	1.02	.13	12.77**	.41

Note. Conditions are control and experimental
[a] Wilks Lamba criterion
[b] Pre-test performance was used as the covariate
[c] Hedges' g is the covariate adjusted mean difference divided by unadjusted pooled within-group *SD*
* $p < .05$, ** $p < .01$

Table 3.
Tests of Statistical Significance for Accuracy Measures

	MANCOVA [b]	ANCOVA [b]			
		GORT-4		WRMT-R Word Identification	
		Accuracy			
	$F(2,101)$ [a]	$F(1,102)$	g [c]	$F(1,102)$	g [c]
Conditions	8.014**	14.03**	.48	.163	.04

Note. Conditions are control and experimental
[a] Wilks Lamba criterion
[b] Pre-test performance was used as the covariate
[c] Hedges' *g* is the covariate adjusted mean difference divided by unadjusted pooled within-group *SD*
* $p < .05$, ** $p < .01$

Table 4.
Usage Rating Profile – Intervention (URP-I) with Descriptives for Items and Factors

	Rating Items	M	SD
1	The amount of time required to use this intervention is reasonable.	5.8	.5
2	I would implement this intervention with a good deal of enthusiasm.	6.0	.0
3	The intervention could be implemented for the duration of time as prescribed.	5.5	1.0
4	The amount of time required for record keeping with this intervention is reasonable.	5.8	.5
5	I am motivated to try this intervention.	6.0	.0
6	I would need consultative support to implement this intervention.	4.8	1.9
7	All pieces of this intervention could be implemented precisely.	5.0	.8
8	The intervention could be implemented with the intensity as prescribed.	4.8	1.9
9	I would have positive attitudes about implementing this intervention.	6.0	.0
10	I understand the procedures of this intervention.	6.0	.0
11	I would know what to do if I was asked to implement this intervention.	6.0	.0
12	Overall, the intervention is beneficial for the child.	5.8	.5
13	Implementation of this intervention would require support from my co-workers.	4.3	1.5
14	Parental collaboration is required in order to use this intervention.	3.3	1.5
15	The requirements for implementing this intervention are unclear.	1.3	.5
16	I would not be interested in implementing this intervention.	1.3	.5
17	The intervention could be implemented exactly as described.	3.3	2.1
18	This intervention is a good way to handle the child's behavior problem.	5.0	.8
19	I could only implement this intervention with assistance from other adults.	3.3	1.0
20	The intervention is a fair way to handle the child's behavior problem.	5.5	.6
21	This intervention is reasonable for the problem behavior described.	5.5	.6
22	I could implement this intervention by myself.	5.3	.5
23	I would need support from my administrator to implement this intervention.	4.8	.5
24	I would be resistant to use this intervention.	1.5	.6
25	This intervention could be implemented as frequently as described.	5.0	1.0
26	This is an acceptable intervention strategy for the child's problem behavior.	5.5	.6
27	I am knowledgeable about the intervention procedures.	5.3	1.0
28	This intervention is an effective choice for addressing a variety of problems.	4.8	1.9
29	This intervention would not be disruptive to other students.	3.0	2.2
30	I have the skills needed to implement this intervention.	5.8	.5
31	Use of this intervention would save time spent on classroom management.	4.0	1.4
32	I understand how to use this intervention	6.0	.0
33	I liked the procedures used in this intervention.	5.5	.6
34	I would have no idea how to implement this intervention.	1.0	.0
35	The directions for using this intervention are clear to me.	5.8	.5

Factors	M	SD
Acceptability (Items: 2, 5, 9, 12, 16*, 18, 20, 21, 24*, 26, 28, 31, 33)	5.29	.38
Understanding (Items: 10, 11, 15*, 27, 30, 32, 34*, 35)	5.50	.35
Feasibility (Items: 1, 3, 4, 7, 8, 17, 25, 29)	4.75	.41
Systems Support (Items: 6, 13, 14, 19, 22*, 23)	3.50	.27

Table 5.
Teacher Responsibilities Fidelity Checklist (Read Naturally)

Items	Percent Collected[a]	Percent Endorsed[b]
Planning & Setting Up		Y/N
A. Session length is 30 minutes.	88	67
B. Students attend 5 sessions per week.	92	100
C. Comfortable workstations with working computers/software, headphones, etc.	91	100
D. Setting promotes students' engagement for entire session.	95	100
E. Ratio of adults to students is no greater than 1:6.	97	99
F. System for indicating when students need teacher.	87	99
Implementing the Steps		
1. Select a Story—Teacher reinforces making a story selection after pass step on previous story—before quitting.	90	100
2. Key Words—Teacher reminds students to click on each key word and listen to the definition and sample sentence.	81	92
3. Prediction—Teacher encourages students to quickly write a brief prediction based on story title, key words, and illustration. If keyboarding skills are weak, teacher may listen to an oral prediction or "take dictation" and type prediction for student—to allow more time on task for reading.	85	100
4. Cold Timing—Teacher encourages students to click on difficult words and do an "honest" cold timing. Teacher may choose to monitor some students' cold timings or set options to require teacher for cold timing.	86	100
5. Read Along—Teacher monitors Read Along to ensure students are reading quietly aloud and following the highlighted sentences with the audio.	87	97
6. Practice—Teacher checks to see that students know their fluency goals and monitors practice to see that students read quietly aloud throughout each practice, click on difficult words, and click on last word read when bell sounds. Teacher confirms that student usually needs 3-10 practices to pass a story.	76	97
7. Quiz—Teacher monitors to be sure students read and thoughtfully answer all questions. Teacher checks to be sure students are not just clicking multiple choice answers until the right one is chosen.	81	100
8. Retell Teacher sets a reasonable time limit and encourages students to write a main idea and several supporting sentences— not to spend too much time on retell, because the emphasis is to have the student spend most of the time reading.	79	100
9. Pass— Teacher ▪ reminds students to continue practicing until s/he is available to come for hot timing. ▪ counts errors, rates expression, and directs students to click on final word. ▪ goes to Pass Requirements Results page and directs student to	82	100

review work, additional practice, and/or retest as indicated.

Monitoring Progress		
A. Teacher monitors student progress regularly using SE reports* at these times:		
▪ when a student seems to be struggling with accuracy or goal rate		
▪ after first six stories in Read Naturally	42	100
after first 12 stories in a level		
▪ at the end of a level		
(*i.e. Needs At-A-Glance for Students Report)		
B. Decisions are made based on data:		
▪ Continue at same level & goal		
▪ Adjust the reading rate goal		
▪ Move to more difficult (or easier) material		
▪ Other decisions:	40	100
▪ Rate of read-alongs		
▪ Number of read-alongs		
▪ Teacher required for cold timing		

Teacher Management		
Teacher uses Teacher Management to:		
▪ Set up class/student information		
▪ Set up story options (goal, level, SE options)		
▪ Reset stories		
▪ Place students (level and goal)	55	100
▪ Print awards and parent letters		
▪ Print stories in a level		
▪ Generate reports		

Communicating with Teachers, Students & Parents		
A. Teacher interactions with students are positive and encouraging.	73	100
B. Teacher monitors students and provides assistance as needed.	73	100
C. Teacher confers with student before making a change in the program.	55	100
D. Teacher communicates with regular classroom teacher and parents by using appropriate reports and calling to discuss progress as necessary.	4	100

Note. Data were collected approximately every two weeks within each of six sites for a total of 72 possible observations.

[a] percent collected is the proportion of possible data collection instances for which data were available

[b] percent endorsed is the proportion of instances that a positive response (yes) was recorded

Table 6.
Student Responsibilities Fidelity Checklist (Read Naturally)

Items	Percent Collected[a]	Percent Endorsed[b]
General Student Checklist		**Y/N**
A. Students are "engaged" for 30 min. session and work independently through the steps.	92	68
B. Students have basic computer skills (password, use of mouse, volume control, etc.).	100	100
C. Students know their individual fluency goals.	82	100
D. Students can pass with 3 to 10 repeated reading practices.	26	60
E. Students can pass a story in one to two 30 min. sessions.	28	100
F. Students understand and use signal for indicating when teacher support is needed.	91	99
Steps Student Checklist		
1. Select a Story— Students understand how to select story and confirm choice.	94	100
2. Key Words— Students are engaged and follow along closely on key words step.	91	94
3. Prediction— Students type in at least a brief phrase or sentence that is a logical prediction.	90	99
4. Cold Timing— Students start timer and read quietly aloud until ending signal, may underline difficult or unknown words, and accurately record score.	88	97
5. Read Along— Students read quietly aloud with narrator and carefully follow along as sentences are highlighted.	94	92
6. Practice— Students start timer and read quietly aloud until ending signal, click on any unknown words for support, and accurately record score.	91	99
7. Quiz— Students attempt to carefully read and thoughtfully answer each question.	79	97
8. Retell Students write a main idea sentence and several supporting sentences with some detail information about the story. *(5 min.)*	78	95
9. Pass— Students read the story at individual goal rate with three errors or less and meet criteria for expression of at least a two according to the description.	83	100

Note. Data were collected approximately every two weeks within each of six sites for a total of 72 possible observations.

[a] percent collected is the proportion of possible data collection instances for which data were available

[b] percent endorsed is the proportion of instances that a positive response (yes) was recorded
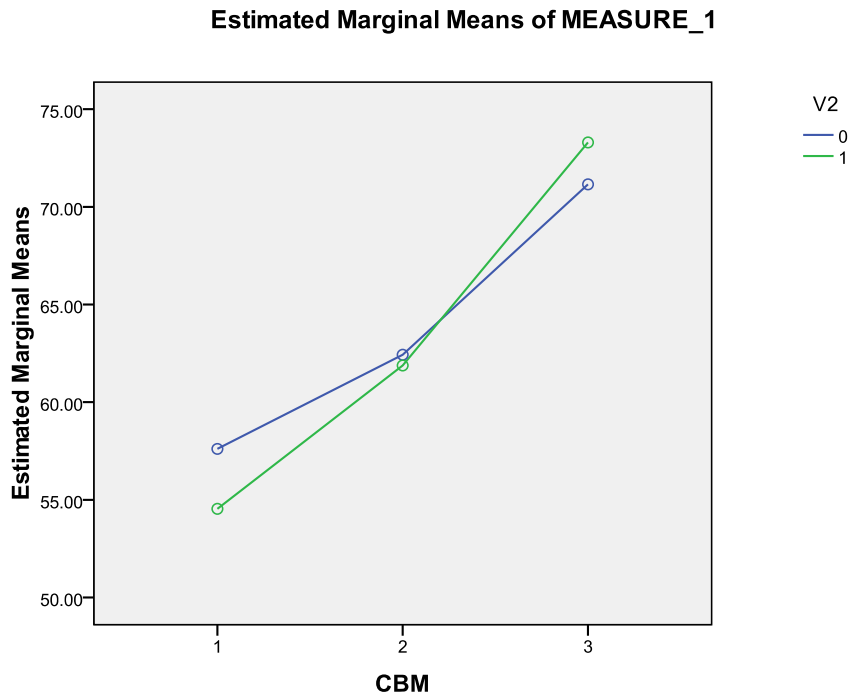
Estimated Marginal Means of MEASURE_1

Figure 1. Depicts Pre-Test (1), Mid-Test (2) and Post-Test (3) values on the X-axis for the median of three CBM-R administrations across the Control (0) and Experimental (1) Groups