

Intensive Intervention in Reading Fluency for At-Risk Beginning Spanish Readers

Maria Guadalupe De la Colina
Bryan Independent School District

Richard I. Parker, Jan E. Hasbrouck, and Raphael Lara-Alecio
Texas A&M University-College Station

Abstract

A study was conducted using a single case, multiple baseline (across subjects) design to study an intensive reading intervention among low-achieving at-risk students in first and second grade Spanish/English bilingual classrooms. The intervention, involving three research-supported techniques, was conducted for 45 minutes per day, three days per week, over 12 weeks, with 74 students from four classrooms (analyses included 53 students with complete data). Dependent measures were oral reading fluency scores and comprehension scores from post-reading questions, collected every two weeks from equivalent probes. It was concluded that implementing intensive reading fluency interventions in bilingual classrooms is feasible and valuable if conducted with fidelity, and if students are highly engaged.

Introduction

The problem of low academic achievement and excessive high school dropout rates among bilingual Hispanic students (Kaufman & Frase, 1990; Snow, Burns, & Griffin, 1998; U.S. Department of Education, 1997) has caused bilingual education to be closely scrutinized. The performance of bilingual Hispanic students has been low on statewide academic assessments. For example, the 1997 Texas State Performance Report for the Texas Assessment of Academic Skills (TAAS) in reading showed a passing rate of 92.4% for White students in grades 3–8 and 10 but only 75% for Hispanic students in the same grades. These academic problems may be due in part to three disadvantages that many bilingual Hispanic students endure in school: (a) low levels of English language proficiency, (b) difficulty with advanced Spanish language skills, and (c) students' parents and extended families having little formal education (Valdivieso & David, 1988; U.S. Bureau of the Census, 1992).

A considerable quantity of literature asserts that literacy instruction in a student's first language (L1) facilitates acquisition of literacy in a second language (L2) (Cummins, 1989; 2000; García, 1994; Genesee, 1987; Jimenez, García, & Pearson, 1995; Krashen, 1992; Lanauze & Snow, 1989; Leasher-Madrid & Garcia, 1985; Medina, 1988; Troike, 1978). It is upon this premise that some of the guidelines for transitional bilingual education are based (González, 1994; Prado & Tinajero, 2000): starting academic instruction in L1, including reading, and transitioning students into academic instruction in L2 as their acquisition of L2 strengthens. A secondary goal is to maintain and continue to develop literacy in the L1 heritage language. Indeed, a major focus of the federal Bilingual Education Act is to assist LEP students to acquire a second language and eventually be mainstreamed into a monolingual program (*Bilingual Education Act*, 1968).

However, continuing public challenges to the efficacy of transitional bilingual programs (August & Hakuta, 1997) have created pressure on these programs to more efficiently increase Spanish reading competency of young bilingual students. More rapid development of Spanish reading competency should lead to more rapid competency in English reading. Bilingual teachers and program administrators are therefore seeking intensive, research-supported Spanish language reading interventions, especially those suitable for low-achieving students in grades 1–4.

Reading Fluency Interventions

Little methodologically rigorous research exists on reading instruction in a second language (August & Hakuta, 1997) or on the effectiveness of Spanish-language reading programs in bilingual classrooms (Snow, et al., 1998). In the absence of research to inform decisions regarding the selection of appropriate reading interventions for bilingual programs, the use of English reading interventions, well supported by a convergence of research evidence, should be examined as possible options for use with struggling readers in these programs. One focus of reading interventions for English-speaking readers has been on improving students' reading fluency because of its key role in the development of skillful reading.

Automaticity theory posits that a fluent reader decodes text automatically—that is, without conscious direction—leaving attention free to be used for comprehension (Laufer, 1973; Laufer & Samuels, 1974). Beginning readers focus their efforts and attention on figuring out each word (Breznitz, 1987). Therefore, the beginning reader's efforts are devoted mostly to decoding, rather than on deriving meaning from the reading. This makes the process of comprehension more difficult and slower (Laufer & Samuels, 1974; Nathan & Stanovich, 1991; Stanovich, 1986). Three instructional practices have been documented as effective in improving the fluency skills

and, thus, the reading comprehension skills of children who speak and read in English: (a) repeated reading, (b) teacher/expert modeling, and (c) student self-monitoring of progress.

Repeated reading

Repeated reading entails rereading a passage to improve automaticity (Koskinen & Blum, 1984). Several studies conducted with students in the elementary grades have demonstrated that repeated reading improves reading accuracy, speed, and comprehension (Dahl, 1974; Dowhower, 1987; Fleisher, Jenkins, & Pany, 1979; Herman, 1985; Homan, Klesius, & Hite, 1993; McCormick & Samuels, 1976; O'Shea, Sindelar, & O'Shea, 1985; Rashotte & Torgesen, 1985; Turpie & Paratore, 1995).

Teacher modeling

The importance of teachers reading aloud to students has been extensively documented (Huck, Hepler, & Hickman, 1987). Carbo (1978) developed a method in which students listen to a tape-recorded story by a fluent reader. This format is attractive to students because they can work on their fluency independently (Rasinski, 1989). Research has demonstrated that providing an expert model of oral reading along with, or slightly ahead of, students produces growth in vocabulary, word knowledge, and visual decoding (Eldredge & Quinn, 1988; Heckelman, 1969; Reitsma, 1988; Skinner, Logan, Robinson, & Robinson, 1997).

Progress monitoring

Providing evidence to students about their progress toward goals can positively affect their performance (Bandura, 1982; Farnham-Diggory, 1972; Fuchs & Fuchs 1986; Hasbrouck, Woldbeck, Ihnot, & Parker, 1999; Shinn, 1989; Skager, 1984; Soltys, 1997; Wang & Peverly, 1986). This evidence of improvement can come from students' own self-monitoring efforts as well as teacher or peer tutor reports (Bentz & Fuchs, 1992). Low-achievers may be as capable of self-monitoring as high-achievers, although they may require support from external progress monitoring systems (Smith, 1991).

A combined strategy: Read Naturally

Read Naturally (RN) (Ihnot, 1997) is a strategy that combines the use of repeated reading, teacher modeling, and student self-monitoring to enhance students' reading fluency skills. RN has been studied in general education, remedial, and special education settings. The major study supporting RN was conducted over a six-year period with 214 second- and third-grade students. The oral reading fluency (ORF) and comprehension scores for these students increased, on average, over 10 percentile points across six months of instruction. Other students using RN have shown gains in reading fluency and comprehension that exceeded typical improvement (Hasbrouck, Ihnot, & Rogers, 1999).

Intensity and Level of Engagement

Important to the success of any reading intervention is its intensity, i.e., how many minutes per day and week active instruction occurs. Intensity of instruction also includes the pace with which new material is presented and the number of interactions between student and teacher. An intensive intervention will keep students actively engaged, responding at a high rate and moving quickly through the lesson and the curriculum (Carnine, Silbert, & Kameenui, 1997). Active practice time is a good index for the opportunity to learn, and it commonly relates to degree of skill improvement (Berliner, 1979; Stallings, 1980; Wyne & Stuck, 1982). While research strongly supports the importance of intensive interventions, especially with children who are struggling (Snow, et al., 1998; Torgeson, 2000), obstacles to providing intensive programs are abundant, including insufficient scheduled time, inattentive or unmotivated students, classroom interruptions, and excessive teacher explanations, which limit student active responding.

Objective measures of instructional intensity typically include frequency and duration (minutes per day and days per week of instructional opportunities). Level of student engagement may be measured by direct observation of student attention and responding, or may be inferred from records and work samples. In the present study, we objectively defined level of engagement as the average number of timed oral reading practices each student performed per week.

Purpose

The purpose of this study was to assess the effectiveness of a fluency intervention, Read Naturally, which combines research-supported components of repeated readings, demonstration, and self-monitoring, using Spanish reading materials with low-achieving, at-risk students in bilingual classrooms in grades 1 and 2. The researchers sought to answer the following question: Over a 12 week period, will a Spanish translation of Read Naturally, implemented 45 minutes per day and three days per week, improve the reading skills of bilingual grade 1 and 2 students? Because research supports the importance of intensity of reading interventions, and because Read Naturally permits intensity of engagement to be objectively monitored, we also asked a supplemental question: Did students' reading improvement vary as a function of level of student engagement in the materials?

Method

Context

This study was conducted in four Spanish/English bilingual classrooms in a small city in southeast Texas. The district is composed of 13,500 students and 12 schools, with a multi-ethnic makeup (African American, 24%; Hispanic, 31%; White, 44%). Sixty-three percent of the students are economically disadvantaged. The district's bilingual/ESL program enrolls nearly 1,000 students. The school district has been making steady progress in Hispanic school performance in recent years. From 1994-98, the Hispanic dropout rate in grades 7-12 dropped from 4% per year to less than 1% per year. Hispanic passing rates on the mandated Texas Assessment of Academic Skills (TAAS) Reading test improved from 69% in 1994 to over 76% in 1998, exceeding by 10 percentage points the improvement of district White students over this same five-year period.

This study's intervention took place in transitional bilingual classrooms during the reading/language arts period, in which students were involved in concentrated phonics practice and ESL-based activity centers. Transitional bilingual education in Texas is offered in pre K-5 grades (Texas Education Agency, 1997). If students have not exited the program by the fifth grade, they enter an English as a Second Language (ESL) program, which starts at sixth grade and continues through 12th grade. Bilingual students are categorized in three levels of English proficiency: beginners, intermediate, and advanced. At the first level, the students receive instruction in their native language, Spanish, and also receive 45 minutes of ESL instruction. At the next two levels, the minutes of ESL taught throughout the day increase at the same rate as the students' Cognitive Academic Language Proficiency (CALP) increases. Increasing English language instruction continues until the students are fully transitioned into the English language instruction. The students should complete the transition into the regular classroom by the fifth grade. Providing that they entered the program at the lower grades (i.e., kindergarten or pre-kindergarten).

Participants

Student participants were 74 Spanish-English bilingual students attending first and second grades, with approximately 18 students per classroom, in four classrooms. Their ESL category was beginner or non-English speaker (NES). Most participants were second-generation Mexican American immigrants; their parents were among the first generation to immigrate to the United States.

All students were selected for the study based upon performance on an initial skills screener. Students were selected only if they could orally read 30-60 words correctly per minute on a first reading of a Spanish story at

either first or second grade level readability, or knew their beginning sounds and could read 50 to 100 sight words. The low performance levels of these students, along with their NES language classification, placed them at risk for successful reading in English, without a targeted and effective intervention (August & Hakuta, 1997; Gersten & Jimenez, 1994; Snow et al., 1998).

Description of the Intervention

For the purpose of this study, materials from first and second grade levels of the Read Naturally (RN) program were translated to Spanish by a team that included three native Spanish speakers, one of whom is a certified bilingual translator. Direct translation was followed by rewriting to ensure natural language in all stories. With the permission of the RN author, stories were translated into Spanish text passages, and audiotapes of each story were professionally recorded in Spanish. These materials were produced for experimental use only.

The intervention was implemented with three groups of students for 45 minutes per day, three times a week for 12 weeks for Group 1, 10 weeks for Group 2, and eight weeks for Group 3. The intervention was conducted during the students' Spanish language arts block. Students not receiving the intervention participated in language arts activities in centers (a mixture of word analysis, vocabulary, and other language arts activities).

RN involves students self-graphing scores of words correct per minute (wcpm) before and after individualized repeated practice. After an initial timing by the teacher on an unpracticed story, students graph this "cold reading" score on bar graphs in their individual folders. They then read along with an audiotape of the story at a listening center. This step involves simultaneous, subvocalized reading of the entire story up to three times. After reading with the audiotape, students practice reading the story independently without the tape. They self-time each individual practice for one minute to determine if they have reached a predetermined performance goal. When the goal is reached, students signal that they are ready for a timing conducted by the teacher. Students then plot this "practiced score" on bar graphs that provide instantaneous and concrete evidence of progress and improvement. Students may pass to the next story only if they reach their targeted fluency goal, make fewer than three oral reading errors, read with appropriate expression, and accurately answer three of four comprehension questions. Students then repeat the process with another story at the same level of difficulty. After the student completes 12 stories at a level, the teacher and the student decide whether the student should continue at the same level with the same goal, adjust the reading rate goal, or move up a level to more difficult reading material (Hasbrouck, Innot, & Rogers, 1999).

RN provides 24 stories to practice at each level, in half-year increments from mid first-grade level through Grade 6. Passages range in length from approximately 60 words in the mid first-grade level to 350 or more words at the sixth-grade level. Students individually pace themselves, completing as many self-timings as possible within the instructional period.

Screening

Prior to the study, students were screened to establish their suitability for participation in the study and to place students into balanced groups. Spanish reading passages at levels 1.0, 1.5, and 2.0, obtained from Spanish basal readers adopted in Texas, were administered to students. Based on screening results, students were classified as reading at grade levels 1.0, 1.5, or 2.0. Students were placed at the level at which they could read 30–60 wcpm. Students who did not know beginning sounds and could not read 50 to 100 sight words were not included in this study. The treatment groups in the study were each composed of a balance of students at each of the three levels. A total of 74 students began the intervention. Due to absences and subsequent missing data, complete data were finally available for only 53 students.

Assessment Instruments

Reading probes

Seven semi-weekly probes were created by the first author, a certified Spanish interpreter. Passage titles were: (a) “El Cumpleaños,” (b) “La Navidad,” (c) “El Cinco de Mayo,” (d) “Los Santos Reyes,” (e) “El Baile del Cinco de Mayo,” (f) “La Comida del Cinco de Mayo,” and (g) “Las Posadas.” The shared theme was important holidays celebrated in Mexico and other Hispanic countries.

The seven probes were designed to be similar in average word length, average number syllables per word, average sentence length, and topic to the passages used in the reading intervention. The range for the total words for the seven probes was 72–80, and number of sentences ranged from 9 to 12. The range for word length was 3.57 to 4.31 (average letters per word), and sentence length ranged from 6.3 to 8.0 (average words per sentence). These countable indicators of readability are in line with typical Level 1 stories from the Spanish basal reading series used in this district. Furthermore, the topics in these passages were related to the students’ culture and background knowledge. To control for invalid results due to passage differences, the administration of these probes was counterbalanced by students within groups and classrooms.

Assessing reading fluency

Students individually read the short passages aloud and were scored for oral reading fluency (ORF). The number of words that they read during one minute, minus errors, was recorded as the words read correctly per minute (wcpm) (Shinn, 1989). Interscorer reliability for ORF was established prior to the study at Kappa = .94.

Assessing comprehension

Four short-answer questions were created for each probe, all supply-type. Students were asked questions of sequence, such as “who?” “did what?” “when?” and “why?” The probes also posed questions of prediction; passage length precluded the use of “main idea” questions. Responses were scored 0 or 1, with .5 points given for partially correct responses. An answer key described correct and partially correct responses to help maintain high interrater reliability on comprehension scoring. In addition, reliability for comprehension scoring was established at Kappa = .89.

Probe equivalence

The seven probes were counterbalanced within groups and classrooms to help negate any differences in readability. In addition, passage difficulty levels were calculated after the study based on the wcpm scores of the students. The mean scores across all students for the seven probes ranged from 53.7 wcpm to 76.3 wcpm, with $F(6,363) = 3.0$ ($p = .007$). Only two probes were outliers, Passage D being significantly easier ($t = 2.93$, $p = .003$), and Passage F being significantly more difficult than the rest ($t = 2.98$, $p = .003$). Since passages were counterbalanced within groups, this degree of non-equivalence was considered tolerable.

Design

This study utilized a multiple baseline, single-case research design. Small groups represented each of the three baselines. The first group began the treatment immediately, and the second and third groups' treatment was lagged by three and five weeks, respectively. All three treatment groups were formed in each classroom to eliminate teacher effects. In addition, initial student reading levels were matched across groups to ensure equivalent groups within each grade level. Finally, the order of administration of the seven reading probes was counterbalanced to eliminate potential effects of differential difficulty of the probes.

The treatment groups were initially formed from 74 bilingual students from two Grade 1 and three Grade 2 classrooms that met the minimum reading criterion. Within each grade level, students were randomly assigned to one of three treatment groups, beginning the treatment at Week 0, Week 3, or Week 5. Student group assignment was then manipulated to better match groups.

Resulting groups contained equal numbers of students reading approximately at grade levels 1.0, 1.5, and 2.0, respectively.

Prior to the intervention (Week 0), all participating students were assessed with the first Spanish reading probe, and all were assessed six more times, every two weeks over a 12-week period. During Week 1, students in Group 1 began the RN treatment three times per week, 45 minutes per day, and continued until the end of the study, through Week 12. At Week 3, students in Group 2 began treatment and continued throughout the 12 weeks. At Week 5, Group 3 began treatment and continued until the termination of the study.

In this design, the dependent measures included: (a) oral reading fluency (ORF) reading scores of wcpm on the semi-weekly equivalent probes, and (b) reading comprehension scores (from post-reading questions) for the same Spanish reading probes. Visual and statistical comparisons between baseline and treatment phases and between groups permitted judgments about the efficacy of the intervention.

Student Level of Engagement

Because student engagement level in the RN program was a concern and could easily be measured from student daily records, after the study each group was sub-divided into “low-engaged” (12 Grade 1; 13 Grade 2) and “high-engaged” (12 Grade 1; 16 Grade 2) students. Level of engagement was based upon the average number of reading attempts per week documented (by a mark on the reading graph) over the 12 weeks in students’ cumulative folders. In this study, low- and high-engaged results were analyzed separately and presented in separate multiple-baseline design graphs.

Treatment Fidelity

In addition to measuring levels of engagement by the student participants, treatment fidelity was monitored through direct observation during weekly visits to individual classrooms by a member of the research team. In addition, the researcher completed a “fidelity of implementation checklist” at mid-program for each classroom. Prior to and after completion of the fidelity checklist by the visiting researcher, teachers monitored themselves on the same checklist. Remedial support was offered by the research team when fidelity concerns were detected by either classroom teachers or by the researcher-observer.

Results

Student Level of Engagement

In theory, the number of recorded timed-readings should be independent of a student’s skill level. Low achieving slow readers could make several timed reading attempts because their reading passages were shorter and easier

than those read by more proficient readers. The level of engagement (average number of timed reading attempts per week) averaged 4.7 across all students, and ranged by student from 1.5 to 7.8 per week, a very large, five-fold difference.

We formed 12 equal-size level of engagement groups (six for each grade level, composed of three High-Engaged (H-Engaged) and three Low-Engaged groups (L-Engaged) according to the following criteria: L-Engaged students ($n = 25$) ranged from 1.5 to 3.0 timed reading attempts per week (mean = 2.6). H-Engaged students ($n = 28$) ranged from 3.2 to 7.8 timed readings per week (mean = 5.8).

Because level of engagement depended in part on teacher-controlled variables (active practice opportunity, efficient organization of materials), it was not surprising that the number of timed reading attempts also varied across the six classrooms. The largest differences were between the two Grade 1 classrooms ($F = 216.63$, $p < .0001$), yielding moderate-size differences ($\eta^2 = .58$). Level of engagement differences also existed between grade levels, with Grade 2 students showing higher engagement levels than Grade 1 students ($F_1 = 21.29$, $p < .0001$), also a moderate effect size difference ($\eta^2 = .31$). Very small but significant differences in level of engagement also were found between the two Grade 2 classrooms, ($F_1 = 9.3$, $p < .0027$), ($\eta^2 = .05$).

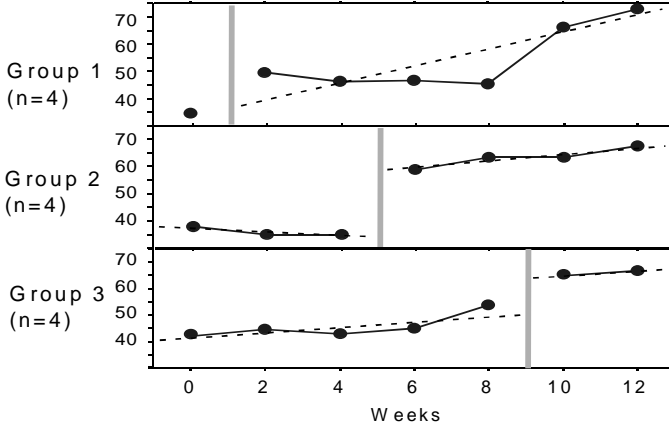
ANOVAs were conducted on the Week 0 ORF scores at each grade level to determine if these level of engagement groups were initially equivalent in ORF skills at the beginning of the study. Non-significant F-test results showed that the 12 groups were essentially similar, and post-hoc Tukey tests showed that none of the six Grade 1 groups differed significantly (at $p < .05$). At Grade 2, only one of six groups differed significantly. Therefore, the multiple baseline, single-case research graphs and analyses could be prepared separately for two levels of engagement.

Visual Analysis of ORF Progress

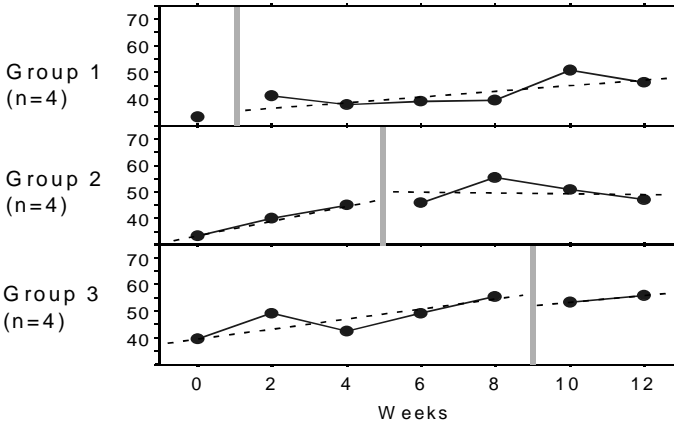
Figure 1 presents multiple baseline graphs of average performance for Grade 1 students in ORF (oral reading fluency), using the well-researched “wcpm” index. The two graphs are for H-Engaged ($n = 16$) and L-Engaged ($n = 13$) students. Regression lines have been plotted to permit more accurate judgment of trends. Vertical intervention lines divide between baseline (pre-treatment) and treatment phases.

Figure 1

Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - High Engaged Students (N=12)



Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - Low Engaged Students (N=12)



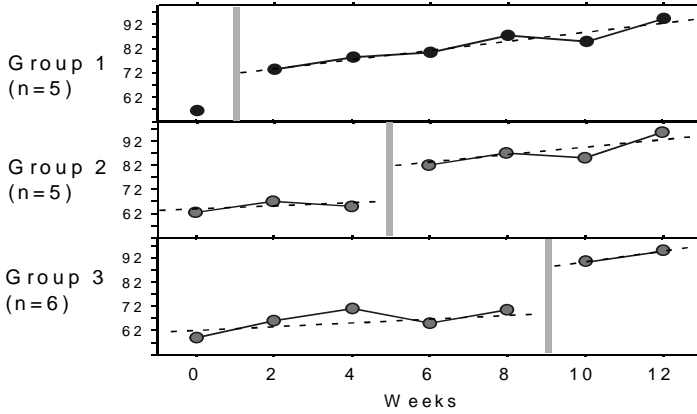
The multiple baseline graphs present only a single data point for the baseline phase for Group 1, so trend lines cannot be applied, and analyses are therefore limited. The strength of the graphs is that each data point represents four to six students, contributing to the stability and generalizability of the results.

Visual analysis of Figure 1 reflects marked improvement in performance between baseline and intervention phases for H-Engage students. For L-Engaged students, the between-phase differences are not as apparent. The type of between-phase differences noted for H-Engaged students appears to be in mean levels and in a “jump” of trend line intercepts at the intervention line. These noted changes represent two of the four types of changes that can be measured between-phases in a time series graph: (a) changes in mean level, (b) “jump” in trend line intercepts at the intervention line, (c) changes in trend line slopes, (d) changes in variability or data “bounce” (Gorman & Allison, 1996). In Figure 1 sizeable slope differences are noted only for Group 2 H-Engaged students (a positive increase) and for Group 2 L-Engaged (a negative decrease). Of these four phase change attributes, only the second and third present convincing evidence of intervention effectiveness. The first, a change in mean level, could be due to gradual improvement, regardless of the intervention.

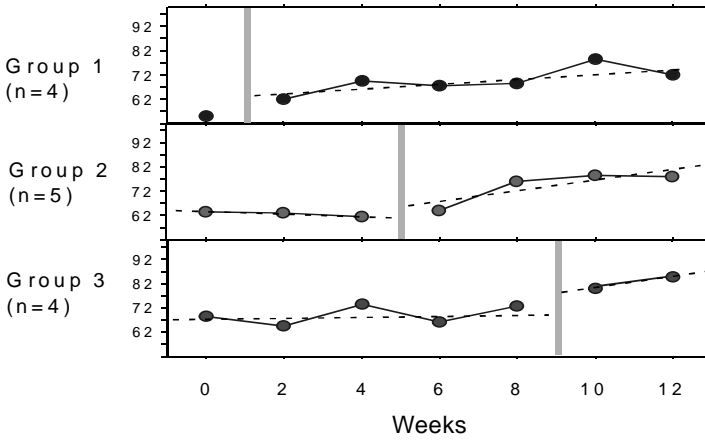
Figure 2 contains the same ORF information for the 16 H-Engaged and 13 L-Engaged Grade 2 students. Visual examination of Figure 2 shows the same type of improvements noted for H-Engaged students only, i.e., an improvement in mean performance level between phases and a perceivable “jump” in slope intercepts at the intervention line. As with Grade 1 students, no large improvements in slope are noted. Also, in keeping with the Grade 1 graph, L-Engaged Grade 2 students showed markedly less of a “jump” in trend line intercept and somewhat less of an increase in mean levels.

Figure 2

Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 2 Bilingual Students - High Engaged Students (N=16)



Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 2 Bilingual Students - Low Engaged Students (N=13)

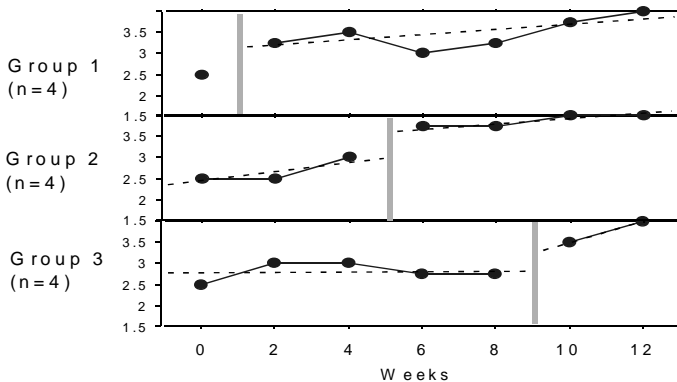


Visual Analysis of Comprehension Progress

Figure 3 presents two multiple baseline reading comprehension graphs for Grade 1 H- and L-Engaged students. Visual analysis shows a clear increase in mean performance levels between phases for H-Engaged students only. Likewise, a notable jump in the intervention line intercepts is noted for the H-Engaged but not for L-Engaged students. In fact, two of the L-Engaged graphs evidence deteriorated performance through drops in intercepts at the point of intervention. In summary, there is some evidence of intervention effectiveness for the H-Engaged students, but not so for the L-Engaged students. Furthermore, it is not even apparent that the L-Engaged students improved in reading comprehension over the 12 weeks.

Figure 3

Improvement Trends for Reading Comprehension related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - High Engaged Students (N=12)



Improvement Trends for Reading Comprehension related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - Low Engaged Students (N=12)

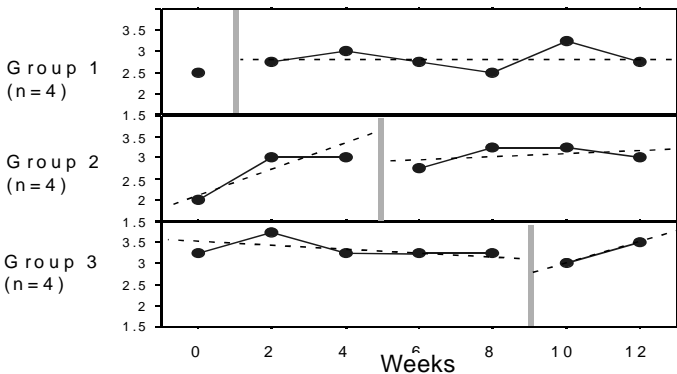
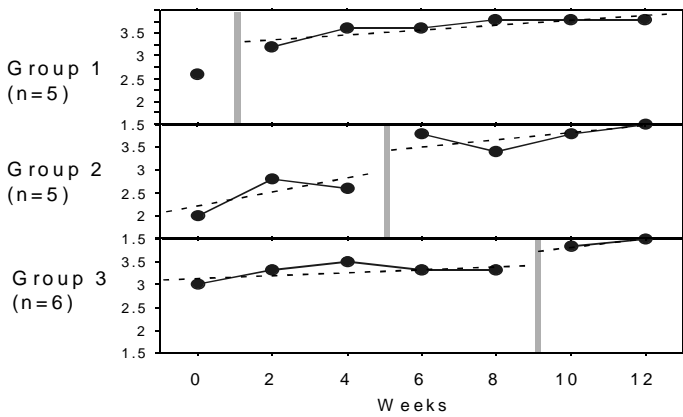


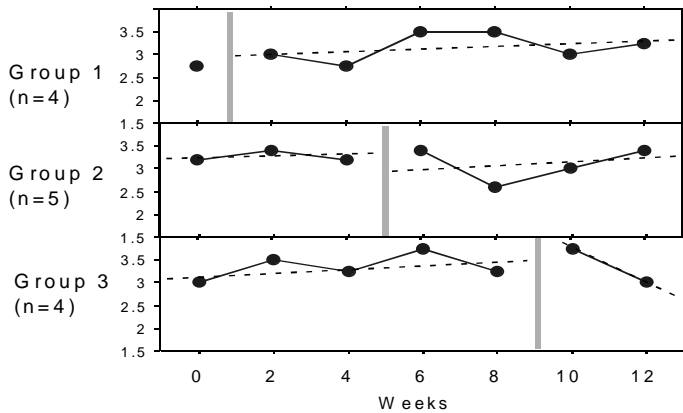
Figure 4 presents the same reading comprehension information for Grade 2 students. The pattern is nearly identical as for Grade 1 students. The H-Engaged students did improve over the 12 weeks (mean level change) and showed evidence of intervention effectiveness (“jump” in trend line intercepts). The L-Engaged students, in contrast, showed neither evidence of improvement of the period of the study, nor evidence of the effectiveness of the RN intervention.

Figure 4

Improvement Trends for Reading Comprehension related to 'Read Naturally' Treatment with Grade 2 Bilingual Students - High Engaged Students (N=16)



Improvement Trends for Reading Comprehension related to 'Read Naturally' Treatment with Grade 2 Bilingual Students - Low Engaged Students (N=13)



Statistical Analysis of ORF Progress

Visual analyses indicated that changes in slope between phases would be unlikely, so no test for slope changes was conducted. Visual examination did indicate probable between-phase differences in intercepts at the intervention line and in mean levels; those differences were the foci of statistical analyses.

Autocorrelation

To legitimately use traditional inferential parametric analyses with time series data, the regression residuals must be serially independent, i.e., not autocorrelated (Kazdin, 1982). An accepted test for autocorrelation in short data series is a lag-1 test, conducted on residuals or “detrended” data (Franklin, Allison, & Gorman, 1997). The Grade 1 groups yielded lag-1 autocorrelation coefficients of $r = -.06, .19, .14$. The Grade 2 groups yielded coefficients of $r = -.03, .16, -.04$ (none significant). Given six separate tests for autocorrelation, all with nearly zero autocorrelation, we concluded that the data were not serially dependent. Having met the assumption of serial independence, tests of inference on mean and slope differences could be conducted as with non-time series group data (Franklin, Allison, & Gorman, 1997).

Test of intercept difference at intervention line

As mentioned earlier, in a time series design, the best evidence for treatment effectiveness is visible in trend line slope changes between phases, or in a jump or gap between trend lines where they intercept at the intervention line. We first analyzed data for this visually pronounced change between phases: the “jump” in slope intercepts. The following is a technical description of the procedure.

Normal output for statistics packages (e.g., SPSS) includes Y-axis intercepts for regression slopes, with intercept standard errors of measurement. By reversing the series for Phase A data, and then recalibrating the time axis for both phase A and B to run from Day 0, we can output Y-intercepts for both phases on the same scale. We can then test for differences in the two Y-intercept values by checking for overlapping confidence intervals (calculated from their SEMeas, or standard error of measurement) or by running a student t-test: Table 1 contains results for tests of differences in intervention line intercepts for the separate graphs in figures 1 and 2.

Table 1 shows significant ($p < .05$) differences in slope line intercepts at the intervention point for all of the H-Engaged groups which could be tested. Only one L-Engaged group (Group 3) also showed a significant change in intercept at the intervention line, but that was a drop rather than a jump, indicating a decrease in performance.

Table 1

Improvement in ORF Between Phases, Through Analysis of Intervention Line Intercepts

Group	n	Phase I		Phase II		't'	.05 sig
		Intcept	SEM	Intcept	SEM		
Grade 1-High Engage							
Group 1	4	N/A	N/A	41.89	5.909		
Group 2	4	34.37	0.973	59.13	1.236	2.20	yes
Group 3	4	49.74	2.898	63.22	0	2.36	yes
Grade 1-Low Engage							
Group 1	4	N/A	N/A	37.90	2.967		
Group 2	4	49.38	0.846	52.96	4.495	2.20	no
Group 3	4	57.53	3.27	53.24	0	2.36	no
Grade 2-High Engage							
Group 1	5	N/A	N/A	73.01	1.775		
Group 2	5	66.22	2.55	81.51	3.001	2.14	yes
Group 3	6	70.33	3.13	90.64	0	2.26	yes
Grade 2-Low Engage							
Group 1	4	N/A	N/A	64.17	2.953		
Group 2	5	61.40	0.176	67.28	3.806	2.14	no
Group 3	4	66.46	1.065	79.72	0	2.36	neg.

Test of Mean Difference Between Phases

The second test conducted was for an increase in mean level between phases, regardless of slope differences or differences in Y-intercepts. A test of mean difference provides a good indication of magnitude of change. Because of the fact that the students were in the earliest grades, and because of the emphasis on reading improvement in all of these bilingual classrooms, overall mean levels from all classrooms should theoretically improve. However, we would expect greater average improvement from students in the H-Engaged groups and those who received the treatment for the greatest period of time (Group 1 vs. Group 3). Table 2 contains the mean scores for each phase, by group, for ORF.

ANOVA results in Table 2 confirm findings from visual examination of the graphs. First, most students did show overall improvement over the 12-week study (though not necessarily due to the intervention). Significant improvement from Phase A to B was made by all groups, except for the Grade 1 L-Engaged group. Strongest findings (highest significance levels) were found for both the H-Engaged group and the Grade 2 L-Engaged group. Contrary to our expectations, Group 1 students did not make the greatest gains, nor did Group 3 students make the least.

Table 2

Oral Reading Fluency Mean Improvement Between Phases for Grades 1 & 2

Group	n	Phase A		Phase B		F tests		Effective Sizes
		M	SD	M	SD	F	P	SD units
Grade 1-High Engage								
Group 1	4	34.5	3.82	54.4	1.75	93.5	.002	5.2
Group 2	4	35.8	3.24	63.1	1.42	274.3	.0005	8.42
Group 3	4	46.2	.71	64.8	3.30	144.7	.001	5.60
Grade 1-Low Engage								
Group 1	4	33.1	5.21	42.4	4.91	4.06	.13	1.78
Group 2	4	39.4	1.41	49.8	3.38	19.5	.02	7.37
Group 3	4	47.1	1.05	54.5	1.81	27.9	.01	7.04
Grade 2-High Engage								
Group 1	5	56.5	3.37	83.3	.71	307.9	<.0001	8.10
Group 2	5	64.9	3.83	87.4	1.62	149.3	.0003	5.87
Group 3	6	66.1	1.56	91.9	3.03	241.4	<.0001	16.5
Grade 2-Low Engage								
Group 1	4	55.1	7.8	69.6	2.6	8.36	.06	1.86
Group 2	5	62.6	2.47	73.9	1.49	106.5	.0005	4.57
Group 3	4	67.2	.65	82.0	3.43	110.6	.001	22.7

The amount of improvement between phases can be communicated by the effect size (standardized mean difference). Effect sizes (standardized mean differences) did not appear to vary systematically by level of engagement or by group. The effect sizes were generally very large in terms of most classroom performance (Wolf, 1986). However, the absolute size of the effects are probably overestimated because of the nature of time series data (Franklin, Allison, & Gorman 1997).

Student progress from Week 0 to Week 12 was also tested. This test replicates a typical pre/post design by not making use of all the intervening time series data. We would expect those students to make most progress who were in H-Engaged groups and those who received the treatment for the greatest period of time (Group 1 versus Group 3).

Table 3 shows significant growth from the first to the last week (over the 12 weeks) for all but one of the 12 groups. At both Grade 1 and Grade 2 levels, largest gains (in terms of effect size) were consistently made by the H-Engaged groups. Also, as predicted, H-Engaged students who received the longest treatment (Group 1) outperformed other students. However, this finding obtained only in H-Engaged groups. In L-Engaged groups, length of treatment did not appear to be predictive of a gain from the first to last week. Effect sizes (in standard deviation units) were very large, and they reflect improvements of a size that would make a practical difference in classroom functioning (Wolf, 1986). However, as noted earlier, effect sizes in time series data tend to be larger than pre/post data (Franklin, Allison, & Gorman, 1997) and should be interpreted with caution.

Table 3

ORF Improvement From First to Last Probe in Grades 1 and 2, Over Twelve Weeks

Group	n	Week 0		Week 12		F tests		Effect Sizes
		M	SD	M	SD	F	P	SD units
Grade 1-High Engage								
Group 1	4	34.5	3.82	72.9	6.50	71.4	.003	10.05
Group 2	4	37.7	7.93	67.6	4.04	31.7	.011	3.77
Group 3	4	45.2	6.29	66.5	3.71	22.4	.017	3.38
Grade 1-Low Engage								
Group 1	4	33.1	5.21	46.1	7.36	4.57	.121	2.49
Group 2	4	33.1	8.30	47.0	4.44	6.53	.083	1.67
Group 3	4	39.4	5.12	55.7	4.36	44.6	.006	3.18
Grade 2-High Engage								
Group 1	5	56.6	3.37	94.3	2.40	930.7	<.0001	11.18
Group 2	5	62.4	5.10	95.5	6.77	153.8	<.0002	6.49
Group 3	6	58.8	5.08	94.3	3.72	209.5	<.0001	6.98
Grade 2-Low Engage								
Group 1	4	55.1	7.82	71.8	2.50	36.6	.009	2.12
Group 2	5	63.7	3.01	77.7	4.78	18.1	.01	4.65
Group 3	4	68.5	3.71	84.4	2.30	104.1	.002	4.28

Statistical Analysis of Comprehension Progress

Visual examination of the comprehension graphs showed similar results to those found in the ORF graphs. That is, there were indications of skill improvement, but not in slope or rate of learning. Rather, the improvement appeared to be in mean level and in a jump in slope intercepts at the intervention line. For comprehension, we therefore performed the same analyses as for ORF.

Autocorrelation

As for ORF, we conducted lag-1 tests for autocorrelation for each data stream. Autocorrelations for Grade 1 subgroups were $r = -.010, .060, .290, -.530, -.470, -.510$, and for Grade 2, $r = .190, -.670, .100, .160, -.250, -.830$. Results varied widely, with the average absolute value being $r = .34$. Because four of the 12 tests resulted in significant autocorrelations of at least moderate size, it was decided to continue with analyses, but mainly for their descriptive usefulness, deemphasizing p-values, which may be grossly underestimated (inflating Type 1 error) for autocorrelated data (Kazdin, 1982).

Test of intercept differences at intervention line

As for ORF scores, we tested comprehension scores for a significant jump or gap between trend line intercepts at the intervention line. Table 4 contains results for tests of differences in intervention line intercepts for the separate graphs in Figures 1 and 2.

Results for comprehension are not as strong as were those for ORF. Table 5 shows that only two groups showed significant jumps in intercept levels, one H-Engaged (Grade 1, Group 3), and one L-Engaged (Grade 2, Group 3). In addition, one L-Engaged group (Grade 1, Group Three) showed a significant decrement in intercept level. The improvement in intercept level shown by the L-Engaged group was deceptive, because performance for that group dropped off sharply, actually producing a negative slope for Phase B.

Table 4

Improvements in Comprehension Scores Indicated by Tests of Intercept

Group	n	Phase I		Phase II		't'	.05 sig
		Intcept	SEM	Intcept	SEM		
Grade 1-High Engage							
Group 1	4	N/A	N/A	3.20	0.23		
Group 2	4	2.93	0.55	3.62	0.24	2.20	no
Group 3	4	2.68	0.11	3.4	0	2.36	yes
Grade 1-Low Engage							
Group 1	4	N/A	N/A	2.75	0.21		
Group 2	4	3.57	0.55	2.85	0.16	2.20	no
Group 3	4	3.25	0.18	2.75	0	2.36	neg.
Grade 2-High Engage							
Group 1	5	N/A	N/A	3.26	0.25		
Group 2	5	3.03	0.52	3.6	0.22	2.14	no
Group 3	6	3.3	0.2	3.66	0	2.26	no
Grade 2-Low Engage							
Group 1	4	N/A	N/A	3.23	0.19		
Group 2	5	3.2	0	3.14	0.77	2.14	no
Group 3	4	3.75	0.06	3.75	0	2.36	yes

Test of mean difference between phases

As for ORF, comprehension scores were tested for mean differences between the two phases. This test provides an index of amount of improvement from one phase to the next, without supporting causal inferences.

Table 5 shows that 10 of the 12 groups improved in comprehension from Phase A to Phase B, the exceptions being two L-Engaged groups. A pronounced pattern is evident: H-Engaged students produced effect sizes of two-to-four times the effect sizes of L-Engaged students. Effect sizes in the range of 1 to 3 SD units are very large—certainly large enough to be noted in daily classroom performance. The p values associated with the F tests cannot be interpreted directly because of autocorrelation, but still do indicate the same patterns seen in the effect sizes.

Table 5

Mean Improvement in Comprehension Between Phases for Grades 1 & 2

Group	n	Phase A		Phase B		F tests		Effective Sizes
		M	SD	M	SD	F	P	SD units
Grade 1-High Engage								
Group 1	4	2.50	.57	3.45	.08	9.28	.05	1.66
Group 2	4	2.60	.72	3.87	.14	13.63	.03	1.76
Group 3	4	2.83	.57	3.25	.35	1.22	.35	.73
Grade 1-Low Engage								
Group 1	4	2.50	.57	2.83	.49	.49	.53	.57
Group 2	4	2.67	.72	3.06	.23	.23	.39	.68
Group 3	4	3.41	.56	3.25	.20	.20	.69	-.28
Grade 2-High Engage								
Group 1	5	2.60	.54	3.63	.21	.21	.007	1.90
Group 2	5	2.46	.50	3.75	.17	.17	.009	2.58
Group 3	6	3.27	.39	3.62	.20	.20	.05	.92
Grade 2-Low Engage								
Group 1	4	2.75	.95	3.16	.27	.27	.50	.43
Group 2	5	3.26	.27	3.10	.22	.22	.08	-.59
Group 3	4	3.25	.41	3.43	.31	.31	.49	.43

Table 6 shows a clear pattern of comprehension improvement from the first to last assessment. H-Engaged students showed greater improvement than L-Engaged students. In fact, growth of all six H-Engaged groups reached statistical significance, whereas none of the six L-Engaged groups did so. Effect sizes of the H-Engaged groups were large enough to be relevant in typical classroom performance; however, no scores approached the size of those obtained for ORF. Across all Grade 1 students, the average gain for H-Engaged students was 38 “percent correct” points, and across Grade 2, the gain was 34 “percent correct” points. H-Engaged Grade 1 students gained 13 percentage points, whereas their L-Engaged classmates gained only five points.

Table 6

Comprehension Improvement in Grades 1 and 2 From First to Last Probe, Over 12 Weeks.

Group	n	Week 0		Week 12		F tests		Effect Sizes
		M	SD	M	SD	F	P	SD units
Grade 1-High Engage								
Group 1	4	2.50	.57	4.00	.00	27.0	.01	2.63
Group 2	4	2.50	1.0	4.00	.00	9.00	.05	1.5
Group 3	4	2.50	.57	4.00	.00	27.0	.01	2.63
Grade 1-Low Engage								
Group 1	4	2.50	.57	2.75	.95	1.58	.71	.43
Group 2	4	2.00	.81	3.00	.00	6.00	.09	1.23
Group 3	4	3.25	.95	3.50	.57	.158	.71	.26
Grade 2-High Engage								
Group 1	5	2.60	.54	3.80	.44	10.2	.03	2.22
Group 2	5	2.00	.70	4.00	.00	40.0	.003	2.85
Group 3	6	3.00	.63	4.00	.00	15.0	.01	1.58
Grade 2-Low Engage								
Group 1	4	2.75	.95	3.25	.50	3.00	.18	.52
Group 2	5	3.20	.44	3.40	.54	1.00	.37	.45
Group 3	4	3.00	.81	3.00	.00	N/A	N/A	00

Discussion

This study examined the intensive Read Naturally fluency intervention for at-risk, beginning L1 Spanish readers in grades 1 and 2. This intervention combined three instructional methods—repeated reading, teacher modeling, and progress monitoring—supported in research literature, but only for at-risk English readers.

The multiple baseline across subjects, single-case research design provided complete data from 53 students from four first- and second-grade transitional bilingual classrooms, in three treatment groups with lagged interventions over 12 weeks. Would the Read Naturally fluency-building strategy, using Spanish stories, administered for 45 minutes, three times a week, over 12 weeks, improve reading fluency and reading comprehension for students with varying levels of engagement in the materials? Because level of engagement was an important variable in this research, oral reading fluency and reading comprehension improvement analyses were conducted separately by different levels of engagement (high versus low).

It was predicted that students who received the intervention throughout the 12 weeks of the study would improve the most in ORF and reading comprehension. The time series analysis made it possible us to detect whether reading improvement was likely due to the intervention. It was also predicted that students who practiced the most (were highly engaged) regardless of the length of the intervention, would improve the most.

The design included several features to strengthen internal validity. Classroom and teacher effects were controlled by embedding treatments within classrooms. Treatment group equivalence was achieved through initial random assignment, followed by matching to balance. Progress monitoring probes were counterbalanced in order of administration to ensure equivalence within and across groups, as well as over time.

Data were analyzed separately for grades 1 and 2, using both visual analysis of graphs and through statistical analyses of: (a) differences in phase-based slope intercepts at the intervention line, (b) pre- post-test differences (between the first and last assessment), and (c) mean differences between baseline and intervention phases.

There were six general findings from this study. What was apparent from visual examination of the multiple baseline graphs was largely confirmed by statistical analysis. The first general finding was that most students—regardless of treatment group or phase—improved measurably in both oral reading fluency and reading comprehension.

The second general finding was that the amounts of improvement shown by students in the high-engagement group were large enough to make a difference in daily classroom performance (Suen & Ary, 1989). The size of improvement can be gauged in part by comparing the growth of these students with national norms and performance standards for ORF scores (Fuchs et al.,

1993; Hasbrouck & Tindal, 1992). According to these guidelines, most H-Engaged students made progress exceeding ambitious goals of three words per week. The students in Grade 1 started the intervention reading at an average of 37.2 wcpm but were at 59.3 wcpm at the end of the intervention, at the expected fluency level for first graders, 60 wcpm (Hasbrouck & Tindal, 1992). The second graders started at an average of 60.7 wcpm and ended at 87.2 wcpm, a significant improvement but still below average for second graders at the end of the school year, 94 wcpm (Hasbrouck & Tindal, 1992). This improvement in wcpm was statistically significant for both grades.

There was also improvement in reading comprehension. The 24 students in Grade 1 improved their comprehension scores significantly by 35 percentage points, from 64% to 89% correct. At Grade 2, students also improved significantly, from 70% to 91% correct. Another measure of improvement, effect sizes, were computed on student improvement from one phase to the next and from the first to last assessment. In both tests, effect sizes were substantial enough to be noticed in classroom performance.

The third general finding was that students at both grades 1 and 2 who were more highly engaged in the instructional materials showed more progress over time. Engagement level was measured as the number of timed readings (regardless of score) attempted by a student over the period of 12 weeks. Engagement level was largely controlled by students themselves. They could largely pace themselves, choosing how much to practice and how frequently to try to improve their performance through timing on the record graph. In Grade 1, H-Engaged groups improved an average of 30 ORF points, compared to only 10 ORF points improvement for L-Engaged groups. At Grade 2 the same pattern obtained: H-Engaged students improved 37 ORF points, compared to only 17 ORF points for L-Engaged groups. These findings are supported by the literature, affirming that active practice time is a good index for opportunity to learn and commonly relates to degree of skill improvement (Berliner, 1979; Stallings, 1980; Wyne & Stuck, 1982). A summary of the “time on task” literature (Calfée & Drum, 1986) concludes that slow readers can be helped by having them “practice, practice, practice” (p. 821). Other researchers concur that practice is a key to improving fluency for poor readers (Adams, 1990; National Reading Panel, 2000).

A further note is warranted on the teacher’s role in student level of engagement. Differences in student levels of engagement were substantial. Neither student skill level nor grade level could account for the differences. Some possible reasons for the differences clearly rested with the teachers. Classroom observations during the study suggested that the amount of student practice was closely tied to teacher involvement. The teachers needed to keep students on task, provide quiet time, and motivate them. The teacher also visited with students individually and provided timing opportunities for students to meet their ORF goals and move to the next level. The teacher was in charge of several required tasks including organizing materials, timing each student,

monitoring students to insure they followed all the steps in the strategy, and structuring uninterrupted time in the daily schedule for practice. Teachers played a key role in accurate and successful implementation of this intervention.

Students' levels of engagement were in some cases related to the teachers' fidelity of program implementation. Several measures were taken to assure fidelity. Teachers were asked to self-monitor to ensure that critical intervention components were in place. Each classroom was visited an average of once per week to monitor implementation, and a fidelity checklist was completed in mid-program for each classroom. The checklist identified common problems such as "lack of focus," "competing activities," "interruptions," and "taking procedural shortcuts." In addition, unexpected events intruded; one teacher had to be out several days because her husband was in the hospital. Her absence appeared to have a negative effect on the intervention.

The fourth general finding from this study was that the type of improvement noted in most groups was not an increase in trend line or rate of learning, but instead, a rather sudden jump in performance, which was maintained for the remainder of the study. This was unexpected, given the focus on "rate of improvement" and "goal lines" found in much of the literature on student progress monitoring. This type of learning may be due to quickly learning and applying a new general strategy, or the reason could be more motivational.

The fifth general finding was that the length of time within the Read Naturally treatment was not a strong predictor of student gains. Many students in the shortest intervention (4 weeks) progressed as much as those participating the longest (12 weeks). This finding was also unexpected and raises questions about the potential efficacy of relatively short (and concomitantly inexpensive) intensive interventions to improve fluency. In this study, it could be that students who were "lagged" into the study in a late-starting group were motivated to work hard to catch up with their classmates who had begun earlier. These are typical of possible confounds when conducting research with multiple treatment groups within a single classroom.

The sixth general finding was that students' oral reading fluency (ORF) improved more than their reading comprehension. This was expected, because the Read Naturally intervention dedicates much more time to fluency practice than to vocabulary and comprehension. It is also true that for second language learners, comprehension is more difficult to improve, because comprehension is frequently compounded with a limited vocabulary and a narrow range of background experiences (Anderson & Roit, 1998; Wink & Putney, 2000). In addition, we can readily expect ORF to improve more quickly and easily in part because the metric used to assess (wcpm) is highly sensitive to growth over time. Its psychometric properties to assess fluency gains are much stronger than those of post-reading questions to assess comprehension. Finally, it has been noted that students must first obtain a level of fluent, or automatic, reading before they can attend sufficiently to text to improve their comprehension (Samuels, 1979).

Conclusions

Based on oral reading fluency and reading comprehension results, we found the Read Naturally strategy provided in Spanish (using repeated reading, teacher modeling, and progress monitoring) was effective in improving the oral reading fluency, and to a lesser degree, reading comprehension, for at-risk, beginning bilingual readers who are highly engaged with the strategy. Results over a 12-week study were substantial, and likely to make a difference in daily reading performance. The commitment of 45 minutes per day, three days per week, in a highly motivating program, was not difficult to maintain with reasonable fidelity in most of the classrooms involved in the study.

Informal observations also indicated that this intervention had a positive impact on students' motivation and self-esteem. The participating students appeared motivated by the fact that their scores improved as a result of their efforts. The RN strategy also reminded students that they had a goal and that they needed to practice reading every day in order to reach it. Teachers were also more likely to praise and encourage students because the regular timings provided direct evidence to them that even these struggling readers were indeed making progress. Following the termination of the study several teachers and students requested to be able to continue using the RN strategy.

This study contributes to the existing foundation on empirically supported reading interventions in Spanish for use in the United States. For monolingual Spanish students to successfully exit from bilingual programs, they need accelerated progress in reading—both in Spanish and in the transition to English (Bernal, 1994). Funding realities and different philosophical and political points of view are creating increased pressure to show improvements in reading for bilingual students in the early grades. This study indicates that if carried out with fidelity, and with high levels of student engagement, short-term intensive interventions can have a noticeable impact in bilingual classrooms.

References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Anderson, V., & Roit, M. (1998). Reading as a gateway to language proficiency for language-minority student in the elementary grades. In R. M. Gersten & R. T. Jimenéz (Eds.), *Promoting learning for culturally and linguistically diverse students* (pp. 42–56). Belmont, CA: Wadsworth.
- August, D., & Hakuta, K. (1997), (Eds.). *Improving schooling for language minority children: A research*. National Research Council and Institute of Medicine. Washington, DC: National Academy Press.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122–147.

- Bentz, J., & Fuchs, L.S.(1992). Curriculum-based measurement and instructional planning. *The Learning Consultant Journal*, 13, 34–37.
- Berliner, D. C. (1979). Tempus educare. In P. L. Peterson & H. J. Walberg (Eds.), *Research on Teaching*, (pp. 122–125). Berkeley, CA: McCutchan.
- Bernal, J. (1994). A historical perspective of bilingual education in Texas. In R. Rodríguez, N. J. Ramos, & J. A. Ruiz-Escalante (Eds.), *Compendium of readings in bilingual education: Issues and practices* (pp. 294–300). San Antonio, TX: Texas Association for Bilingual Education.
- Bilingual Education Act. (1968). 20th Cong., Pub. Law. 90-247.
- Breznitz, Z. (1987). Increasing first graders' reading accuracy and comprehension by accelerating their reading rates. *Journal of Educational Psychology*, 79 (3), 236–242.
- Calfee, R., & Drum, P. (1986). Research on teaching reading. In M. C. Wittrock *Handbook of research on teaching*, 28, 821. New York: Macmillan.
- Carbo, M. (1978). Teaching reading with talking books. *The Reading Teacher*, 32, 267–273.
- Carnine, D. W., Silbert, J., & Kameenui, E. J. (1997). *Direct instruction reading* (3rd ed.) Columbus, OH: Merrill Publishing Co.
- Cummins, J. (1989). *Empowering language minority students*. Sacramento, CA: California Association for Bilingual Education.
- Cummins, J. (2000). Bilingual literacy, empowerment, and transformative pedagogy. In J. V. Tinajero & R. A. Devillar (Eds.), *The power of two languages: Effective dual-language use across the curriculum* (pp. 9–19). New York: McGraw-Hill.
- Dahl, P. R. (1974). *An experimental program for teaching high-speed word recognition and comprehension skills*. Washington, DC: National Institute of Education, Office of Research. (ERIC Document Reproduction Service ED 099812, Final Report Project No. 3–1154).
- Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, 22 (4), 389–406.
- Eldredge, J. L., & Quinn, D. W. (1988). Increasing reading performance of low-achieving second graders with dyad reading groups. *Journal of Educational Research*, 82 (1), 40–46.
- Farnham-Diggory, S. (1972). *Cognitive processes in education: A psychological preparation for teaching and curriculum development*. New York: Harper & Row.
- Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. *Reading Research Quarterly*, 15, 30–48.

- Franklin, R. D., Allison, D. B., & Gorman, B. S. (1997). *Design and analysis of single-case research*. Hillsdale, NJ: Erlbaum.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53 (3), 199–208.
- Fuchs, L. S. (1993). Enhancing instructional programming and student achievement with curriculum-based measurement. In J. J. Kramer (Ed.), *Curriculum-based assessment* (pp. 65-104). Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska.
- García, E. (1994). The education of linguistically and culturally diverse students: Effective instructional practices. In R. Rodríguez, N. J. Ramos, & J. A. Ruiz-Escalante (Eds.), *Compendium of readings in bilingual education: Issues and practices* (pp. 3–13). Austin, TX: Texas Association for Bilingual Education.
- Genesee, F. (1987). *Learning through two languages*. Cambridge, MA: Newbury House.
- Gersten, R. M., & Jimenez, R. T. (1994). A delicate balance: Enhancing literacy instruction for students of English as a second language. *The Reading Teacher*, 47 (6), 438–449.
- González, J.M. (1994). Bilingual education: a review of policy and ideologies. In R. Rodríguez, N. J. Ramos, & J. A. Ruiz-Escalante (Eds.), *Compendium of readings in bilingual education: Issues and practices* (p. 8). Austin, TX: Texas Association for Bilingual Education.
- Gorman, B. S., & Allison, D. B. (1996). Statistical alternatives for single-case designs. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 159–214). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hasbrouck, J. E., Ihnot, C., & Rogers, G. H. (1999). Read Naturally: A strategy to increase oral reading fluency. *Reading Research & Instruction*, 39(1), 27–38.
- Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, 24, 41–44.
- Hasbrouck, J. E., Woldbeck, T., Ihnot, C., & Parker, R. I. (1999). One teacher's use of curriculum-based measurement: A changed opinion. *Learning Disabilities: Research & Practice*, 14 (2), 118–126.
- Heckelman, R. G. (1969). A neurological impress method of remedial reading instruction. *Academic Therapy*, 5 (4), 277–282.
- Herman, P. A. (1985). The effect of repeated readings on reading rate, speech pauses, and word recognition accuracy. *Reading Research Quarterly*, 20, 553–565.

- Homan, S. P., Klesius, J. P., & Hite, C. (1993). Effects of repeated readings and nonrepetitive strategies on students' fluency and comprehension. *Journal of Educational Research, 87* (2), 94–99.
- Huck, C. S., Hepler, S., & Hickman, J. (1987). *Children's literature in the elementary classroom* (4th ed.). New York: Holt.
- Ihnot, C. (1997). *Read naturally*. St. Paul, MN: Read Naturally.
- Jimenez, R. T., García, G. E., & Pearson, P. D. (1995). Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal, 32*, 67–97.
- Kaufman, P., & Frase, M. J. (1990). *Dropout rates in the United States: 1989*. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings, 6*, 126–151. New York, NY: Oxford University Press.
- Koskinen, P. S., & Blum, I. H. (1984). Repeated oral reading and the acquisition of fluency. In J. Niles & L. Harris (Eds.), *Changing perspectives on research and in reading/language processing and instruction*, (pp. 183–187). Rochester, NY: National Reading Conference.
- Krashen, S. (1992). *Fundamentals of language education*. Torrance, CA: Laredo Publishing.
- Laberge, D. (1973). Attention and the measurement of perceptual learning. *Memory and Cognition, 1*, 268–276.
- Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.
- Lanauze, M., & Snow, C. E. (1989). The relation between first- and second-language skills: Evidence from Puerto Rican elementary school children in bilingual programs. *Linguistics and Education, 1*, 323–340.
- Leasher-Madrid, D., & Garcia, E. (1985). The effect of transfer in bilingual proficiency. In E. Garcia & R. Padilla (Eds.), *Advances in bilingual education research* (pp. 53–70). Tucson: University of Arizona Press.
- McCormick, C., & Samuels, S. J. (1976). *Word recognition by second graders: The unit of perception and interrelationships among accuracy, latency, and comprehension*. [Research report No. 102]. Minneapolis, MN: University of Minnesota Research Development and Demonstration Center in Education of Handicapped Children.
- Medina, M. (1988). *Native language proficiency and Spanish achievement in an eight year maintenance bilingual education program*. Paper presented at the annual meeting of the National Association for Bilingual Education, Houston, TX.
- Nathan, R. G., & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory Into Practice, 30* (3), 176–184.

- 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 Institute of Child Health and Human Development.
- O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1985). The effects of repeated readings and attentional cues on reading fluency and comprehension. *Journal of Reading Behavior, 17* (2), 129–142.
- Prado, E. B., & Tinajero, J. V. (2000). Literacy instruction through Spanish: Linguistic, cultural, and pedagogical considerations. In J. V. Tinajero & R. A. Devillar (Eds.), *The power of two languages: Effective dual-language use across the curriculum* (pp. 42–53). New York: McGraw-Hill.
- Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly, 20* (2), 180–188.
- Rasinski, T. V. (1989). Fluency for everyone: Incorporating fluency instruction in the classroom. *Reading Teacher, 42*, 691.
- 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.
- Samuels, S.J. (1979). The method of repeated readings. *The Reading Teacher 3*, 403–408.
- Shinn, M. R., (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford Press.
- Skager, R. W. (1984). *Organizing schools to encourage self-direction in learners*. Oxford / Hamburg: Pergamon Press / Unesco Institute for Education.
- Skinner, C. H., Logan, P., Robinson, S. L., & Robinson, D. H. (1997). Demonstration as a reading intervention for exceptional learners. *School Psychology Review, 26* (3), 437–447.
- Smith, M. S. (1991). *The difficulties of a curriculum helper in an urban school*. Paper presented at the Bergamo Conference, Dayton, OH. (ERIC Document Reproduction Service No. ED 357090)
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press
- Soltys, M. L. (1997). *Effects of setting personal learning goals on adolescent achievement in language arts skills*. Dissertation Abstracts (AAT 9717997), Wilmington College, DE.
- Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher, 9*, 11–16.

- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 11*, 360–406.
- Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale, NJ: Lawrence Erlbaum.
- Texas Education Agency. (1997). *Assessment requirements for students of Limited English Proficiency in the state testing program*. Austin, TX: Author.
- Torgeson, J. K. (2000). *Lessons learned from intervention research in reading: A way to go before we rest*. Paper presented at the Council for Exceptional Children conference, Vancouver, B.C., Canada.
- Troike, R. (1978). Research evidence for the effectiveness of bilingual education. *Bilingual education paper series, 2*. Los Angeles: California State University, National Dissemination and Assessment Center.
- Turpie, J. J., & Paratore, J. R. (1995). Using repeated reading to promote reading success in a heterogeneously grouped first grade. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer (Eds.), *Perspectives on literacy research and practice: Fortyfourth yearbook of the national reading conference* (pp. 255–264). Chicago, IL: National Reading Conference, Inc.
- U.S. Bureau of Census. (1992). *Population projections of the United States, by age, sex, race, and Hispanics origin: 1992 to 2050* (No. P25-1092). Washington, DC: U.S. Government Printing Office.
- U.S. Department of Education (1997). National Center for Education Statistics. 1997. *Digest of Education Statistics*. Washington, DCAuthor.
- Valdivieso, R., & David, C. (1988). *U.S. Hispanics: Challenging issues for the 1990s*. Washington, DC: Population Trends & Public Policy.
- Wang, M. C., & Peverly, S. T. (1986). The self-instructive process in classroom learning contexts. *Contemporary Educational Psychology, 11*, 370–404.
- Wink, J., & Putney, L. G. (2000). Turning transformative principles into practice: Strategies for English-dominant teachers in a multilingual context. In J. V. Tinajero & R. A. Devillar (Eds.), *The power of two languages: Effective dual-language use across the curriculum* (pp. 175-186). New York: McGraw-Hill.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Sage: Beverly Hills, CA.
- Wyne, M. D., & Stuck, G. B. (1982). Time and learning: Implications for the classroom teacher. *Elementary School Journal, 83*, 67-75.