



Research Abstract IV
March 2009

Does *Tune Into Reading* Improve FCAT Scores?

* A descriptive synthesis of three years of FCAT Developmental Scale Score (DSS) gains for students who used *Tune Into Reading* in Elementary, Middle, and High School.

* Primary funding by FLDOE/Just Read! Florida.

Susan Homan, Ph.D.
Cindy Calderone, Ph.D.
Robert Dedrick, Ph.D.
University of South Florida

Key Findings:

- Over three consecutive years, struggling readers that used *Tune Into Reading* (TiR) outperformed non-TiR users on the Florida Comprehensive Assessment Test (FCAT) in year-over-year gains by a range of 32.6% to 117.9%.
- No other independent studies have been found that show a reading intervention having a positive effect on FCAT scores over multiple years.

TABLE OF CONTENTS

Introduction	3
Brief Overview of Research Studies that Used Tune Into Reading (TiR) 2004-2008	3
Purpose of Abstract IV	4
Method/Procedure: “Achievement Level” vs. “Developmental Scale Score”	4
Results – DSS Change	5
Discussion	6
Conclusions	7
Appendix A – DSS Change Comparison Charts	8
Appendix B – What is Tune In to Reading?	12
Appendix C – What is FCAT?	13
Bibliography	14

Introduction

This is the fourth abstract in a series that reports on continuing research using an interactive singing software program with struggling readers. Abstract I reported the pilot study results from a middle school population (2004-2005). Abstracts II and III reported results from the replication and expansion studies (2005-2006), involving 300 plus struggling readers in six schools in three school districts.

Brief Overview of Research Studies that Used Tune Into Reading (TiR) 2004-2008

In the spring of 2004, my research team became intrigued with the use of singing to improve reading for struggling readers. A pilot study demonstrated that 8th graders reading at 4th grade level made an average instructional reading level gain of 1.2 years in nine weeks (Biggs, Homan, Dedrick, Rasinski & Minick, 2008). Three years of subsequent studies, funded by the Florida Department of Education and Just Read! Florida, validated the initial study results. After using TiR, the mean instructional reading level for more than 1000 struggling readers increased by more than one year at the end of the 9-week study. Reports of each study are available at: (<http://www.coedu.usf.edu/main/departments/ce/Homan.html>).

Table 1 provides a list of research studies using the TiR intervention with struggling readers from 2004 to 2008.

Table 1: 4-Year Research Overview

Study Year	School Year	Description	Grade Levels	# of Study Subjects
1	04/05	Pilot study and sustainability analysis	7 th , 8 th	46
2	05/06	Replication/expansion	5 th , 8 th , & 11 th	252
3	06/07	English language learners	6 th	79
		Title 1 elementary level	3 rd , 4 th , 5 th	302
4	07/08	4 elementary schools 2 Boys & Girls Clubs 1 rural high school 8 Juvenile Justice sites	4 th , 5 th 9 th , 10 th , 11 th	393

Total Study Subjects: 1,072

Purpose of Abstract IV

From 2005 to 2009 the USF research team has presented results from the studies listed in Table 1 at over a dozen national conferences including The International Reading Association, The National Reading Conference and The American Reading Forum. In addition, results have been published in two peer-reviewed journals *Reading Psychology* (Biggs, Homan, Dedrick, Rasinski & Minick, 2008), and *Reading and Writing Quarterly* (Rasiniski, Homan, & Biggs, 2009).

As the research team presented these consistent, positive study results, we were typically asked questions about the intervention's impact on standardized, high-stakes test scores. At presentations we gave in Florida, these questions were about the impact on the reading portion of the Florida Comprehensive Assessment Test (FCAT). Therefore, the purpose of Abstract IV is specifically to answer the question "Does using *Tune Into Reading* improve FCAT scores?"

Method/Procedure: "Achievement Level" vs. "Developmental Scale Score"

To answer the foregoing question, we performed a descriptive synthesis of three years of results from three individual studies comprising a total of six grade level groups.

Most Florida educator's, parents, and students are familiar with the "Achievement Levels" that are commonly reported for the FCAT. These levels are numbered 1-5 with 5 being the highest proficiency grouping. Although it may seem logical to use the Achievement Levels to measure student progress from year to year, according to the Technical Assistance paper (DPS: 03-065-FL DOE), the proper method of tracking gains is to use the amount of change from one year to the next in the Developmental Scale Score (DSS) of the student:

Additionally, the following explanation is from the 2007 Assessment and Accountability Briefing Book, published by FL DOE.

"A different measurement, called a developmental scale score (DSS) is used to understand whether a student has "gained" in achievement. For the individual student reports (started in 2002), student scores are linked to the scale score of 100 to 500 and converted to scores on the developmental scales. The FCAT developmental scores range from 0 to about 3000 across Grades 3-10 and link two years of student FCAT data that track student progress over time. By using FCAT developmental scores, parents and educators can assess changes in scores across years and monitor the student's academic progress from one grade to the next. Each year, student scores should increase according to the student's increased achievement."

Therefore, the analysis in this abstract is based on the amount of DSS change from one year to the next for students in the studies.

Criteria for inclusion in the descriptive synthesis were:

1. students had a matched DSS reading score from two consecutive years
2. students used the TiR reading intervention for at least nine weeks
3. FCAT DSS reading scores from two consecutive years were available for the matched comparison students, or
4. FCAT DSS reading scores from two consecutive years was available for the matched grade level from the school/district/and state.

The DSS reading scores of students who used TiR were compared to DSS scores of matched and/or randomized samples. Randomized control groups were used whenever possible, but when not available we used the district, state, and relevant school level DSS gain scores as comparison groups. The changes in Reading DSS from 2005-2006, 2006-2007, and 2007-2008 are shown in Table 2 below.

Results – DSS Change

Table 2 presents grade levels, DSS change, and incremental gains from studies spanning three years of research with struggling readers using TiR.

Table 2: Synthesis Results

Year	Grade	DSS Change Non-TiR	DSS Change TiR	Incremental Gain
June 05 – June 06	5 th Grade	152	228	50.0%
	8 th Grade	187	248	32.6%
	11 th Grade	65	88	35.4%
June 06 – June 07	5 th Grade	78	170	117.9%
June 07 – June 08	4 th Grade	207	312	50.7%
	5 th Grade	71	116	63.4%

* Note: For detailed results for each group, please see Appendix A.

Descriptive synthesis of six groups of students over three years provides a consistent pattern of TiR student reading growth as measured by FCAT DSS reading scores (full analysis will be available on my webpage Summer, 2009). Students in both the TiR and non-TiR groups were reading below level as measured by FCAT (Levels 1 or 2). While the TiR students implemented the TiR intervention, the non-TiR students were using a variety of state approved intervention programs.

Our review of the data provided evidence of a dramatic increase in one-year DSS change in all three of the years when comparisons were possible. The consistency and size of the gain was unexpected and remarkable.

Discussion

No Child Left Behind (NCLB) legislation has provided a national focus on struggling readers. There are now so many different reading interventions available, that independent research validating the efficacy of the products is of great help to educators making decisions on which reading intervention to choose for their students.

One highly anticipated federally-funded, multi-year study has examined the impact of reading and mathematics software interventions on standardized test scores. The first report from that study entitled *Effectiveness of Reading And Mathematics Software Products: Findings from The First Student Cohort* was published by the Institute of Education Sciences (IES) in March 2007. The report from the second cohort, was released in February, 2009. The results from both cohorts were not encouraging. They reported that test scores were not significantly higher in classrooms using the selected software products. The products included Leapfrog, Read 180, Academy of Reading, and Knowledge Box. In fact, the first report states:

“Fourth grade reading products did not affect test scores by amounts that were statistically different from zero.”

(NCEE 2007 – 4006. p. xi)

Additionally, the report on the second cohort states:

“Similarly, in fourth grade, the measured product effect in the second year is not statistically significantly larger than the effect in the first year.”

(NCEE 2009-4042, p. xx)

At the state level, educators can find some information on interventions at the Florida Center for Reading Research (FCRR) website. However, there were minimal studies reported using Florida’s FCAT, our high stakes assessment tool for measuring student progress. After a search of the Just Read! Florida website and the FCRR website, as well as dissertations from all universities in Florida, my research team found very few independent studies to validate a positive effect of any reading intervention on FCAT scores. We found no studies of interventions with multi-year student FCAT gains.

We are continuing to search for publications and technical reports using FCAT results. My research team would appreciate information on independent studies of reading interventions with progress measured by FCAT scores. (homan@tempest.coedu.usf.edu).

Currently, we can report finding no evidence of an intervention other than *Tune Into Reading* that has independent research supporting multi-year student gains on FCAT.

This makes the results reported in this abstract even more unusual. The FCAT reading gains for students using TiR over the past three years are strong and impressive. They

appear consistent for elementary, middle, and high school students. Innovative interventions that can produce consistent FCAT Developmental Scale Score gains over a range of grade levels are rare.

Conclusions

Students who used Tune Into Reading have consistently outperformed non-TiR users on the reading portion of the FCAT over a 3-year period. Over 1,000 students have now participated in controlled studies using the high stakes FCAT assessment. These results are both unvarying and unique.

From 2005 to 2008 students from three school districts representing nine different schools and four different grade levels (4th, 5th, 8th, and 11th) achieved Developmental Scale Score gains that demonstrate striking improvement for students using TiR. While teachers and other curriculum have certainly contributed to struggling readers' gains, over three years and nine sites the TiR pattern of improvement is difficult to ignore as the primary source of reading improvement.

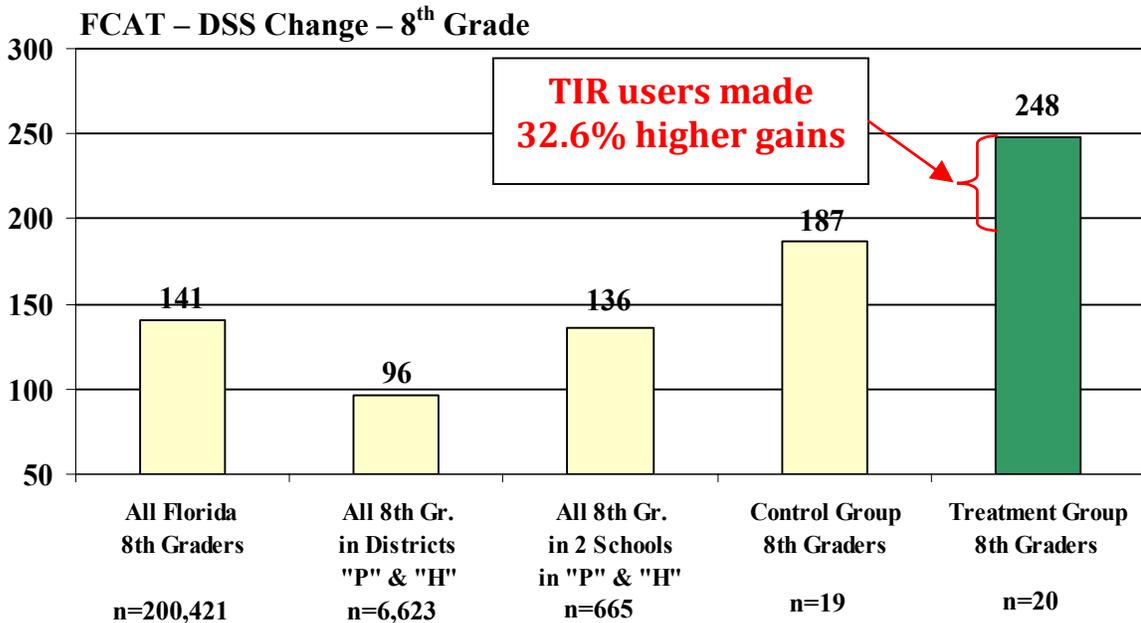
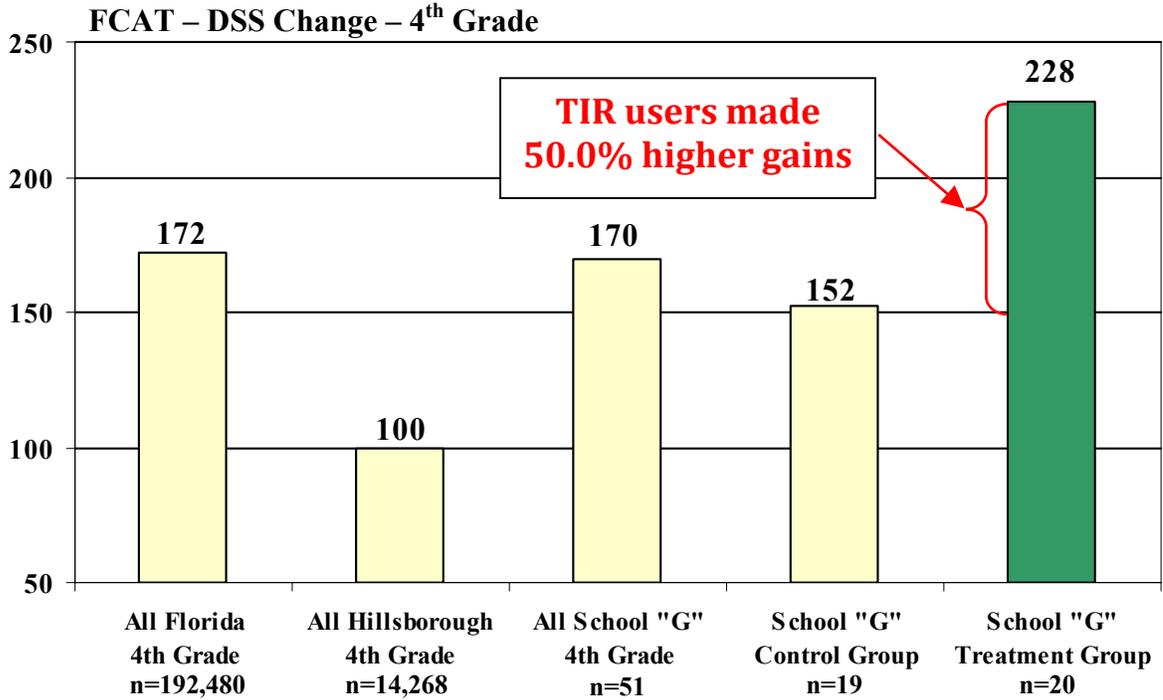
These dramatic and consistent results are extremely encouraging news for our struggling readers and the educators who work with them.

Appendix A – DSS Change Comparison Charts

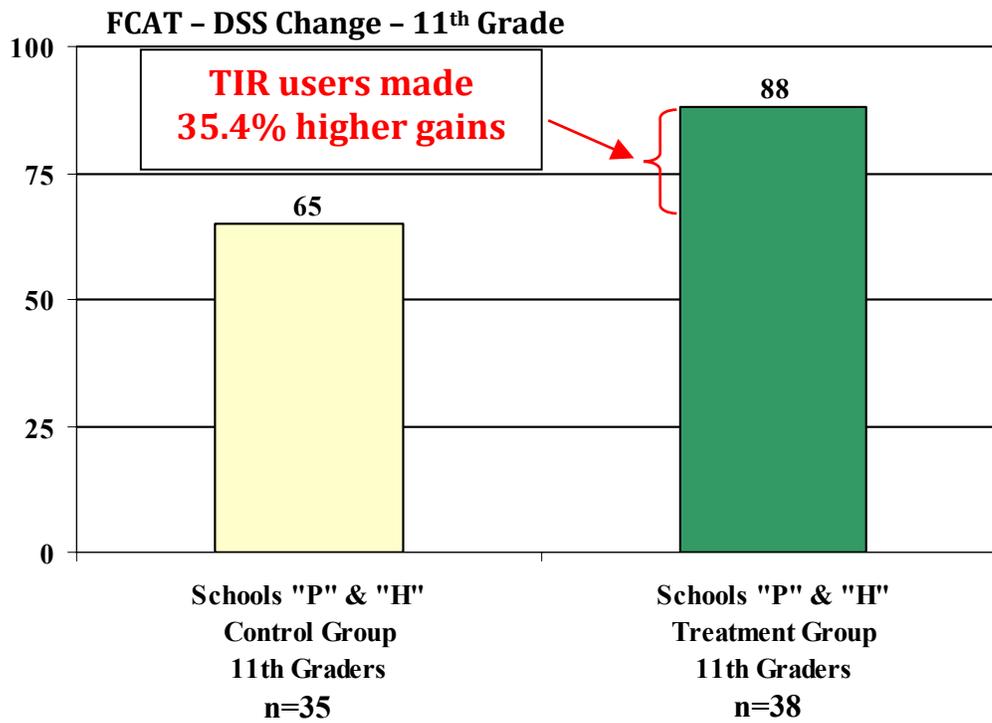
The charts in this appendix compare the DSS gains of students who used TiR to the DSS gains of control and comparison groups. The sample sizes are noted below each chart.

** Standard deviations and effect sizes available in the full report.*

Year 1 (2005 – 2006)

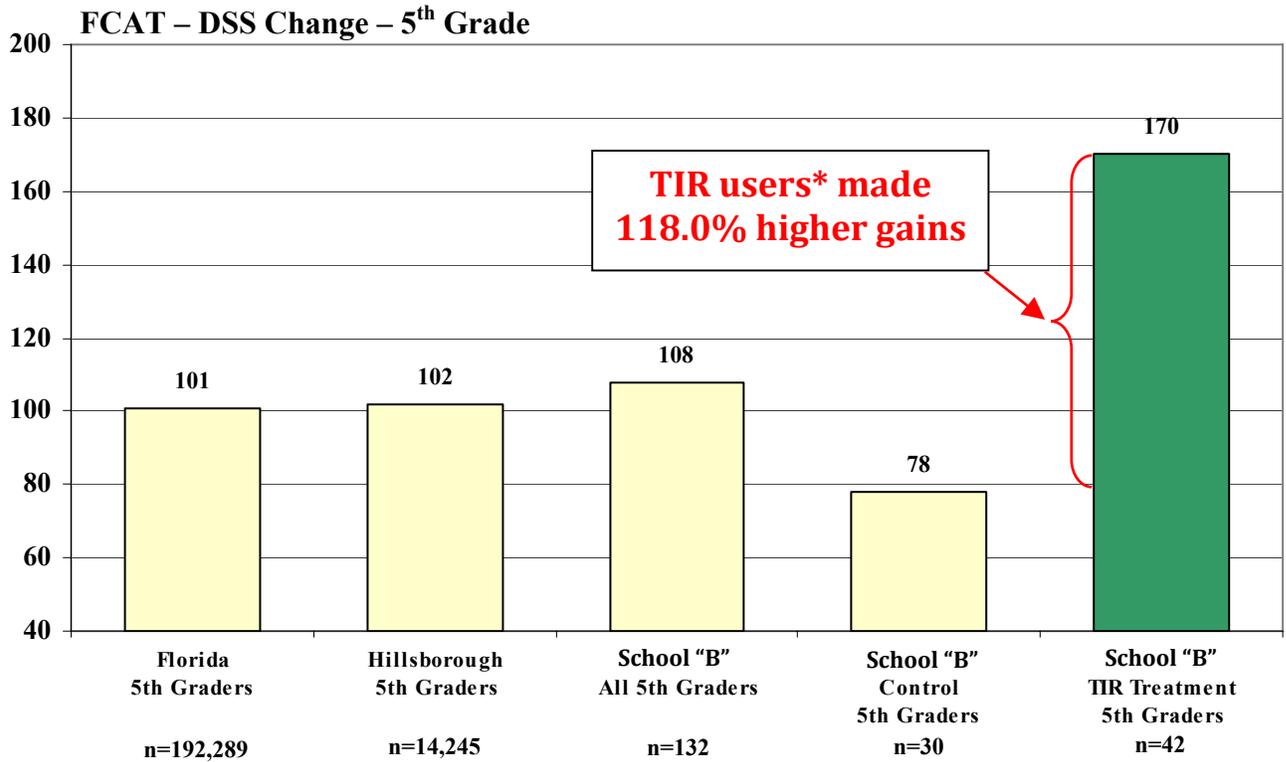


Year 1 (2005 – 2006), cont'd



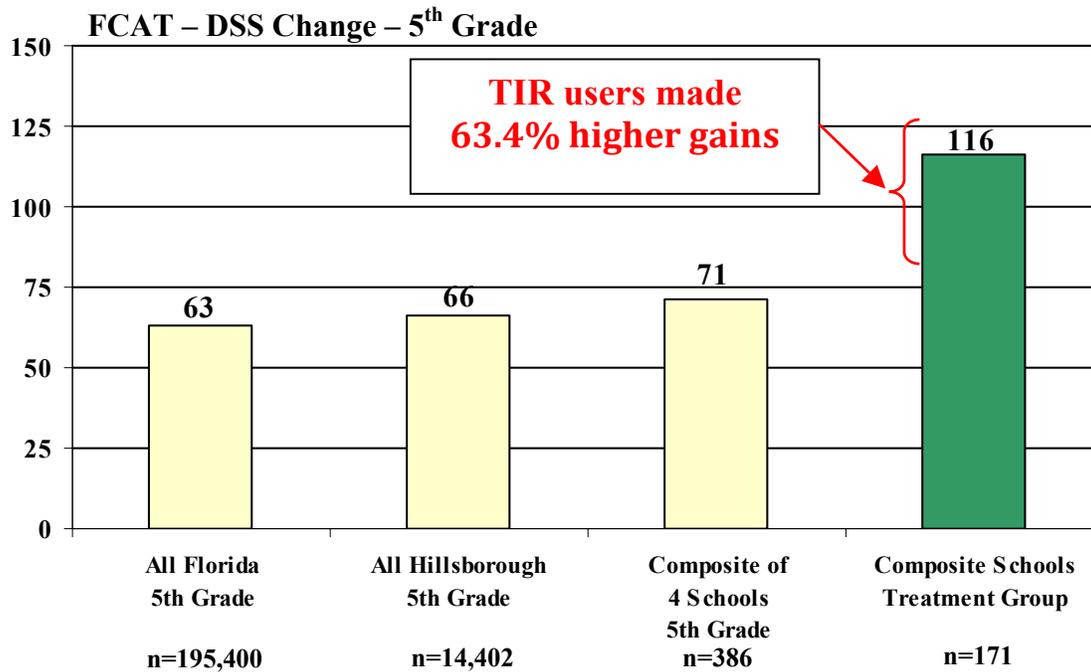
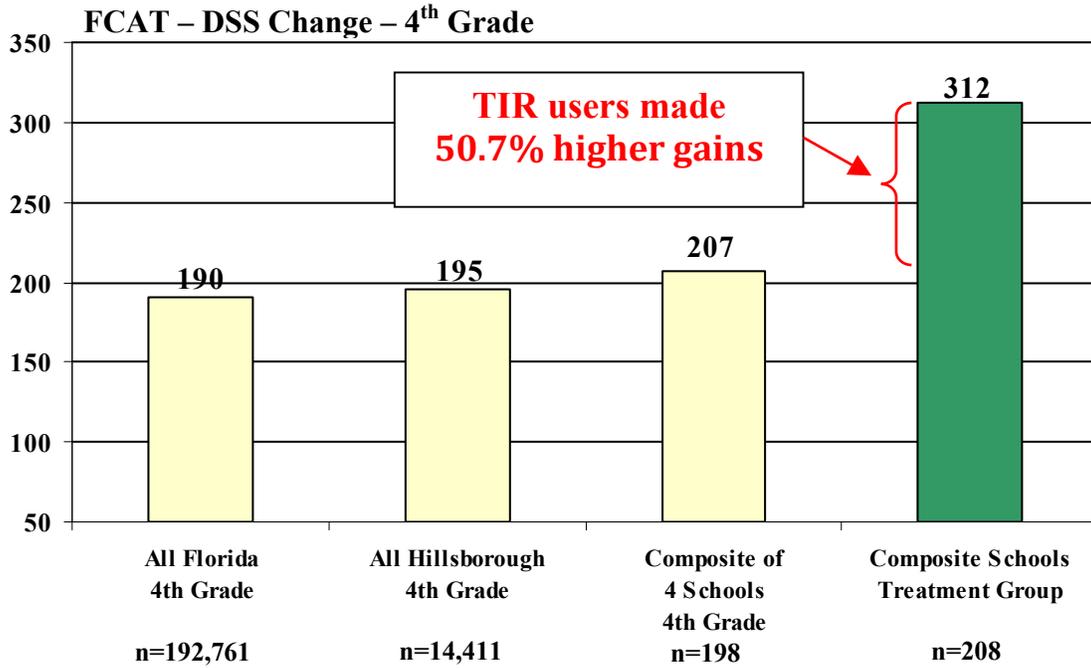
Note: No state or district 11th grade data is available at FL DOE website.

Year 2 (2006 – 2007)



* Note: These students used TIR for 18 weeks, or approximately twice the normal 9-week usage period.

Year 3 (2007 – 2008)



Appendix B – What is Tune In to Reading?

Tune in™ to Reading (TiR) is a web-based reading intervention program that has several unique features including providing real time pitch recognition and feedback to the user. The inclusion of pitch recognition is important because Lamb and Gregory (1993) found that pitch discrimination is significantly correlated with reading ability. The scoring mechanism, which is part of the TiR software, accommodates each individual's vocal range. It contains a portfolio sign-in menu that aligns to the particular vocal range of each participant. Each student uses an individual soundproof microphone headset for listening, singing, and recording. TiR contains over 600 songs analyzed for readability level. The song reading levels range from 2nd to 12th grade. Students sing songs, progressing from lower to higher readability levels, and read and reread the lyrics while attempting to improve their singing. In this way, repeated reading is integrated into the singing program.

Two different formats of textual presentations are utilized in the software program. The first format, linear sheet music, allows students to read the lyrics silently three times while listening to the background music and tempo. This practice aligns with the recommended number of repetitions suggested by Samuels (1979). The next step is a graphic textual view. This alternative text format provides a visual display of words broken into syllables without the accompanying musical staff. The software automatically places each syllable accented at the appropriate pitch within each student's personal vocal range.

Along with visual tracking of the words, the software provides a guideline for accurate pitch and tone, with a real time track line of the student's voice while they are singing and recording a song. Students can actually see what they are singing in terms of pitch. The instant feedback provides a fundamentally interactive experience that promotes visual tracking, which encourages the student to improve pitch throughout the song and during additional repeated singings (readings). After singing each time, the student receives a score. These scores, ranging from 0-100, represent the accuracy of the student's pitch and tone.

Appendix C – What is FCAT?

Florida Comprehensive Assessment Test

The primary purpose of the FCAT is to assess student achievement of the higher order thinking skills for reading, writing, math, and science. Students take the FCAT in grades 3 through 11. In grades 4, 8, and 10, students take the writing portion of the test, in grades 5, 8, and 11, students take the science portion of the FCAT, and students in grades 3 through 10 take the reading and mathematics portions. Criterion-referenced tests are designed to identify an individual's status with respect to an established standard of performance. For the FCAT, these established standards are the Sunshine State Standards. The FCAT's secondary purpose is to compare the performance of Florida students to students across the nation, which is accomplished by using a norm-referenced test (NRT) for reading and math. The current NRT used is the Stanford Achievement Test 10 (*SAT 10*), published by Harcourt Assessment, Incorporated. It is a research-based norm-referenced achievement test that provides information on student performance based on its nationwide standardization program conducted in the spring and fall of 2002 based on the K-12 population.

After a student takes the FCAT in reading and mathematics, the student receives a Developmental Scale Score (DSS) that ranges from 0 to 3000. These scores provide additional information to help interpret scores from the FCAT SSS test. Developmental scores are used because simply looking at the scale scores that the FCAT reports, which range from 100 to 500, do not reflect students' progress from previous years. Students should receive higher developmental scores as they move from grade to grade according to increased achievement. Since reading and mathematics are tested every year, this score is used to help parents and schools understand students' year-to-year progress. Based on the scale score, the student is then assigned one of five Achievement Level Classifications ranging from 1 to 5.

A level five score indicates that the student has had success with the most challenging content of the SSS and has answered most of the test questions correctly, including the most challenging questions. Students who earn a level four score have had success with challenging content of the SSS, and have answered most of the test questions correctly but may have had only some success with questions concerning the most challenging content. A level three score means that the student has had partial success with the challenging content of the SSS, but their performance is inconsistent. They may have answered many of the test questions correctly, but they are generally less successful with the most challenging questions. Students at this level are considered on grade level in reading and mathematics. A student who earns a level two score has had limited success with the challenging content of the SSS. A level one score indicates little success with the challenging content of the SSS.

Bibliography

Assessment and Accountability Briefing Book (2004).
Fcat.fldoe.org/pdf/fcataabb.pdf. FLDOE.

Biggs, Homan, Dedrick, Rasinski & Minick. (2008). Using an Interactive Singing Software Program: A Comparative Study of Middle School Struggling Readers. *Reading Psychology*. 29:195-213.

Effectiveness of Reading And Mathematics Software Products: Findings from The First Student Cohort (2007). Institute of Education Sciences, U.S. Department of Education. Report to Congress.

Effectiveness of Reading And Mathematics Software Products: Findings from Two Student Cohorts (2009). Institute of Education Sciences, U.S. Department of Education. Report to Congress.

Homan, Calderone, & Dedrick. (2009). *Using Tune Into Reading with Struggling Readers at Four Hillsborough County Elementary Schools*. Research Report: Just Read! Florida.

Lamb, S. J. & Gregory, A. H. (1993). The Relationship Between Music and Reading in Beginning Readers. *Educational Psychology*, 13(1), 19-27.

Samuels, S. J. (1979). The Method of Repeated Readings. *The Reading Teacher*, 41, 756-760.

Technical Assistance Paper (2002). *FCAT Developmental Scale*. Florida Department of Education.