TECHNICAL REPORT #9:

Silent Reading Fluency Test: Reliability, Validity, and Sensitivity to Growth for Students Who Are Deaf and Hard of Hearing at the Elementary, Middle School, and High School Levels

Susan Rose, Patricia McAnally, Lauren Barkmeier, Sean Virnig, and Jeff Long

RIPM Year 3: 2005 – 2006

Date of Study: October 2005 – May 2006

August 2008

The College of Education & Human Development

University of Minnesota

IDEAs that Work
U.S. Office of Special Education Programs

Produced by the Research Institute on Progress Monitoring (RIPM) (Grant # H324H30003) awarded to the Institute on Community Integration (UCEDD) in collaboration with the Department of Educational Psychology, College of Education and Human Development, at the University of Minnesota, by the Office of Special Education Programs. See progressmonitoring.net

Correspondence may be sent to:

Susan Rose, Ph.D.
Dept. of Ed. Psychology
University of Minnesota
56 East River Road
Minneapolis, MN 55455
612-624-6387 (V/TTY)
srose@umn.edu

Abstract

The purpose of this study was to identify the technical adequacy of progress-monitoring materials that reflect student growth in reading, specifically with students who are deaf or hard of hearing. The Silent Reading Fluency Test (SRFT) was developed in response to the need for a low-cost, reliable, and valid reading measure to inform instruction in programs for deaf and hard of hearing learners. The SRFT was designed using the format of the *Test of Silent Contextual Reading Fluency (TOSCRF)* as a means for monitoring student progress. Content for the SRFT consisted of passages from the *Reading Milestones* (RM) and the *Reading Bridge* (RB) series. In this study we investigated four different aspects of technical adequacy including interrater reliability, criterion and content validity, alternate form reliability, and differentiation of performance scores of students at the elementary, middle and high school levels when given at quarterly intervals. The study was conducted as part of a program-wide progress monitoring study. Participants in the study included $N = 101$ students who were deaf or hard of hearing in grades 3 through 12 and nine reading teachers licensed in the field of deaf education. The results of this investigation indicate that the SRFT is a relatively valid and reliable measure of reading fluency, particularly at the elementary level as indicated through statistically significant correlations with the TOSCRF, the NWEA: MAP- RIT scores, and teacher ratings. The SRFT appears to be sensitive to growth at the elementary level, however, its application to middle school and high school levels may be questionable.

The Technical Adequacy of the Silent Reading Fluency Test

The federal and state mandates for accountability and documentation of student progress have been particularly challenging for teachers of students with hearing loss due to the lack of valid, reliable, and functional assessment tools. Standardized achievement tests normed on hearing students may or may not include a sample of deaf students and thus raise concerns regarding technical adequacy and bias (Luckner, Sebald, Cooney, Young III, & Muir, 2006). Traditionally, these standardized reading tests emphasize achievement in specific areas of reading such as comprehension and vocabulary and are typically administered annually. While helpful as pre/post, or summative measures of achievement, the information gleaned from annual achievement test results have little impact on informing instruction at the time of instruction or assisting teachers in decision making in a timely manner. In addition, traditional reading achievement tests tend to be insufficiently sensitive to reflect small units of reading progress. Achievement test data suggest that the average deaf or hard of hearing student gains about 0.3 grade level per year, with less gain occurring after the student has achieved the 3rd grade reading level (Paul, 1998). With the increased emphasis on accountability in serving children with hearing loss as well as the need for verification of the effectiveness of placement, instructional interventions, and communication modalities, technically reliable measures that can signal subtle growth in reading progress are required.

Carver (1974), Jenkins and Pany (1978), and Deno (1985) pointed out that norm-referenced, standardized achievement tests do not assess students on how much of their own curriculum they have mastered. Based on criticisms of the traditional method for assessing student achievement on tests that Carver (1974) termed “psychometric” (p. 512), there has been a shift toward alternative measures that are designed to detect growth in individual students and

thus guide teaching instruction, i.e., formative evaluation (Deno, 1992). Deno, Mirkin, and their associates (Deno, Mirkin, & Chiang, 1982; Deno,1985) at the University of Minnesota Institute for Research on Learning Disabilities (IRLD) developed Curriculum-Based Measurement (CBM) procedures in the late 1970s and early 1980s (Marston, 1989). These measures are widely-used nationally and provide teachers with tools that are inexpensive, can be administered quickly and frequently, and inform teachers in making instructional decisions. Research over the past 20 years has demonstrated the technical adequacy of CBM procedures. Unlike summative assessments, CBM is an "on going measurement system designed to account for student outcomes" (Fuchs & Fuchs, 1993, p. 2). Research has also demonstrated that the data obtained from the use of CBM procedures can positively affect teaching and learning (Deno, 1992). Oral reading fluency, in which students read aloud from a passage for one minute, is the most frequently used reading measurement procedure particularly at the elementary levels (Harris & Hodges, 1995, p. 85; Pressley, 2002, pp. 292-294; Samuels, 2002, pp. 167-8; Snow, Burns, & Griffin, 1998, pp. 4 and 7). Fluency has been identified as a key indicator in the process of measuring students' reading progress.

In the National Research Council report, *Preventing Reading Difficulties in Young Children*, Snow, Burns, and Griffin (1998) stated that "adequate progress in learning to read English…beyond the initial level depends on sufficient practice in reading to achieve fluency with different texts" (p. 223). Constructing meaning from text depends strongly on the reader's skill in recognizing words accurately and reading with fluency. Because of the critical role these skills play in comprehension, both should be regularly assessed in the classroom (Snow et al., 1998).

Logan (1997) described fluency as the ability to recognize words quickly and autonomously and suggested that the most important property of fluent reading is the ability to perform two difficult tasks simultaneously. These two tasks are rapid word identification and comprehension of text. In other words, a student who is a fluent reader would have to recognize words quickly and accurately and also comprehend the text.

The most recent conceptualizations of fluency have been extended beyond word recognition processes and now include comprehension processes as well (Thurlow & van den Broek, 1997). The current concepts of what is involved in becoming a fluent reader have been enlarged to include both the word recognition/decoding process as well as the comprehension process. The National Reading Panel Report (National Institute of Child Health and Development, NICHD, 2000) recognized that effective readers are also fluent readers.

A large study on fluency achievement of students conducted by the National Assessment of Educational Progress (Pinnell, Pikulski, Wixson, Campbell, Gough, & Beatty, 1995) found a close relationship between fluency and reading comprehension. Students who are low in fluency generally have difficulty constructing meaning from the text. Given this information, it is not surprising that the National Research Council report strongly encourages frequent, regular assessment of reading fluency to permit "timely and effective instructional response when difficulty or delay is apparent" (p. 7).

CBM as well as a number of informal procedures can be used in the classroom to assess reading fluency via informal reading inventories, miscue analysis, and reading speed calculations to name a few. All of these procedures require students to read orally and therefore, are not appropriate for use with a majority of deaf and hard of hearing learners. Hammill, Wiederholt, and Allen (2006) developed a tool for monitoring the progress of students' reading fluency

which does not require oral reading. This test presents a promising format for use with deaf and hard of hearing students. The test they developed, the *Test of Silent Contextual Reading Fluency* (TOSCRF), measures the speed with which students can recognize individual words in a series of printed passages that become progressively more difficult in their content, vocabulary, and grammar (Hammill et al., 2006).

In the *Examiner's Manual,* the TOSCRF test developers (Hammill, et. al.2006) listed a wide variety of interrelated silent reading skills measured by the test including:

- Recognize printed words and know their meaning
- Use one's mastery of syntax and morphology (i.e., grammar) to facilitate understanding the meaning of written sentences and passages
- Incorporate word knowledge and grammar knowledge to quickly grasp the meaning of words, sentences, paragraphs, stories, and other contextual materials
- Read and understand contextual material at a pace fast enough to make silent reading practical and enjoyable (p.2).

The authors maintain that its "results…can be used confidently to identify both poor and good readers" (Hammill et al., 2006, p. 2). They further stated that the TOSCRF can be used to measure contextual fluency in a comprehensive reading assessment, to monitor reading development, and to serve as a research tool.

The format of the TOSCRF requires students to identify words printed without spaces between them by drawing lines indicating word boundaries, a format that is similar to the format of children's word-search puzzles. Guilford (1959) used several word search measures in

developing his Structure of Intellect model and Guilford and Hoepfner (1971) used word search measures to help establish the Convergent Production of Symbolic Transformations factor.

Meeker and Meeker (1975) developed the *Structure of Intellect Learning Abilities Test*, a battery of tests to evaluate a variety of cognitive abilities. The second edition of this test (Meeker, Meeker, & Roid, 1985) used the contextual word-strings-without-spaces format in three of the four sets that measure word recognition.

In summary, the works of Guilford, Hoepfner, the Meekers, and Roid provide a rationale for the TOSCRF format that is based on theory and is validated thoroughly in the research literature. These early works provide the TOSCRF with a firm foundation that helps establish its content-validity (Hammill et al., 2006, pp. 36-37).

Because there are few good assessment tools of student achievement in general, and reading in particular, that are available for deaf and hard of hearing learners (Paul, 2001), the format used in the TOSCRF appeared to be a promising format for use with this population. The tool we developed for monitoring reading progress of deaf and hard of hearing students, the *Silent Reading Fluency Test* (SRFT), is based on Deno's (1985) conceptualization of CBM procedures for progress monitoring and on the format of the TOSCRF.

Using the format of the *Test of Silent Contextual Reading Fluency* (Hammill, Wiederholt, & Allen, 2006) as a model, similar materials were developed using passages from the text in the *Reading Milestones* (Quigley, McAnally, Rose, & King, 2001) and *Reading Bridge* (Quigley, McAnally, Rose, and Payne, 2003) series. These two reading series, which contain controlled vocabulary and linguistic structures, were developed for use with deaf and hard of hearing students, English language learners, and students with language disabilities. In addition, these reading series were selected because, in a national survey on reading materials used in programs

for deaf and hard of hearing learners, *Reading Milestones* was reported as being used by 30% of the programs which was three times the number of programs that used the second most frequently cited basal reader (LaSasso & Mobley, 1997). A progress monitoring tool based on *Reading Milestones* and *Reading Bridge* may provide a progress monitoring system for programs that use these two reading series and an effective measure for other programs as well.

Purpose

The purpose of this study was to identify the technical adequacy of the SRFT. That is, the content and criterion validity, alternate form reliability, and inter-rater reliability of the SRFT as well as to identify its ability to discriminate higher level readers from less skilled readers and to determine its ability to reflect student progress in reading when given at frequent intervals. The SRFT was developed specifically for use with students who demonstrate significant language differences including students with hearing loss and English Language Learners. The SRFT followed the design of the TOSCRF (Hammill, et. al., 2006) and the characteristics of CBM, that is, the process and design of the measures can be administered quickly and at frequent, systematic intervals (Deno, 1985). The quantitative results of the SRFT can often be used as an indicator of student progress in reading, and can provide teachers, parents, and students with information needed to determine the effectiveness of instruction.

Research Questions

The following research questions were addressed in this study: (1) Are the SRFT scores relatively valid and reliable indicators of the general reading performance of deaf and hard of hearing learners? (2) Are Forms A and B of the Silent Reading Fluency Test (SRFT) equivalent? (3) Do the SRFT scores differentiate performance of students at various levels of reading? (4) Is

the SRFT sensitive to reading growth of deaf and hard of hearing students when given at quarterly intervals?

It was anticipated that the SRFT would prove to be a relatively valid and reliable indicator of students' reading performance at the elementary level. We also anticipated that the SRFT would allow frequent administration (e.g., weekly, monthly or quarterly) for the purpose of informing teachers of students' progress in reading.

Method

Participants and Setting

This study was conducted as part of a school-wide progress monitoring program. A sample of $N = 101$ students, 3rd through 12th grade, and nine teachers participated in the program. All of the students qualified for special education services due to identified hearing loss and educational need. Thirty–six students had losses in the mild to moderate range and 61 students had losses ranging from severe to profound. Four students did not have recorded information regarding the nature and severity of their hearing losses. Forty-six of the 101 participants were female; 55 were male, and 23% ($n = 23$) had additional diagnosed disabilities (e.g., ADD, emotional/behavioral disability, mild cognitive delay). All of the students were enrolled in the academic program. Twenty-seven participants were elementary students, 16 were middle school students, and 58 were high school students. Participants included 46% residential students (students living at the school during the week), and 54% day students (students living with their families in the community and attending the school daily during regular school hours). Six of the high school students attended the community public schools for a part of the school day.

Measures

Silent Reading Fluency Test development. The SRFT was developed by the first two authors. The format and protocol followed as closely as possible those used to develop the *Test of Silent Contextual Reading Fluency* as explained in the test manual (Hammill, Weiderholt, and Allen, 2006). The passages in *The Silent Reading Fluency Test* were adapted from the text in the *Reading Milestones* (RM) series, which has six levels, ten books per level, and *Reading Bridge* (RB) series, which has two levels, two books per level. The beginning passages incorporate preprimary/first-grade vocabulary and simple grammatical structures with graduated complexity in subsequent passages.

Reading levels in the *Reading Milestones* series range from preprimary/first through third grades and in the *Reading Bridge* series, from fourth through fifth grades. The language structures in the *Reading Milestones* and *Reading Bridge* series use vocabulary and linguistic structures designed to meet the needs of deaf or hard of hearing students at each level as indicated by the results of the national research investigations conducted by Stephen P. Quigley and colleagues (1976). In the *Reading Milestones* series, the last story in each book incorporates the targeted vocabulary words and language structures used in the first five stories of the book. Most of the passages in the test protocol that the students were given to read were taken from story 6 in the last book of each level of the *Reading Milestones* series. In the *Reading Bridge* series, passages were taken from selections near the end of each of the four books. Two forms were constructed for the SRFT (see Form A and Form B in Appendix A). Each of the passages in the two forms was similar in number of words to the passages in the TOSCRF. The vocabulary and the grammatical structures reflected the controlled language used in *Reading Milestones* and *Reading Bridge.* In addition to the passages from the RM and the RB series, Form B from the

Test of Silent Contextual Reading Fluency (TOSCRF) was added to the testing materials as a criterion reference. Although the items on the TOSCRF extend from preprimer/first grade through 12th grade levels of difficulty, the items on the SRFT ranged from preprimer/first grade through 5th grade levels. The passages from RM and RB were printed in the same format as those in the TOSCRF. They were printed in upper case without punctuation or spaces between the words. For example:

A K I N G L I V E D I N A G O L D C A S T L E T H E C A S T L E

W A S B Y A R I V E R N E A R T H E M O U N T A I N S

The task required students to read each passage and draw lines between the boundaries of words (e.g., A/KING/LIVED/IN/A/GOLD/CASTLE/THE/CASTLE/WAS BY/A/RIVER/ NEAR/THE/MOUNTAINS).

Criterion and Content Validity Measures

Anastasi and Urbina (1997) suggested that "…criterion-prediction (also referred to as *criterion-related validity)* validation procedures indicate the effectiveness of a test in predicting an individual's performance in specific activities" (p. 118). Based on this description, the criterion-prediction validity for the SRFT refers to the ability of this assessment to measure an individual's performance in reading, specifically in word identification, comprehension and fluency. In discussing content-description validity, Anastasi and Urbina (1997) suggested that these "validation procedures involve the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured" (pp. 114-115).

The SRFT was correlated to the TOSCRF to establish criterion validity. The TOSCRF has been correlated to five other tests of reading that measured word identification, reading fluency, comprehension, and general reading ability resulting in coefficients ranging from .57 to .89 (Hammill, Wiederholt, and Allen (2006). According to Hopkins (2002), the magnitude of these correlations would be LARGE (.50-.69) to VERY LARGE (.70-.99). The correlations between the TOSCRF and the SRFT ranged from .84 to .90. Hammill and colleagues (2006, p. 38) included 46 students who were deaf or hard of hearing in the validity studies using the Test of Silent Word Reading Fluency (TOSWRF; Mather, Hammill, Allen & Roberts, 2004) as the criterion reading measure.

The TOSCRF (Form B) was also used as one of the content validity measures. In designing this assessment tool, Hammill and colleagues selected sentences from the *Gray Oral Reading Tests—Fourth Edition* (GORT-4; Wiederholt, & Bryant, 2001). Hammill and colleagues (2006) stated that "the works of Guilford, Hoepfner, M. Meeker, R. Meeker, and Roid provide a rationale for the TOSCRF that is based on theory and is validated thoroughly in the research literature" (p. 37). The works also indicated that the format used in the TOSCRF (word strings without spaces) has been successfully used to measure word identification speed "providing the TOSCRF with a firm foundation which helps establish it's content-description validity" (p. 37). In summary, the test developers stated that considerable evidence has been collected and presented to "show that the TOSCRF is a valid measure of general reading and reading fluency and can be used with confidence" (p. 49).

The *Measures of Academic Progress* (NWEA: MAP), an assessment system developed by the Northwest Evaluation Association (Northwest Evaluation Association [NWEA], 2003) was used as a second source for the confirmation of content validity. There is substantial

evidence supporting the validity of the NWEA: MAP for measuring the achievement level and growth of students in the major subject areas (e.g., Reading, Language, Math, and Science; NWEA, 2003). The concurrent validity statistics for the MAP show that it was compared to seven other tests that assess student learning. The correlation coefficients on the seven tests ranged from .75 to .87. When the reading subtest of the MAP was compared to the two tests (the Stanford Achievement Test, 9th Edition, and the Colorado State Assessment Program) taken by the same students, the coefficients ranged from .82 to .87 and .84 to .87 respectively. The range of the correlation coefficients for the Iowa Tests of Basic Skills was .77 to .84.

The reliability of the MAP was calculated in two different manners. One method used marginal reliability that utilizes "the test information function to determine the expected correlation between the scores of two hypothetical tests taken by the same students" (NWEA, 2003, p.54). The marginal reliability estimates ranged from .90 to .95 across second through tenth grade levels. Test-Retest reliability correlations ranged from .76 to .91 across second through tenth grade levels. The NWEA maintains that the content is valid; the scores are constructed in a valid manner, and that the Association "has a great deal of confidence in the validity of the assessments for their intended uses" (NWEA, 2003, p. 52).

A third content validity measure used was teacher ratings of each student's reading, writing, math, and general communication abilities. Teachers were provided with a class list and a rating scale of 1 to 5 with 1 indicating a low general reading ability and 5 an above average general reading ability in comparison to their peers. Teachers submitted their ratings for each student during the Fall semester.

Procedure

Data collection. Administration of the SRFT and TOSCRF were part of a school-wide progress monitoring program that included administration of the MAP, CBM MAZE passages, the SRFT, TOSCRF and teacher ratings of students' performances. A school-wide inservice training in the purpose and principles of progress monitoring was provided followed by a brief training session in the administration of the SRFT and TOSCRF. Classroom teachers of the deaf administered the progress monitoring measures using directions and practice items provided with the progress monitoring packets. Each student completed Forms A and B of the SRFT and Form B of the TOSCRF within a three day time span during each quarter of the academic year. The order of administration of the progress monitoring measures was counterbalanced, that is one third of the students completed the SRFT-Form A followed by Form B followed by the TOSCRF-Form B; one third of the students completed the SRFT-Form B, TOSCRF-Form B and SRFT-Form A; the remaining students completed TOSCRF-Form B, SRFT-Form A followed by Form B. Using the same standard instructions as used by the TOSCRF, the teachers presented practice forms to familiarize the students with the test format and response requirements before administering the test materials. The teachers used stop watches and flashing lights to indicate the end of the testing period.

The SRFT (Forms A and B) and the TOSCRF (Form B) were given to students as a 3-minute timed task. The teachers administered each form to the class of students in less than ten minutes, including the time necessary for explaining the directions, completing the practice items, and completing the task.

Scoring. Scoring the SRFT and the TOSCRF consists of counting the words the student identified correctly through all passages attempted. The numerical scores for each form reflect

total words correctly identified in all passages attempted by the student, incorrect words identified, and the correct words minus the incorrect words. For the purposes of this study, only the number of words correctly identified was used which conforms to the scoring practices of the TOSCRF.

The NWEA: MAP was administered by the teachers in the Fall and Spring of the Academic year within ten days of administration of the progress monitoring measures.

Data analysis. Assessment of cross-sectional reliability and validity was based on Pearson correlation coefficients using the CORR procedure of the SAS software. Sensitivity to growth was assessed using growth curves based on hierarchical linear modeling (HLM; Singer & Willett, 2003, chap. 3). The SAS procedure MIXED was used for the HLM analysis.

Results

Interscorer Agreement

Two raters independently scored the SRFT and the TOSCRF. The two sets of data were analyzed to establish inter-rater reliability. The average agreement between the scores of the two raters across all measures (SRFT-Forms A and B and the TOSCRF-Form B) was 92%. A review of the discrepancies in scoring indicated that most of the discrepancies occurred when the line or slash between two words was questionable, that is, ambiguous slash marks over letters or extending across word boundaries. Discrepancies between the two raters were resolved through discrepancy scoring by a third independent rater.

Alternate Form Reliability

Alternate form reliability was calculated between mean scores for the SRFT-Forms A and B by computing Pearson-Product Moment correlations. The correlation coefficient (.92) between Forms A and B indicates that the forms have a high level of agreement.

Descriptive Statistics for the SRFT and the TOSCRF

Descriptive statistics of the assessment results are presented in the appendix. The means of correct- word scores for elementary, middle school, and high school students for the Fall, Winter, and Spring administrations of the SRFT-Forms A and B and TOSCRF-Form B are presented in Table 1 and Figures 1 and 2. The information displayed in the table and graphs indicates that the test differentiates among student scores at these three levels, that is, students in upper grades have a mean higher score on the SRFT than students in lower grade levels.

Insert Table 1 and Figures 1 & 2 about here

Content/ Construct Validity

To establish content and construct validity, a triangulated set of variables was used. The SRFT was correlated with the TOSCRF, with the reading subtest of the NWEA: MAP (NWEA, 2003) and with teacher ratings of students' reading abilities. The SRFT-Forms A and B were compared individually with the TOSCRF resulting in a correlation coefficient of .84 between Form A and the TOSCRF and a coefficient of .90 between SRFT-B and TOSCRF.

The correlations between the reading scores of the SRFT and the reading subtests RIT scores of the NWEA: MAP were computed for each level (elementary, middle school, and high school) on SRFT-Forms A and B on the fall and spring data. The correlations range from .58 to .85 for the fall data and from .45 to .84 for the spring data (Table 2).

Insert Table 2 about here

The results of an application of a Pearson-Product Moment correlation comparing teacher ratings of each student's general reading ability and student scores on SRFT-Forms A and B across grade levels are presented in Table 3. Four weeks into the Fall term, teachers rated each student on a scale of 1 to 5 with 1 indicating low general reading ability and 5 indicating above average general reading ability. The results showed satisfactory correlations between teacher ratings and student scores at the elementary level on SRFT Forms A and B (Fall = .85 and .74; Spring = .81 and .84), but low correlations at the middle school level (Fall = .62 and .69; Spring = .64 and .60) and high school level (Fall = .60 and .58; Spring = .45 and .54).

Insert Table 3 about here

Sensitivity to Growth

Sensitivity to growth was assessed by estimating the linear slope of the measures over three assessment period (Fall, Winter, and Spring). Table 4 shows the estimated intercept and slope of the measures by grade level (elementary, middle, and secondary). All the intercept estimates were statistically significant at the $\alpha = .001$ level. Therefore, Table 4 shows only the *p*-values for the test of the slope. The slopes indicate the average increase per wave of measurement. For example, the predicted increase in elementary level SRFT-A was 12.68 per assessment wave. This translated into an increase of 2(12.86) = 25.72 over all three waves (Fall to Spring). As the last column of the table indicates, the slopes were significant for the elementary and secondary levels but not for the middle grades. That is, the positive increases in the middle grades indicated by the slope values were not statistically reliable. On the other hand, the gains for the other grade levels were statistically reliable.

Insert Table 4 about here

Discussion

The purpose of this study was to examine (1) if Forms A and B of the Silent Reading Fluency Test (SRFT) are approximately equivalent forms, (2) the criterion and content validity of the SRFT, (3) if the SRFT and the TOSCRF show a differentiation in the performance scores of students at the elementary, middle school, and high school levels, and (4) if the SRFT is sensitive to reading growth of deaf and hard of hearing students when given at frequent intervals. To establish the alternative form reliability of the SRFT, we examined the two forms of the test to determine equivalency. The correlation of .92 between Forms A and B indicates that the forms are satisfactorily equivalent.

We also examined the scores of students at elementary, middle, and high school levels. The data indicate that the SRFT differentiates among student scores at these three academic levels. That is, the mean scores on the SRFT of students in upper grades were higher than the mean SRFT scores of students in lower grades (see Table 1 and Figures 1 and 2).

The means, standard deviations, and the slopes of the trend lines derived from students' scores at the three time points were examined to determine if the SRFT is sensitive to a change in reading performance of deaf and hard of hearing learners (see Table 1 and Figures 1 and 2). Student scores at all three levels show increases from the fall to winter testing. However, the slopes of the middle grades were not significant. Details of the HLM analysis (not presented) indicated a much greater variability for the middle school scores leading to larger standard errors for the slopes and thus a lack of statistical significance. As the middle score slopes were similar

or even greater in magnitude than the other levels, the measures may be found to have statistically significant slopes in samples with less variability than ours. To summarize the SRFT and TOSCRF appear to measure a change in reading performances among deaf and hard of hearing learners at quarterly intervals, at least at the elementary and secondary levels.

To examine content validity, student SRFT scores were compared to the NWEA: MAP scores completed during the same two week period in the fall and spring. The correlations between the scores of the SRFT and the reading subtests of the NWEA: MAP-RIT scores show relatively low correlations (range from .45-.85, see Table 2) between the two tests, particularly at the middle and high school levels. Assuming that the NWEA: MAP is a valid test of reading performance for deaf and hard of hearing learners, these correlations may indicate that the two tests do not measure the same underlying construct of reading skills or that the SRFT and/or the NWEA: MAP is not an accurate measure of reading performance of deaf and hard of hearing students.

The SRFT scores were compared to TOSCRF raw scores to establish content validity. The correlations between the SRFT, Forms A and B, and the TOSCRF were relatively high (.84 and .90) indicating that these two tests probably do measure the same underlying construct of reading skills.

Correlations were calculated between teacher ratings of general reading ability and the student scores on the SRFT. The results show satisfactory correlations between teacher ratings and student scores at the elementary level but low correlations at the middle school and high school levels. Teacher ratings may have been influenced by the fact that during the first quarter of the school year, students were reassigned to different classes for reading. Other factors that may have influenced the lower correlations at the middle school and high school levels is that

these two academic levels have rotating classes and teachers are not as familiar with their students as teachers at the elementary level who work with the same group of students all day. Another factor to consider is that middle school and high school studies focus on content areas. Since reading is not a content area, teachers may be less aware of the students' reading skills because less time is spent on direct instruction in reading.

Conclusions

Content validity of the SRFT appears to be acceptable when used with elementary level students with hearing loss. However, the application of the SRFT with middle and high school level students is questionable. The results of the data analysis produced two positive indicators, that is, (1) the SRFT appears to be sensitive to reading growth at the elementary and secondary levels, and (2) the mean SRFT score for students in upper grade levels was higher than students in lower grade levels. However, the low correlations between NWEA: MAP-RIT scores and teacher ratings at the middle and high school levels preclude a more optimistic conclusion, as does the lack of significant slopes for middle school. Given the construction of the SRFT, that is, using text written at the 1st through 5th grade levels, the outcomes may indicate that the range of reading difficulty is too narrow to show reading growth in students reading beyond the 5th grade level.

References

Anastasi, A., & Urbina, S (1997). *Psychological testing* (7th ed.). New York: Macmillan.

Carver, R.P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist, 29,* 512-518.

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(3), 219-232.

Deno, S.L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36,* 5-10.

Deno, S.L., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 18,* 16-26.

Fuchs, L.S. & Fuchs, D. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review,* 02796015, vol. 22, 1, 27-49.

Guilford, J.P. (1959). Three faces of intellect. *American Psychologist, 14,* 469-470.

Guilford, J.P., & Hoepfner, R. (1971). *The analysis of intelligence.* New York: McGraw-Hill.

Jenkins, J., & Pany, D. (1978). Standardized achievement tests: How useful for special education? *Exceptional Children, 44,* 448-453.

Hammill, D.D., Wiederholt, J.L., & Allen, E.A. (2006). *TOSCRF: Test of Silent Contextual Reading Fluency, examiner's manual.* Austin, TX: PRO-ED.

Harris, T.L., & Hodges, R.E. (1995). *The Literacy Dictionary: The vocabulary of reading and writing.* Newark, DE: International Reading Association.

Hopkins, W.G. (2002, June 23). A scale of magnitudes for the effect statistics. In *A New View of Statistics.* Retrieved from http://www.sportsci.org/resource/stats/effectmag.html.

LaSasso, C., & Mobley, R. (1997). *Results of a national survey of reading instruction for deaf students.* Monograph of Collected Papers from the 23rd Annual Conference, The Association of College Educators—Deaf and Hard of Hearing, Santa Fe, NM.

Lennon, R.T. (1956). Assumptions underlying the use of content validity. *Educational and Psychological Measurement, 16,* 294-304.

Logan, G.D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly, 13,* 123-146.

Luckner, J.L., Sebald, A.M., Cooney, J., Young III, J.Y., & Muir, S.G. (2006). An examination of the evidence-based literacy research in deaf education. *American Annals of the Deaf, 150*(5), 443-455.

Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M.R. Shinn (ed.), *Curriculum-Based Measurement: Assess Special Children.* New York: The Guilford Press, 18-78.

Mather, N. Hammil, D.D., Allen, E.A. & Roberts, R. (2004). *Test of Silent Word Reading Fluency.* Auston, TX: PRO-ED.

Meeker, M., & Meeker, R. (1975). *Structure of intellect learning abilities test.* El Segundo, CA: SOI Institute.

Meeker, M., Meeker, R., & Roid, G.H. (1985). *Structure of intellect learning abilities test—1985 edition.* Los Angeles: Western Psychological Services.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103. New York: Macmillan.

Messick, S. (1995, September). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

National Institute of Child Health and Human Development (NICHD). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Northwest Evaluation Association. (2003). *Technical manual for the NWEA Measures of Academic Progress and Achievement Level Tests.* Portland, OR: Northwest Evaluation Association.

Paul, P.V. (2001). *Language and deafness.* San Diego, CA: Singular.

Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., & Beatty, A.S. (1995). *Listening to children read aloud.* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Pressley, M. (2002). Metacognition and self-regulated comprehension (pp.291-309). In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed.), Newark, DE: International Reading Association.

Quigley, S.P., McAnally, P., Rose, S., & King, C. (2001) *Reading milestones*. Austin, TX: PRO-ED.

Quigley, S.P., McAnally, P., Rose, S., & Payne, J. (2003). *Reading bridge.* Austin, TX: PRO-ED.

Quigley, S.P., Wilbur, R., Montanelli, D., & Steinkamp, M. (1976). *Syntactic structures in the language of deaf children.* Urbana-Champaign: University of Illinois, Institute for Child Behavior and Development.

Samuels, S.J. (2002). Reading fluency: Its development and assessment (pp. 166-183). In A.E. Farstrup & S.J. Samuels (eds.), *What Research Has to Say About Reading Instruction* (3rd ed.), Newark, DE: International Reading Association.

Singer, J. D., & Willett, J. B., (2003). *Applied Longitudinal Data Analysis*. New York: Oxford.

Snow, C.E., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Thurlow, R., & van den Broek, P. (1997). Automaticity and inference generation. *Reading and Writing Research Quarterly, 13,* 165-184.

Appendix

Table 1

Means and Standard Deviations of Silent Reading Fluency Scores/TOSCRF Scores

	Fall	*Winter*	*Spring*
SRFT-A			
Elementary	40.88	58.91	52.32
	(20.31)	(26.85)	(17.29)
Middle School	81.27	92.79	113.30
	(28.37)	(29.45)	(36.36)
High School	96.59	109.84	118.95
	(36.06)	(38.69)	(42.91)
SRFT-B			
Elementary	55.15	71.91	68.84
	(25.17)	(32.36)	(26.53)
Middle School	94.67	116.64	126.62
	(32.13)	(28.88)	(36.05)
High School	112.13	123.23	130.36
	(33.09)	(32.91)	(41.74)

Table 2

Correlations of SRFT and NWEA: MAP

	Elementary	*Middle*	*HS*
N	25	15	44
SRFT-A	.75	.62	.59
SRFT-B	.74	.69	.58

Table 3

Correlation of SRFT to Teacher Ratings – Fall

	Form A	Form B
Elementary	0.85	0.82
Middle	0.51	0.58
HS	-0.34	-0.47

Table 4

Silent Reading Fluency Linear Trends Across Elementary, Middle, and High School Levels

Levels	Intercept	Slope	Slope *p*-value
Elementary			
SRFT-A	42.02	12.68	< 0.001
SRFT-B	56.46	12.43	< 0.001
TOSCRF	47.00	12.09	< 0.001
Middle			
SRFT-A	82.42	8.86	0.32
SRFT-B	95.50	17.18	0.07
TOSCRF	93.55	7.35	0.12
Secondary			
SRFT-A	90.97	20.13	< 0.001
SRFT-B	111.95	11.38	< 0.001
TOSCRF	106.67	7.52	< 0.001

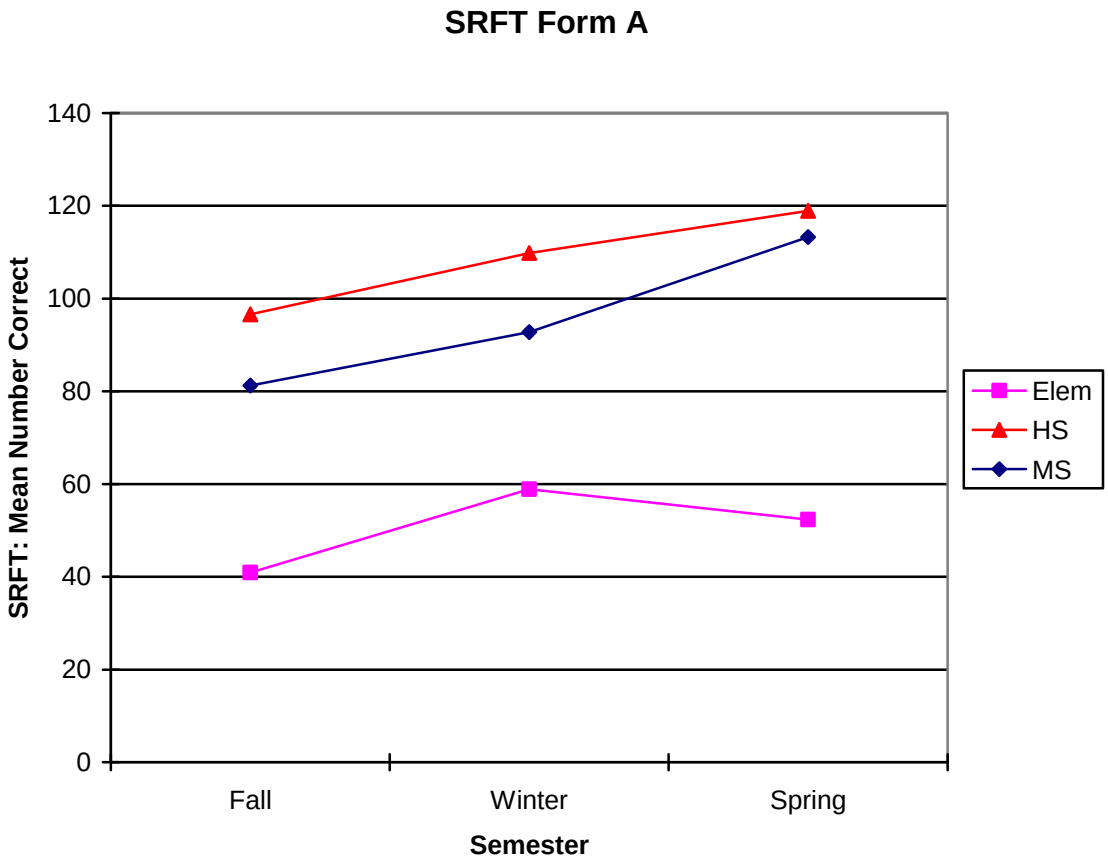

Figure 1. SRFT- Form A: Mean number correct for Fall, Winter and Spring semesters at the Elementary, Middle School and Secondary Levels.

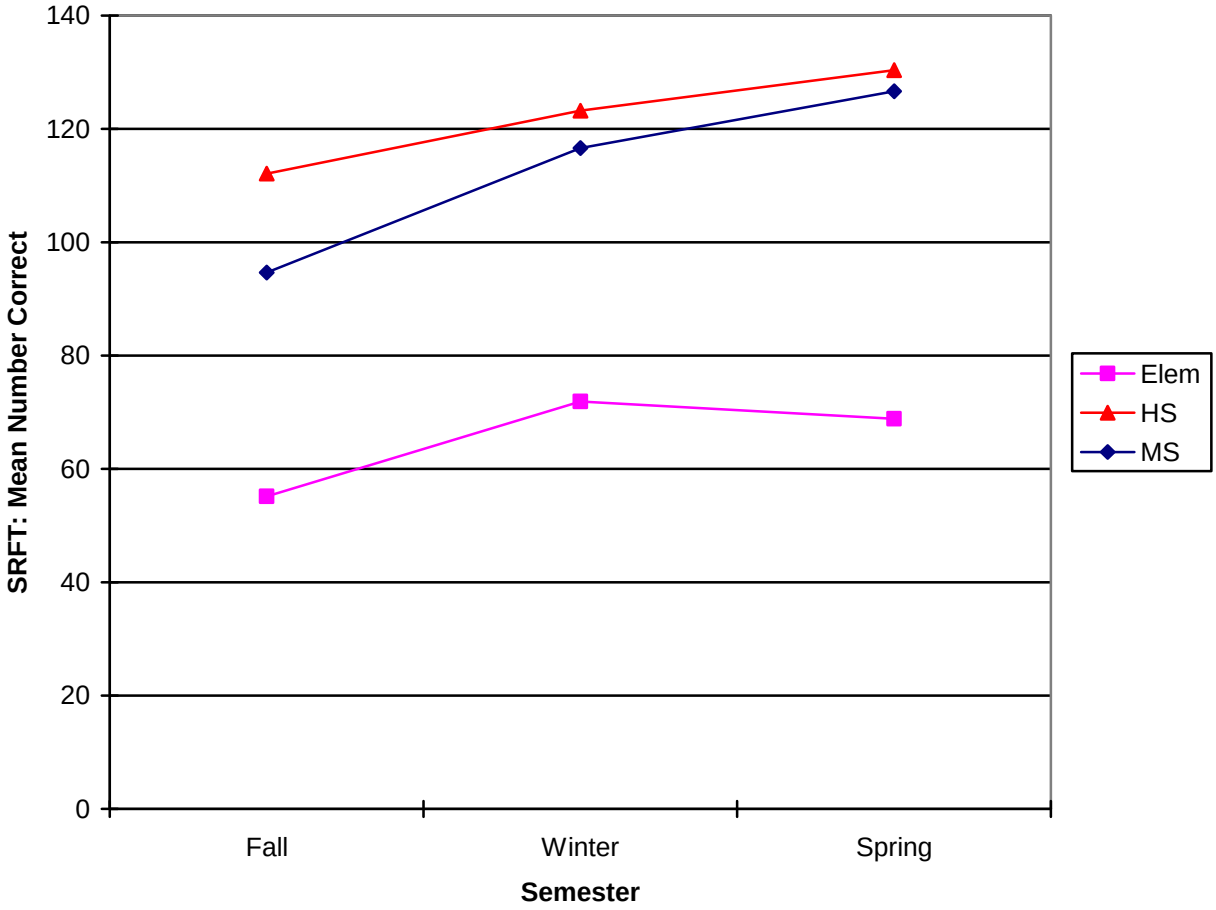

Figure 2. SRFT- Form B: Mean number correct for Fall, Winter and Spring semesters at the Elementary, Middle School and Secondary Levels.