



TECHNICAL REPORT #11:

Assessing Written Expression for Students Who Are Deaf or Hard of Hearing: Curriculum-Based Measurement

Shu-fen Cheng and Susan Rose

RIPM Year 3: 2005 – 2006

Date of Study: Summer 2005

August 2008

The College of Education
& Human Development

UNIVERSITY OF MINNESOTA



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

Abstract

The purpose of this study was to address the technical adequacy of progress monitoring scoring procedures of written expression for use with students with hearing loss. Twenty-two secondary-school deaf and hard of hearing students completed two story starter and two picture prompt written expression probes. Each writing sample was scored using seventeen different written expression scoring procedures. Technical characteristics for each of scoring procedures were examined in terms of alternate-form reliability and criterion-related validity.

Results indicate that the most reliable and valid predictors of general writing proficiency for students with hearing loss were correct word sequences and correct minus incorrect word sequences. The findings from this study support the use of the CBM scoring procedures as reliable and valid tools for assessing written expression for secondary -level students with hearing loss.

Assessing Written Expression for Students Who Are Deaf Or Hard of Hearing:
Curriculum-based Measurement

Individuals who are deaf or hard of hearing consider writing skills as a a a critical component in job-related communication and community participation (Costello, 1977); however, proficiency in writing has been regarded as an incredible challenge by students who are deaf or hard of hearing. The literature reveals that the majority of students who are deaf or hard of hearing fail to master the complex writing process of English. Research-based evidence shows that students who are deaf or hard of hearing, when compared with their hearing peers, generate fewer words, produce shorter and simpler sentences, show less complex grammar and descriptive words, use the same words repeatedly make significantly more mechanical errors, and have difficulty using function words and creating coherent texts (Ivimey & Lachterman, 1980; Marschark, Lang, & Albertini, 2002; Marschark, Mouradian, & Halas, 1994; Singleton, Morgan, DiGello, Wiles, & Rivers, 2004; Yoshinaga-Itano & Snyder, 1985; Yoshinaga-Itano, Snyder, & Mayberry, 1996a; Yoshinaga-Itano, Snyder, & Mayberry, 1996b). Researchers and educators have focused on improving the writing proficiency of students with hearing loss for more than a century. While some gains have been made, the written expression of students with hearing loss continues to be a challenge (Albertini & Schley, 2003; Paul, 2001; Rose, McAnally, & Quigley, 2004).

As a result of the Individuals with Disabilities Education Improvement Act (IDEA) and the No Child Left Behind Act (2001), programs serving children with disabilities are required to provide accountability for students' academic progress. The language in IDEA 2004 states that "All children with disabilities are included in all general state and district wide assessment programs . . . with appropriate accommodations and alternate assessments, where necessary and as indicated in their respective individualized education programs." (Section 1412(c)(16)(A)). This mandate places unique pressure on programs serving students

who are deaf and hard of hearing. The challenge is in determining assessment procedures that can inform and guide instruction and that are technically robust, that is, reliable and valid, as well as efficient and practical.

Curriculum-Based Measurement (CBM) is a formative assessment process designed to monitor student progress in basic skill areas including reading, spelling, mathematics, and written expression (Deno, 1985). CBM was originally designed to provide special education teachers with an efficient and effective process to formatively evaluate the effects of their instruction on students' academic performance. Thus, curriculum based measures were designed to be general outcome indicators that are sensitive to students' growth over a short period of time (e.g. weekly, monthly). In order for these measures to be functional, Deno, Marston & Mirkin (1982) set standards that include strong technical adequacy, and procedures that were easy to use, required minimal training, easy to interpret, informative, time efficient and economical. An extensive history of research demonstrates that CBM, in the area of reading, written expression, and mathematics is a reliable and valid as a formative assessment process and can be used effectively to collect student performance frequently to support educational decisions (Deno, 1985; Deno, 2003; Deno & Fuchs, 1987; Deno, Fuchs, Marston, & Shin, 2001; Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, Hamlett, & Stecker, 1991; Marston, 1989; see www.progressmonitoring.net).

The majority of CBM studies have focused on reading and pre-reading skills, however, written expression using CBM procedures has received increased attention. While variety of writing measures have been investigated over the past 20 years, selected measures have been designated as meeting CBM standards (McMaster & Espin, 2007). CBM writing indicators include number of words written, words spelled correctly, correct word sequences, and correct minus incorrect word sequences (Deno, Marston, & Mirkin, 1982; Deno, Mirkin, & Marston, 1980; Espin, Scierka, Skare, & Halverson, 1999; Espin, Shin, Deno, Skare,

Robinson, & Benner, 2000; Videen, Deno, & Marston, 1982). These scoring procedures have been demonstrated as reliable and valid as well as efficient and sensitive to students' academic progress for elementary, middle and secondary school students (Deno, et al., 1980; Deno, et al., 1982; Espin, et al., 1999; Espin, et al., 2000; Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002; Marston, 1982; Marston & Deno, 1981; Shinn, 1981; Tindal & Parker, 1989; Tindal & Parker, 1991; Videen, et al., 1982; Watkinson & Lee, 1992). More recently, CBM research has been extended to diverse groups, such as English Language Learners (Wiley, & Deno, 2005; Graves, Plasencia-Peinado, & Deno, 2005) and students with hearing loss at the elementary level (Chen, 2002).

Chen (2002) investigated the technical characteristics of CBM in written expression with elementary level students with hearing loss. Fifty-one third and fourth grade students who are deaf or hard of hearing participated in this study. All participants were asked to produce four 3-minute writing samples. The findings suggest that the number of total words written, words spelled correctly, words spelled incorrectly, different words used, correct word sequences, incorrect word sequences, and correct minus incorrect word sequences serve as technically robust measures in indexing third and fourth grade level students' growth in writing. Specifically, the number of words spelled correctly, the number of correct word sequences, and the total number of words written had relatively stronger correlations ($r > .80$) with criterion measures across two time points (i.e., winter and spring) compared with other writing variables.

Chen's study provides teachers of deaf or hard of hearing students with a promising formative approach to quickly and effectively evaluate the growth of written expression for students with hearing loss placed in third and fourth grade classes. However, we are not sure whether Chen's results would apply to high school level students with hearing loss. In addition, researchers using CBM have reported that the most robust scoring procedures (e.g.

total number of words written, total number of words spelled correctly) used at the elementary level to predict student performance must be reviewed for use at the secondary level. For example, Espin et. al. (1999, 2000) found that the technical characteristics of simple measures (e.g., total words written, words spelled correctly, three minute writing sample) did not prove to be technically adequate indicators of secondary students' written expression when compared with more complex measures (e.g., correct word sequences). This finding corroborates their hypothesis that as students get older, their writing becomes more complex. Similarly, we also hypothesize that the process of written expression would become more complex for secondary-school students who are deaf or hard of hearing, and the requirement of written expression for secondary-school students would no longer just focus on word spelling or conventions, but emphasize more advanced linguistic skills. We were interested in exploring the development of CBM written expression measures for high school level students with hearing loss. In addition to examining existing CBM indicators, we investigated the use of alternative indicators to specifically address the unique characteristics in written expression for students with hearing loss. These potential indicators involve variables considering semantic diversity of words and clauses (e.g. number of different words), syntax (e.g. number of subject-verb agreements), and the use of word knowledge (e.g., number of morphemes used correctly).

The purpose of this study was to examine the technical characteristics of various scoring procedures used to evaluate written expression of secondary- school students who are deaf or hard of hearing. Specifically, we examined the reliability and validity of 17 different written expression scoring procedures. Two research questions were addressed in this study:

1. What is the alternative-form reliability of various scoring procedures in written expression for secondary-school students who are deaf or hard of hearing?

2. What is the criterion-related validity of various scoring procedures in written expression for secondary-school students who are deaf or hard of hearing?

Method

Participants

Participants in the study were 22 students who are deaf or hard of hearing. All of the students were attending a special summer program that provided intensive instruction in the areas of reading, writing, and mathematics in preparation for the state basic skills tests. Students were residents of a Midwestern state and represented 18 different school districts. Twelve of the students attended specialized programs for deaf and hard of hearing students while the remaining 10 students participated in general education classroom instruction with support from teachers certified in deaf education and or speech-language pathologists. Students were identified in grade levels that ranged from 7th through 12th. Fourteen males and 8 females participated in this study, with a mean age of 15.59 years (range of 12 to 20 years). The majority of the students were classified as Caucasian ($n = 18$). Two students were African American, and two students were Hispanic American. Fourteen of the students were identified on their IEPs as having severe to profound hearing loss, with 8 students as having moderate to severe hearing loss. Fifteen students were identified by their teachers as having additional disabilities, and the remaining students ($n = 7$) as having no additional disabilities.

Materials

The written expression materials included four 3-minute writing tasks with each task using different prompts. The writing prompts used in this study were replicated from a study designed to investigate the reliability and validity of writing measures with hearing students (McMaster & Campbell, 2005). The writing prompts were carefully selected and intended to be free from cultural bias and appropriate for students with a wide range of age and skill

levels, as well as English Language Learners (ELL), students with language and learning disabilities, and students who were deaf or hard of hearing.

Narrative story starters. The narrative prompts used to elicit a story, were “On my way home from school, a very exciting thing happened....” and “One day, we were playing outside the school and....” Each narrative story starter was printed at the top of a sheet of lined paper. Each student was provided with two sheets of lined paper for writing their stories.

Picture prompts. The picture prompts included (1) a picture of students playing ball outside of a school, and (2) a picture of students boarding a bus outside of a school. For the picture with students playing ball, the corresponding written prompt was “Write about a game that you would like to play”, printed at the top of a sheet of paper, and followed by lines printed on the same sheet. For the picture with students boarding a bus, the corresponding written prompt was “Write about a trip you would like to take with the students in your class.” The corresponding prompt was printed at the top of a sheet of lined paper.

Scoring Procedures

The writing samples produced by students with hearing loss were scored using the standard procedures of existing CBM and alternative scoring procedures. These procedures were selected based on previous research on CBM written expression (see McMaster & Espin, 2007). Additional scoring procedures were developed by the researchers based on a review of the literature focusing on language assessment and written expression of students who are deaf or hard of hearing (Albertini & Schley, 2003; Albertini, Bochner, Dowaliby & Henderson, 1996; White, 2007).

Existing CBM scoring procedures:

(A word is defined as any series of letters separated from another series of letters by a space.)

Total words written (TWW). The number of words written in the sample.

Words spelled correctly (WSC). The number of correctly spelled English words in the sample.

Words spelled incorrectly (IWS). The number of incorrectly spelled English words in the sample.

Correct word sequences (CWS). The number of two adjacent, correctly spelled words that are semantically and syntactically acceptable within the context of the sentence to a native English speaker.

Incorrect word sequences (IWS). The number of two adjacent words that are not semantically or syntactically acceptable within the context of the sentence to a native English speaker.

Correct minus incorrect word sequences (CMIWS). The result of subtracting the number of correct word sequences minus the number of incorrect word sequences.

Alternative scoring procedures:

Correct subject-verb agreements (C-SVA). The number of correct occurrences of subject-verb agreements in the sample (e.g. Singular subjects require singular form of verbs; plural subjects require plural form of verbs).

Incorrect subject-verb agreements (I-SVA). The number of incorrect occurrences of subject-verb agreements in the sample.

Total number of subject-verb agreements (T-SVA). The number of correct subject – verb agreements plus the number of incorrect subject-verb agreements written in the sample.

Different words (DW). The number of different words written in the sample (e.g. .

Type-token ratio (TTR). The number of different words divided by the total number of words written in the sample.

Correct clauses (CC). The number of grammatically and semantically correct clauses including independent and dependent clauses in the sample. (A clause is a group of related

words that has both a subject and a verb, usually forming part of a compound or complex sentence).

Incorrect clauses (IC). The number of incorrect clauses in the sample, including independent and dependent clauses.

Total number of clauses (TC). The number of correct plus incorrect clauses in the sample.

Correct morphemes (CM). The number of morphemes used correctly within the context of the sample. (A morpheme is defined as the smallest unit of grammatical structure.)

Incorrect morphemes (IM). The number of morphemes used incorrectly within the context of the sample. .

Total number of morphemes (TM). The number of correct morphemes plus the number of incorrect morphemes in the sample.

Criterion Variables

Teacher ratings of student writing proficiency. The classroom teachers were asked to rate each student's writing skills on a 5- point Likert-type scale (1 = less skilled, 3 = averaged skilled, 5 = more skilled) when compared to their peer group. The teachers were encouraged to use the full range of ratings to describe their students.

Spontaneous Written Story of TOWL -3. The *Test of Written Language- 3rd edition* (TOWL-3: Hammill & Larsen, 1996) is a comprehensive test for evaluating written language designed for students from 7 years to 17 years 11 months of age. It is divided into two formats: 1) spontaneous written language samples, and 2) contrived samples of written expression. In the spontaneous portion of the test, students are directed to look at a stimulus picture and develop a story about the picture. The writing sample is scored based on English language conventions, language use, and story construction. In the contrived subtests,

students are required to either develop sentences based on target words or concepts, or correct sentences that contain semantic errors.

For the purpose of this study, only the spontaneous written story subtest from Form B was used. The directions for administration provided in the administration manual were signed and spoken. Students had 15 minutes to write a story in response to a black and white stimulus picture with a futuristic scene of space ships, astronauts, outer space, and so on.

Procedures

Probe administration. All written expression samples, including the TOWL-3 and the 4 three -minute picture/story starter probe samples were administered over a period of one week during the summer school session. The classroom teachers administered the assessments following specified procedures. Each day the students were asked to complete one of the five writing tasks. Teachers adapted the planned administration schedule to meet the demands of other activities (e.g. a field trip, class schedule changes). The order of administration was counterbalanced across classes and sessions.

Each teacher was provided with a writing packet accompanied by the directions for the day and asked to complete the activity with the students and return the written language samples to the school office by the end of the day. The classroom teachers read the directions orally and signed the directions to the students. The students were given one minute to think about what they wanted to write and then prompted to begin writing. The students were given three minutes to write each of the four picture/ story starter writing probes and 15 minutes to write in response to the TOWL-3 picture probe.

Scorer training. The primary investigator designed a training protocol and rubric for each method of scoring and met individually with the scorers in two-hour training sessions. Training included a description of the scoring tasks, demonstration, and practice with the scoring procedures. In each session, the researcher reviewed the rules for scoring with the

scorers, and practiced using a writing sample. Then, each scorer practiced using a second sample with the researcher's guided assistance. Each scorer then scored 5 writing samples independently. The researcher then compared each set of raters' scores with her own scores, and calculated the inter-rater agreement for the scoring procedures. Any scorer whose scores had less than 90% agreement was asked to meet with the researcher for further training. Once 90% agreement was attained, the scorers were given a set of writing samples to score.

The TOWL-3 writing samples were scored by a different set of trained scorers familiar with TOWL-3 scoring procedures. Inter-rater agreements for the TOWL-3 were calculated by dividing the lower score by the higher score and multiplying by 100. Inter-rater agreement of TOWL-3 averaged 82.2% and ranged from 45% to 96.8%. Inter-rater agreements below 9 were scored by a third trained scorer. Scores resulting from the highest inter-rater agreements were used in the final analysis.

Data Analyses

The primary research questions addressed the alternative-form (picture versus story starter prompt) reliability and criterion-related validity of scoring procedures. To investigate the alternate-form reliability for each scoring procedure, bivariate Pearson product-moment correlation coefficients were computed between the scores obtained from the two story starter samples or two picture prompt samples. To address whether the alternate-form reliability coefficients differed between the different writing prompts, the correlations were compared after transforming the correlation coefficients to Fisher's r' statistic (Fisher, 1921). After the transformation, the z test statistic was used to compare each statistically significant correlation.

To examine the criterion-related validity evidence, Pearson product-moment correlation coefficients were calculated to determine the relationship between the mean scores derived from the probes administered to students and the TOWL-3 raw scores.

Spearman rank-order correlation coefficients were computed to examine the relationship between the mean scores from each probe and teacher ratings. To determine whether significant validity differences existed between the correlations by types of writing prompts, the correlation coefficients were transformed using Fisher's r' statistic (Fisher, 1921).

Results

Inter-rater Agreement

To examine the inter-rater agreement of the CBM scoring procedures in written expression, 30% of the packets were randomly selected from the total sample and scored by 6 different raters and the researchers. Inter-rater agreement was calculated by dividing agreements by agreements plus disagreements and multiplying by 100.

Agreements for each scoring procedure across all probes are included in Table 1. The inter-rater agreements ranged from 68.8% to 100%. With the exception of two, all inter-rater and most of the scoring procedures were above 90% (i.e., 24 out of 34). The inter-rater agreements attaining 100% across writing prompts were: number of different words written (DW), number of correct subject-verb agreements (C-SVA), the total number of subject-verb agreements (T-SVA), and the total number of clauses (TC).

 Insert Table 1 about here

Descriptive Statistics

Means and standard deviations are reported for story starters and picture prompts by scoring procedures in Table 2. For the story starter probes, the deaf and hard of hearing students generally received higher scores in response to story starter 1 than to story starter 2 as measured by TWW, WSC, CWS, CMIWS, DW, C-SVA, CC, TC, CM, and TM. The pattern of picture prompts was similar to that of the story starters. Students scored higher in

response to picture prompt 1 than to picture prompt 2 as measured by TWW, WSC, CWS, IWS, CMIWS, DW, C-SVA, CM, and TM. Generally, students' scores were higher in response to the story starters than to the picture prompts as measured by TWW, WSC, CWS, IWS, C-SVA, T-SVA, CC, TC, CM, and TM.

Insert Table 2 about here

Inter-rater Reliability

The measurement error associated with the scoring processes is of particular of interest in this study. In order to explore how clearly and unequivocally the scoring guidelines assist scorers in evaluating writing samples, intra-class correlations (ICC) were used to measure inter-rater reliability. Intra-class correlations may be conceptualized as the ratio of rating variance to total variance. A two-way random effects model was used to investigate inter-rater reliability in this study in that the scorers were conceived as a random selection from among all possible scorers, and target writing samples were chosen at random from a pool of writing samples. Two-way random effects can be defined as a random sample of k scorers selected from a large population, where each scorer rates each writing sample, that is, each scorer rates n target writing samples altogether (Shrout & Fleiss, 1979).

Table 3 summarizes the intra-class correlation coefficients for each scoring procedure for students who are deaf or hard of hearing. With the exception of WSI, TWW, and IM, the scoring procedures produced intra-class correlations above .80, indicating that approximately 80% or higher observed variance is true variance for these scoring procedures. C-SVA, I-SVA, and T-SVA had the highest intra-class correlation coefficients across all four of the writing tasks.

Insert Table 3 about here

Alternate-form Reliability

In general, the alternate-form reliability coefficients of picture prompts were slightly higher than those of story starters. Type-token ratio (TTR) was not significant across writing tasks. The scoring procedures counting the numbers of incorrect word pairs (IWS), subject-verb agreements (I-SVA), clauses (IC), and morphemes (IM) for story starter tasks were not statistically significant. For picture prompts, WSI and I-SVA were not statistically significant.

For story starters, the currently used scoring procedure with the highest reliability coefficient was CWS ($r = .745$); whereas the alternative scoring procedure with the highest reliability coefficient was C-SVA ($r = .733$). For picture prompts, CMIWS had the highest reliability coefficient ($r = .895$) among currently used CBM scoring procedures, whereas, TC had the highest reliability coefficient of $r = .803$ among the alternative scoring procedures.

To test the significance between the two types of writing prompts, each correlation was transformed to r' . The transformed correlation coefficients were compared across writing prompts using the z test statistic (Fisher, 1921). Only the IC was found to differ significantly between the two types of writing prompts ($z = -2.52$). No reliable differences were found in other scoring procedures.

Insert Table 4 about here

Criterion-related Validity

Criterion-related validity was investigated by examining the relationships between the scoring procedures and criterion measures, including TOWL-3 and teacher ratings. The mean

of the two writing tasks within each prompt (i.e., story starter and picture prompt) was calculated for the correlation analysis.

Table 5 displays the aggregated means and standard deviations for currently used and alternative CBM scoring procedures, including TWW, WSC, WSI, CWS, IWS, CMIWS, DW, TTR, C-SVA, I-SVA, T-SVA, CC, IC, TC, CM, IM, and TM.

 Insert Table 5 about here

Correlation of CBM current scoring procedures with criterion test scores. Pearson correlations were calculated between scores from the writing tasks administered to the students and the score of the subtest of TOWL-3. Across two writing prompts, the strongest correlations were observed for CWS and CMIWS. The weakest correlations were WSI, TTR, I-SVA, IC, and IM for story starters. For picture prompts, the weakest correlations observed were TTR, IC, WSI, TC, I-SVA, and IM. Other scoring procedures were moderately to strongly correlated with the score of the subtest of TOWL-3.

 Insert Table 6 about here

Correlation of CBM current scoring procedures with teacher ranking of student writing proficiency. Spearman rank-order correlation coefficients were calculated between scores from the probes administered to the students and the teacher ratings of student writing skills. Scoring procedures that were moderately to moderate-strongly correlated with the teacher ratings include CWS, CMIWS, DW, C-SVA, T-SVA, CC, TC, CM, and TM. The highest correlations were CWS ($r = .696$ to $.729$, $p < .01$) and CMIWS ($r = .666$ to $.819$, $p < .01$). Among the alternative scoring procedures, T-SVA was strongly correlated with

teacher ratings for story starters; whereas, CC had the highest correlation coefficient with teacher ratings for picture prompts.

Insert Table 7 about here

To investigate the significance between two types of writing prompts on criterion measures (TOWL-3 and teacher ratings), correlations between the scoring procedures and the criterion measures were transformed to r' . The transformed correlation coefficients were compared across writing prompts using the z test statistic (Fisher, 1921). Again, an alpha value of .05 was adopted; thus, z values greater than 1.96 or less than -1.96 were considered as the index of significant differences in the strength of the correlations. No reliable differences were found.

Discussion

The purpose of this study was to examine currently used CBM written expression scoring procedures and alternate scoring procedures with secondary level students who are deaf or hard of hearing. In addressing this question, the alternate-form reliability and criterion-related validity of various scoring procedures were examined.

Prior to addressing the technical characteristics of seventeen different scoring procedures, we would like to mention our pilot study. In our preliminary analyses (inter-rater agreement and inter-rater reliability), we wanted to identify measures that could be efficiently, consistently and frequently administered and scored. First, the results of examination of inter-rater agreement suggest that many of current CBM scoring procedures used were scored reliably following training. With the exception of two, the inter-rater agreement percentages were above 90%. The agreement coefficient of correct minus incorrect word sequences was the lowest for both writing prompts. A possible explanation for the lower coefficients is that

correct minus incorrect word sequences is derived from correct word sequences minus incorrect word sequences. If there is a difference between scorers in both, though it is small and within the criteria of judging agreement, it would penalize the correct minus correct word sequence score because the difference would be doubled rather than neutralized, making the difference exceed the criteria of judging agreement. If this occurs several times, the average inter-rater agreement decreases. Second, inter-rater reliability examined by intra-class correlations strongly explained the true scores for most of the scoring procedures. All additional CBM scoring procedures were highly reliable. The results revealed that the scoring guidelines for alternative scoring procedures are explicit and clear for scorers to understand and apply. This reduces the rater's measurement error, and increases the opportunity of observed variance close to the true variance.

Alternate-form Reliability

Differences were found in the alternate-form reliabilities for the various CBM scoring procedures examined in this study. Alternate-form reliability coefficients were somewhat lower than desirable when using story starters as writing prompts for five scoring procedures: incorrect word sequences, type-token ratio, incorrect subject-verb agreements, incorrect clauses, and incorrect morphemes. The highest reliability coefficients for story starter prompts were found for correct word sequences correct minus incorrect word sequences, correct subject-verb agreements, total subject-verb agreements, correct clauses, and correct morphemes. When using picture prompts as stimuli for writing, incorrect word sequences, type-token ratio, and incorrect subject-verb agreements were unacceptably low. When using picture prompts, the most consistent reliability coefficients were found for correct word sequence, correct minus incorrect word sequence, correct subject-verb agreement, correct clauses, incorrect clauses, total number of clauses and correct morphemes. Generally, alternate-form reliabilities were higher for picture prompts than for story starters. The

strongest and most consistent reliability coefficients were observed for the CBM indicators, correct word sequences and correct minus incorrect word sequences, across the two types of stimulus prompts.

The alternate-form reliability results in this study are similar to those conducted with middle school hearing students. Researchers studying CBM written expression procedures suggest that correct word sequences and correct minus incorrect word sequences are reliable measures in indexing students' general proficiency in written expression at the secondary level (McMaster & Espin, 2007; Espin, et al., 2000). It is worthwhile to mention that alternative scoring procedures for correct subject-verb agreements, correct clauses, and correct morphemes show the most promise for serving as additional indices for written expression.

Story starters generated lower alternate-form reliability coefficients than picture prompts. The low reliability for counting the number of incorrectly spelled words word pairs (IWS), subject-verb agreements incorrect clauses, and incorrect morphemes in the story starter probes may be due to topic sensitivity. We found that the means of these scoring procedures in story starter 2 were all slightly larger than those in story starter 1 and the standard deviations of these scoring procedures in story starter 2 were more variable than those in story starter 1. The purpose of CBM is to gather stable samples of student performance over time to evaluate student growth; this variability from topic to topic makes these scoring procedures inappropriate to use (Espin, et al., 2000).

It is worth noting that our reliability coefficients for each prompt were limited to two, 3 minute written expression samples. Fuchs, Deno, and Marston (1983) found that in order to obtain a reliability coefficient above .70, aggregation across four writing samples is necessary. In addition, although the alternate-form reliability coefficients ranged from moderate to moderately-strong due to the small sample size, most of the scoring procedures achieved

statistical significance. The results suggest that applying CBM scoring procedures in indexing performance of written expression for students with hearing loss is promising.

Criterion-related Validity Evidence

The pattern of criterion-related validity on the subtest of TOWL-3 revealed that the CBM scoring procedures had slightly higher correlations for the story starters than for the picture prompts. Among all scoring procedures, correct minus incorrect word sequences serves as the best indicator of writing performance for secondary-school students with hearing loss. Correct clauses produced the highest correlation coefficient with TOWL-3 among alternative scoring procedures. However, the validity coefficients on teacher ratings were complex and dependent on the character of scoring procedures. For existing CBM scoring procedures, validity coefficients of the scoring procedures for counting words (i.e., total words written, words spelled correctly, words spelled incorrectly) were stronger for story starters than for picture prompts, whereas the validity coefficients of the scoring procedures considering grammar and writing conventions (e.g., correct word sequences, incorrect word sequences, correct minus in correct word sequences) were stronger for picture prompts than for story starters. Correct word sequences and correct word minus incorrect word sequences had the strongest correlation with the criterion measures (i.e., TOWL-3 and teacher ratings) across writing prompts, indicating that these two scoring procedures appeared sufficiently robust to corroborate their use with students with hearing loss. The criterion related validity findings of this study are similar to research conducted with hearing students (Deno, et. al., 1983;Espin et al., 1999; Gansle, et al., 2004; Malecki & Jewell, 2003).

Several indicators relate well to teacher ratings across different types of writing prompts: number of different words, number of correct subject-verb agreements, number of correct clauses, number of correct morphemes, and total number of morphemes. Story starters produced higher coefficients with TOWL-3 and teacher ratings than picture prompts. Correct

clauses had the strongest relationship with the subtest of TOWL-3 and teacher ratings. Generally, the validity coefficients with teacher ratings were relatively lower than those with the TOWL-3. One possible explanation for lower correlations with teacher ratings involves teachers' familiarity with students. Recall that all of the participants were recruited from a summer school program; several of the students were from different school programs. Therefore, the teachers weren't as familiar with these students. This may lead teachers to assign unreliable ratings to students' writing performances based on limited information. In this case, the teacher ratings may be less useful as a criterion measure.

Regarding scoring efficiency, most of the scoring procedures that are technically adequate and are promising indicators for use with deaf and hard of hearing students were time-consuming, including correct word sequences, correct minus incorrect word sequences, different words, correct subject-verb agreements, correct clauses, correct morphemes, and total morphemes. These indicators require scorers to carefully follow the relatively more complex scoring directions compared with selected CBM scoring procedures, such as total words written and words spelled correctly. Thus, teachers may need to consider a trade-off in which accuracy is gained but efficiency is lost or vice versa when selecting procedures for indexing student performance in written expression.

With regard to suitability of CBM writing indicators, our findings show similar patterns with previous research done with hearing students. The best indicators for students at the elementary level are not necessarily the best indicators for students at the secondary level (Deno, et al., 1982; Espin, et al., 1999; Espin, et al., 2000, Tindal & Parker, 1989). Total words written and words spelled correctly have been reported as the most technically adequate indicators for student with hearing loss at the elementary level (Chen, 2002). However, these measures did not prove to be strong scoring metrics of writing performance for students with hearing loss at the secondary level in our study. Based on these results, we

assume that as students with hearing loss progress in their written English expression, more complex scoring procedures that incorporate grammar, syntax, semantics, and written expression conventions are better candidates of predicting writing performance.

Differences Related to Types of Writing Prompts

Our data showed few differences between the two types of writing prompts in terms of technical adequacy. With the exception of incorrect clauses, all scoring procedures showed quite small differences and were not statistically significant across story starter writing and picture prompt writing samples. Our data reveal that by using most of the CBM indicators to score written expression for students with hearing loss, stimuli used in this study are basically interchangeable. For criterion-related validity evidence, the results revealed that no differences were found between two types of writing prompts across all scoring procedures. The relationship between all scoring procedures and criterion measures (i.e., TOWL-3 and teacher ratings) indicates that the use of two different types of writing prompts provides indirect evidence of measuring the same construct.

Limitations

Several limitations restrict the generalization of this study and should be discussed and interpreted with caution. The first limitation of this study is related to the participants. Although the deaf or hard of hearing populations across the country are heterogeneous, the participants in this study were mainly Caucasian. Thus, the results may not be applicable to other culturally diverse groups of students with hearing loss.

Another limitation of the study was the small number of participants. Due to the small sample size, it may be inappropriate to generalize the findings to all students with hearing loss. In addition, generalizations of the findings from this study should be limited to secondary-school students with hearing loss. Further research at various grade levels is

needed to provide sufficient evidence of the technical adequacy of scoring procedures at each grade level for students with hearing loss.

A further limitation was that the criterion measures used in this study did not accurately or fully represent general writing proficiency. For example, one of the criterion measures was teacher ratings. Teachers of deaf or hard of hearing children were not equally acquainted with all students in their classrooms due to the nature of the summer school program. This may result in assigning unreliable ratings to students' general writing performance. It may have served to decrease the correlations between scoring procedures and teacher ratings.

Further, the investigation of reliability and validity of the scoring procedures examined in this study did not involve exploring the sensitivity of the writing measures as indicators of growth in the progress monitoring process. The findings of this study should be restricted from this perspective.

A final limitation involves the fidelity of the administration procedures. Data collection was conducted as part of the routine within the summer program and done by the classroom teachers of deaf or hard of hearing children, and no available information was provided to substantiate the fidelity of the administration protocols.

Implications for Practice and Future Research

Several educational implications can be obtained from this study. First, teachers of deaf or hard of hearing children can have confidence in using CBM scoring procedures including total words written, words spelled correctly, correct word sequences, incorrect word sequences, and correct minus incorrect word sequences. Unlike traditional standardized tests, CBM scoring procedures can be incorporated by teachers of deaf or hard of hearing children to evaluate the effectiveness of daily instruction, and program evaluation. In addition, these scoring procedures are simple to use, easy to interpret results, efficient both monetarily

and time wise, may reduce the burden of testing, and increase teachers' attention to routinely monitoring the progress of written expression among students with hearing loss.

Generally, this study demonstrated that the two types of writing prompts (i.e., story starters and picture prompts) were interchangeable for students with hearing loss. Teachers may use either story starters or picture prompts.

Results from this study suggest the need for additional research. Research is needed to determine the sensitivity and the technical characteristics of these scoring procedures in a longitudinal perspective as a means of detecting progress in written expression. Our findings reveal that the additional alternative writing indicators serve as promising potentials for CBM practice, further technical adequacy studies with different perspectives (e.g., different grade levels, sample length, other writing prompts, different cultures, hearing students, and/or English Language learners) are needed to determine whether the alternative scoring procedures are feasible and sustainable across diverse groups, and a variety of writing variables.

Since there is a substantial literature base documenting the unique needs of students with hearing loss, future research may identify the effect of using alternative scoring procedures to determine desirable benchmarks or proficiencies for each grade level. Ultimately, future studies regarding effectiveness of interventions using evidence-based scoring procedures that can be used at short term intervals (e.g. weekly, monthly) and that can inform instruction will have a significant impact on increasing written expression among students who are deaf or hard of hearing.

References

- Albertini, J.A., & Schley, S. (2003). Writing: Characteristics, instruction, and Assessment. In M. Marschark, P.E. Spencer, Deaf Studies, Language and Education, NY: Oxford University Press, 123-135.
- Chen, Y-C. (2002). *Assessment of reading and writing samples of deaf and hard of hearing students by curriculum-based measurements*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, Minnesota.
- Costello, E. (1977). Continuing education for deaf adults: A national needs assessment. *American Annals of the Deaf*, 122, 26-32.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S.L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37(3), 184-192.
- Deno, S.L., & Fuchs, L.S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, 19(8), 1-16.
- Deno, S.L., Marston, D., Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children Special Education and Pediatrics: A New Relationship*, 48, 368-371.
- Deno, S.L., Fuchs, L.S., Marston, D.B., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30, 507-524.
- Deno, S.L., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children*, 48, 368-371.

- Deno, S.L., Mirkin, P.K., & Marston, D. (1980). *Relationships among simple measures of written expression and performance on standardized achievement test* (Research Report No. 22). Minneapolis, MN: University of Minnesota Institute for Research on Learning Disabilities.
- Espin, C.A., Scierka, B. J., Skare, S., & Halverson, N. (1999). Criterion-related validity of curriculum-based measures in writing for secondary students. *Reading and Writing Quarterly*, 15(1), 23-28.
- Espin, C.A., Shin, J., Deno, S.L., Skare, S., Robinson, S., & Benner, B. (2000). Identifying indicators of written expression proficiency for middle school students. *The Journal of Special Education*, 34(3), 140-153.
- Fuchs, L.S., Deno, S.L., & Marston, D. (1983). Improving the reliability of curriculum-based measures of academic skills for psychoeducational decision making. *Diagnostique*, 8, 135-149.
- Fuchs, L.S., Deno, S.L., & Mirkin, P. (1984). Effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, 21, 449-460.
- Fuchs, L.S., Fuchs, D., Hamlett, C.L., & Stecker, P.M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28, 617-641.
- Gansle, K.A., Noell, G.H., VanDerHeyden, A.M., Naquin, G.M., & Slider, N.J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternate measures for curriculum-based measurement in writing. *School Psychology Review*, 31(4), 477-497.

- Graves, A.W., Plasencia-Peinado, & Deno, S.L. (2005). Formatively evaluating the reading progress of first-grade English learners in multiple-language classrooms. *Remedial and Special Education, 26*(4), 215-225.
- Hammill, D.D., & Larsen, S.C. (1996). *Test of Written Language* (3rd ed.). Austin, TX: Pro-Ed.
- Ivimey, G.P., & Lachterman, D. H. (1980). The written language of young English deaf children. *Language and Speech, 23*, 351-378.
- Malecki, C.K., & Jewell, J. (2003). Developmental, gender, and practical considerations in scoring curriculum-based measurement writing probes. *Psychology in the Schools, 40*(4), 379-390.
- Marschark, M., Lang, H.G., & Albertini, J.A. (2002). *Educating deaf students: From research to practice*. New York: Oxford University Press.
- Marschark, M., Mouradian, V., & Halas, M. (1994). Discourse rules in the language productions of deaf and hearing children. *Journal of Experimental Child Psychology, 57*, 89-107.
- Marston, D. (1982). *The technical adequacy of direct, repeated measurement of academic skills in low-achieving elementary students*. Unpublished doctoral dissertation, Minneapolis, MN: University of Minnesota.
- Marston, D. (1989). Curriculum-based measurement: What it is and why do it? In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp.18-78). New York: Guilford Press.
- Marston, D., & Deno, S.L. (1981). *The reliability of simple, direct measures of written expression* (Research Report No. 50). Minneapolis, MN: University of Minnesota, Institute for Research on Learning Disabilities.

- McMaster, K., & Campbell, H. (2005). *Technical Features of New and Existing Measures of Written Expression: An Examination Within and Across Grade Levels* (Research Report No. 9). Minneapolis, MN: University of Minnesota Research Institute on Progress Monitoring.
- McMaster, K. & Espin, C. (2007). Technical features of Curriculum-based measurement in writing: A literature review. *The Journal of Special Education, 41*, 68-84.
- Paul, P. V. (2001). *Language and deafness*. San Diego, CA: Singular.
- Rose, S., McAnally, P.L., & Quigley, S.P. (2004). *Language learning practices with deaf children* (3rd ed.). Austin, TX: PRO-ED.
- Shinn, M.R. (1981). *A comparison of psychometric and functional differences between students labeled learning disabled and low achieving*. Unpublished doctoral dissertation, Minneapolis, MN: University of Minnesota.
- Singleton, J.L., Morgan, D., DiGello, E., Wiles, J., & Rivers, R. (2004). Vocabulary use by low, moderate, and high ASL-proficient writers compared to hearing ESL and monolingual speakers. *Journal of Deaf Studies and Deaf Education, 9*, 86-103.
- Shrout, P.E., Fleiss, J.L. (1979). Intraclass correlations: Use in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.
- Tindal, G.A., & Parker, R. (1989). Assessment of written expression for students in compensatory and special education programs. *The Journal of Special Education, 23*(2), 169-183.
- Tindal, G.A., & Parker, R. (1991). Identifying measures for evaluating written expression. *Learning Disabilities Research and Practice, 6*, 211-218.
- Videen, J., Deno, S.L., & Marston, D. (1982). *Correct word sequences: A valid indicator of writing proficiency in written expression* (Research Report No. 84). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

- Watkinson, J.T., & Lee, S.W. (1992). Curriculum-based measures of written expression for learning-disabled and nondisabled students. *Psychology in the Schools, 29*, 184-191.
- White, A.H. (2007). A toll for monitoring the development of written English: T-Unit Analysis using the SAWL. *American Annals of the Deaf, 152*, 29-41.
- Wiley, H.I., & Deno, S.L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessments. *Remedial and Special Education, 26(4)*, 207-214.
- Yoshinaga-Itano, C., & Snyder, L. (1985). Form and meaning in the written language of hearing-impaired children. *The Volta Review, 87*, 75-90.
- Yoshinaga-Itano, C., Snyder, L., & Mayberry, R. (1996a). How deaf and normally hearing students convey meaning within and between written sentences. *The Volta Review, 98*, 9-38.
- Yoshinaga-Itano, C., Snyder, L., & Mayberry, R. (1996b). Can lexical/semantic skills differentiate deaf or hard of hearing readers and nonreaders? *The Volta Review, 98*, 39-61.

Table 1

Inter-rater Agreement for Scoring Procedures in Written Expression for Different Prompts
(*N* = 22)

| <i>Scoring Procedures</i> | <i>Inter-rater Agreement</i> | |
|---|------------------------------|-----------------------|
| | <i>Story Starter</i> | <i>Picture Prompt</i> |
| Existing CBM Scoring Procedures | | |
| TWW | 93.8% | 100.0% |
| WSC | 93.8% | 93.8% |
| WSI | 100.0% | 93.8% |
| CWS | 87.5% | 87.5% |
| IWS | 87.5% | 87.5% |
| CMIWS | 81.3% | 68.8% |
| Alternative CBM Scoring Procedures | | |
| DW | 100.0% | 100.0% |
| TTR | 96.0% | 100.0% |
| C-SVA | 100.0% | 100.0% |
| I-SVA | 100.0% | 93.8% |
| T-SVA | 100.0% | 100.0% |
| CC | 93.8% | 100.0% |
| IC | 93.8% | 93.8% |
| TC | 100.0% | 100.0% |
| CM | 75.0% | 93.8% |
| IM | 81.3% | 93.8% |
| TM | 87.5% | 87.5% |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA= correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

Table 2

Means and Standard Deviations for Each Scoring Procedure across All Probes for Students with Hearing Loss (N = 22)

| CBM Scoring Procedures | Story Starter 1 | | Story Starter 2 | | Picture Prompt 1 | | Picture Prompt 2 | |
|--|-----------------|-------|-----------------|-------|------------------|-------|------------------|-------|
| | M | SD | M | SD | M | SD | M | SD |
| Existing CBM Scoring Procedures | | | | | | | | |
| Total Words Written | 51 | 16.87 | 46.68 | 18.53 | 49.18 | 15.79 | 43.95 | 16.07 |
| Words Spelled Correctly | 50.23 | 16.53 | 45.86 | 18.59 | 48.18 | 15.99 | 43.43 | 16.31 |
| Words Spelled Incorrectly | .77 | .973 | .82 | .80 | 1.00 | .98 | .52 | .60 |
| Correct Word Sequences | 41.27 | 20.22 | 35.36 | 22.67 | 40.05 | 20.39 | 34.90 | 15.59 |
| Incorrect Word Sequences | 15.23 | 7.46 | 17.00 | 10.07 | 14.91 | 6.52 | 14.19 | 8.68 |
| Correct minus Incorrect Word Sequences | 26.05 | 24.31 | 18.36 | 28.56 | 25.14 | 24.35 | 20.17 | 18.15 |
| Alternative Scoring Procedures | | | | | | | | |
| Different Words | 34.77 | 9.89 | 34.82 | 13.04 | 36.18 | 10.30 | 29.67 | 9.80 |
| Type-token Ratio | .7 | .10 | .76 | .07 | .74 | .08 | .70 | .10 |
| Correct Subject-verb Agreements | 6.41 | 3.74 | 5.5 | 3.56 | 4.91 | 3.22 | 4.57 | 2.46 |
| Incorrect Subject-verb Agreements | 1.0 | 1.16 | 1.86 | 2.21 | 1.05 | 1.17 | 1.71 | 1.52 |
| Total Subject-verb Agreements | 7.4 | 3.84 | 7.41 | 3.76 | 5.95 | 3.09 | 6.29 | 2.88 |
| Correct Clauses | 3.09 | 2.88 | 2.59 | 3.13 | 1.86 | 1.89 | 1.90 | 1.76 |
| Incorrect Clauses | 4.27 | 1.78 | 4.36 | 2.04 | 4.18 | 1.68 | 4.48 | 2.73 |
| Total Clauses | 7.36 | 2.63 | 6.95 | 3.36 | 6.05 | 2.54 | 6.38 | 2.77 |
| Correct Morphemes | 55.82 | 19.59 | 51.23 | 20.73 | 52.36 | 18.57 | 46.95 | 17.00 |
| Incorrect Morphemes | 2.09 | 2.18 | 2.27 | 2.59 | 1.50 | 1.47 | 2.33 | 2.48 |
| Total Morphemes | 57.91 | 19.18 | 53.50 | 21.25 | 53.86 | 18.28 | 49.29 | 17.74 |

Table 3

Intra-class Correlations of CBM Scoring Procedures for Students with Hearing Loss (N = 22)

| Scoring Procedures | Inter-rater Reliability Coefficients | | | |
|--------------------|--------------------------------------|-----------------|------------------|------------------|
| | Story Starter 1 | Story Starter 2 | Picture Prompt 1 | Picture Prompt 2 |
| TWW | 1.000 | .999 | .999 | .750 |
| WSC | .999 | .999 | .999 | .998 |
| WSI | .851 | .741 | .920 | .197 |
| DW | .999 | .999 | .998 | .999 |
| TTR | .996 | .974 | .997 | .992 |
| CWS | .997 | .989 | .992 | .998 |
| IWS | .974 | .972 | .858 | .986 |
| CMIWS | .992 | .978 | .977 | .993 |
| C-SVA | 1.000 | .992 | 1.000 | .972 |
| I-SVA | 1.000 | .988 | 1.000 | .793 |
| T-SVA | 1.000 | .990 | 1.000 | .971 |
| CC | .867 | .977 | .964 | .982 |
| IC | .799 | .970 | .844 | .954 |
| TC | .985 | .982 | .967 | .958 |
| CM | .998 | .990 | .999 | .993 |
| IM | .945 | .712 | .865 | .862 |
| TM | .998 | .998 | .999 | .998 |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA= correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

Table 4

Alternate-Form Reliabilities for Writing Prompts: Students with Hearing Loss (N = 22)

| CBM Scoring Procedures | Story Starter | Picture Prompt |
|--|---------------|----------------|
| Currently Used Scoring Procedures | | |
| TWW | .574** | .659** |
| WSC | .563** | .651** |
| WSI | .437* | .211 |
| CWS | .745** | .834** |
| IWS | .321 | .625** |
| CMIWS | .729** | .895** |
| Alternative Scoring Procedures | | |
| DW | .606** | .558** |
| TTR | .401 | .182 |
| C-SVA | .733** | .718** |
| I-SVA | .056 | .254 |
| T-SVA | .697** | .612** |
| CC | .730** | .782** |
| IC | .182 | .767** |
| TC | .622** | .803** |
| CM | .645** | .663** |
| IM | .300 | .591** |
| TM | .589** | .631** |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA= correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

*p<.05, 2-tailed; **p<.01, 2-tailed

Table 5

Descriptive Statistics for Mean CBM Existing and Alternative Scoring Procedures: Students with Hearing Loss (N = 22)

| CBM Scoring Procedures | Story Starters | | Picture Prompts | |
|---------------------------------------|----------------|-------|-----------------|-------|
| | M | SD | M | SD |
| Existing Scoring Procedures | | | | |
| TWW | 48.84 | 15.70 | 45.81 | 13.91 |
| WSC | 48.05 | 15.53 | 45.02 | 14.03 |
| WSI | .80 | .75 | .79 | .62 |
| CWS | 38.32 | 20.04 | 36.45 | 16.29 |
| IWS | 16.11 | 7.16 | 14.79 | 6.76 |
| CMIWS | 22.20 | 24.59 | 21.67 | 19.46 |
| Alternative Scoring Procedures | | | | |
| DW | 34.80 | 10.30 | 32.52 | 8.66 |
| TTR | .73 | .07 | .72 | .07 |
| C-SVA | 5.95 | 3.39 | 4.67 | 2.64 |
| I-SVA | 1.43 | 1.28 | 1.38 | 1.08 |
| T-SVA | 7.41 | 3.50 | 6.05 | 2.68 |
| CC | 2.84 | 2.79 | 1.86 | 1.73 |
| IC | 4.32 | 1.47 | 4.33 | 2.10 |
| TC | 7.16 | 2.70 | 6.19 | 2.54 |
| CM | 53.52 | 18.29 | 48.79 | 15.54 |
| IM | 2.18 | 1.92 | 1.95 | 1.77 |
| TM | 55.70 | 18.03 | 50.74 | 15.65 |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA= correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

Table 6

*Correlations between CBM Existing and Alternative Scoring Procedures and TOWL-3:
Students with Hearing Loss (N = 22)*

| CBM Scoring Procedures | Story Starter | Picture Prompt |
|---------------------------------------|---------------|----------------|
| Existing Scoring Procedures | | |
| TWW | .668** | .437* |
| WSC | .666** | .443* |
| WSI | .180 | -.226 |
| CWS | .801** | .681** |
| IWS | -.544** | -.560** |
| CMIWS | .811** | .765** |
| Alternative Scoring Procedures | | |
| DW | .663** | .486* |
| TTR | -.267 | -.092 |
| C-SVA | .777** | .635** |
| I-SVA | -.074 | -.383 |
| T-SVA | .729** | .469* |
| CC | .795** | .679** |
| IC | -.230 | -.189 |
| TC | .696** | .306 |
| CM | .682** | .512* |
| IM | -.260 | -.413 |
| TM | .665** | .462* |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA=correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

* p<.05, 2-tailed; ** p<.01, 2-tailed

Table 7

Spearman Rank-order Correlations between Existing and Alternative CBM Scoring Procedures and Teacher Ratings: Students with Hearing Loss (N = 22)

| CBM Scoring Procedures | Story Starter | Picture Prompt |
|--------------------------------|---------------|----------------|
| Existing Scoring Procedures | | |
| TWW | .580** | .411 |
| WSC | .579** | .411 |
| WSI | -.090 | -.055 |
| CWS | .696** | .729** |
| IWS | -.381 | -.568** |
| CMIWS | .666** | .819** |
| Alternative Scoring Procedures | | |
| DW | .663** | .478* |
| TTR | -.153 | -.036 |
| C-SVA | .620** | .574* |
| I-SVA | -.198 | -.305 |
| T-SVA | .674** | .414 |
| CC | .571** | .689** |
| IC | -.356 | -.365 |
| TC | .470* | .348 |
| CM | .562** | .460* |
| IM | -.012 | -.039 |
| TM | .572** | .467* |

Note: TWW=total words written, WSC=words spelled correctly, WSI=words spelled incorrectly, CWS=correct word sequences, IWS=incorrect word sequences, CMIWS=correct minus incorrect word sequences, DW=different words, TTR=type-token ratio, C-SVA=correct subject-verb agreements, I-SVA=incorrect subject-verb agreements, T-SVA=total subject-verb agreements, CC=correct clauses, IC=incorrect clauses, TC=total clauses, CM=correct morphemes, IM=incorrect morphemes, TM=total morphemes

** p<.05, 2-tailed; * p<.01, 2-tailed